% see Manuri 2017; Knapp et al 2018 Table 1 for lidar metric definitions.

\section{Discussion}

Logical: It is desirable to develop models that have clear biological interpretation; the variables pools arrived at via our approach can be reasonably interpreted to have direct analogs to ecologically significant variables. Lidar height, density, and distribution metrics are analogous to variables used in aerial stand volume tables that are used in forest inventory. Also, the inclusion of the vertical distribution variable, in our case height skewness, may help the model to account for intermediate tree crown in the over-story and suppressed trees in the understory \cite{Li et al 2008}. Close agreement in magnitude between RMSE and RMSEcrossval suggests that the model is not overfit and suitable for generalization.

Issues influenced the analysis of this data set: 1) the potential for multicollinearity when a large number of independent variables are involved; 2) the choice of variables for, and the comprehensibility of, regression equations.

Latent variables – overfitting may occur, independent (validation) data used to determine the optimal number of latent variables (?? Maybe Naesset e tal 2005). Close agreement in magnitude between RMSE and RMSEcrossval suggests that the model is not overfit and suitable for generalization.

Regional Lit comparison: Lidar based regional models have been developed to estimate biomass in boreal, temperate deciduous, temperate coniferous, and tropical forests \cite{Naesset 2004, Nelson 2004,Lefsky 2005,Naesset 2008, Asner 2012}.

Lefsky et al. (2002) were successful at developing a unified equation for predicting aboveground biomass in multiple biomes (Temperate Deciduous Broadleaf, Temperate Coniferous Needleleaf, Boreal Coniferous Needleleaf), there was no replication in each biome. When all sites are considered together, mean height squared is the best overall predictor of above-ground biomass. (Lefsky et al 2002)

Lefsky et al 2005 A we were able to look at 5 sites within the Temperate Coniferous Needleleaf biome, using sites with varying environment and composition. They created equations that predicted stand structure variables (e.g. aboveground biomass and LAI) across an environmental and compositional gradient. Aboveground biomass were consistently predictable along a productivity and species composition gradient from the true fir forests of Mt. Rainier to Ponderosa pine forests at Metolius, and at the high productivity forests of Cascade Head, the coast range, and H.J. Andrews (Lefsky et al 2005 A).

(Lefsky et al 2005 A and 2002) offer a regional confirmation of the continental-scale hypothesis offered in Lefsky et al. (2002), in which the geographic generality of an equation predicting aboveground biomass was demonstrated.

Lefsky et al 2005 A found broad consistency in lidar-stand structure relationship over this region, and a relative lack of importance of environmental conditions (Lefsky et al 2005 A).

While the range of environmental conditions and composition examined in this paper is narrower than in Lefsky et al. (2002), the number of site locations examined is larger, and thus confirms the result for the Pacific Northwest region of the USA (Lefsky et al 2005 A).

Given this wide range of conditions, and the earlier results of Lefsky et al. (2002 and 2005 A), it is reasonable to ask if, in forests dominated by coniferous species, tree architecture is constrained to the point where a unified relationship between lidar measurements and stand structure might exist for these forests generally. In existing studies of this type (Lefsky et al. 2002 and 2005 A) there has been an attempt to have a structural or temporal sequence of stands at one or more study locations.

Density is mass/volume.

PIPO fit comparison: Weaker RMSE than those that have were built on small data sets covering a limited lidar footprint.

\subsection{model fit}

The models perform best in the central range of the data and tend to underpredict regions with high biomass and over predict regions with low biomass.

Iid - linear issues: Over predicted low values (Li et al 2008, Kim et al 2009)

Power relationship –

* Log(biomass) Lim and Treitz 2004, Naesset 2011, Shao et al 2018
* Log, log – Gobakken 2012
* Stephens uses P30^2 (Stephens 2012)

Height shape, L-moments (Latifi, Hernandez 2014, Hernadnez 2015, Gonzalez 2014, li 2014, Valbuena 2016, vega 2016)

Separate models by productivity (Gobakken 2012); Naesset also did: Commonly used stratification criteria such as age class and site quality, which is correlated with at least the dominant conifer tree species in Norway, may therefore be useful for an efficient stratification of inventories. In the present study, stratification according to age class and site quality was efficient to represent distinct forest types in practical inventories (Naesset 2002).

‘Results

3.1. Regressions models for BA and AGB estimation

The 10 best models were selected for BA and AGB estimation from linear, multivariate linear and non-linear regressions based on their residual distribution, R2 and RMSE. The selected BA and AGB models had R2 of more than 0.70 and 0.85, respectively (see Tables 3 and 4). All return density and height percentiles metrics failed to fulfil good model requirements due to insignificant parameters, low R2, high RMSE and non-normal residual distribution. The best percentile parameters for the models were a + b × P40 and a + b × P25 + c × VAR (see Models No. 1 and 8 in Tables 3 and 4). (Manuri 2017)

Similar to Sheridan et al. (2014), this study also confirmed the heteroscedasticity of the residuals using these parameters. These models were thus excluded in the next step of the analysis. (Manuri 2017)

The power models using CRP variable explained 80.9% and 90.9% of the BA and AGB variations, respectively. These fits were similar or slightly better than those of squared CRP (QCRP) linear models. However, the regressions between predicted and observed values from all linear models had better-fitted lines than power models with slopes not significantly different from 1 and intercepts not significantly different from 0. A 10-fold cross-validation confirmed the lowest RMSE for all CRP-related models (Fig 5.). (Manuri 2017)

All models tended to have low precision in estimating small AGB. Only tree species were recorded on the plots (Manuri 2017). Thus, the biomass values of \*\* aspen groves\*\*\* did not completely represent the actual biomass in the plots. The consequence of excluding trees with a small dbh from plots is that the biomass estimate may not be representative of the distributions of laser canopy heights, especially those with few trees.

These models also tended to under-estimate higher biomass and over-estimate low biomass. The other reason could be due to the model form. Similar trend was found by Englhart et al. (2013) when applying power model. They further suggested to use 2 equations: i.e. power model for low AGB and linear for high AGB values. (Manuri 2017)’

[22] There were slight, nonlinear trends in the relationships between mean DBH and μlidar and σlidarand total basal area and μlidar for coniferous forests ([Figure 3](http://onlinelibrary.wiley.com/doi/10.1029/2008JG000870/full#jgrg498-fig-0003)). We tested various data transformations to improve linear model fits and found that natural log transformations improved models of basal area for both forest types and mean tree height for deciduous forests. More complicated, nonlinear models may have explained the data better; however, our sample sizes were small and there were few degrees of freedom available to fit nonlinear models well. Therefore, we continued the analysis with linear regression models. In spite of the slight nonlinear trends, the scatterplots show that the randomly sampled points did not capture the nonlinear patterns in the data as well as the stratified sample. Hawbaker 2009

[23] Our second goal was to evaluate the two sampling designs and their influence on prediction errors for vegetation structure and biomass. To accomplish this goal, we used the predictor variables selected in the countywide models but estimated regression coefficient values using only the data from each sample design. By using the previously established countywide model form, each sampling design could be evaluated as to how well it captured the trends. This resulted in four models for each tree measurement corresponding to the two sample designs and two forest types (coniferous random, coniferous stratified, deciduous random, and deciduous stratified). Hawbaker 2009

[24] For each model, we calculated the root mean squared error (RMSE) using the model predictions and observations from the data set used to construct the model, referred to as RMSEmodel. Then, we validated models by making predictions for the other sample design and calculated the RMSE from its predictions and observations, referred to as RMSEvalidate. Thus, models built with the stratified sample were validated with the random sample and vice versa. We expected that models built using the stratified sample would have RMSEmodel values similar to RMSEvalidate. Because the random samples may not have included values at the edges of the data distribution, which are more difficult to predict, we expected RMSEmodel for models built using the random sample would be less than RMSEvalidate. Hawbaker 2009

Næsset, E., Gobakken, T., Solberg, S., Gregoire, T. G., Nelson, R., Ståhl, G., & Weydahl, D. (2011). Model-assisted regional forest biomass estimation using LiDAR and InSAR as auxiliary data: a case study from a boreal forest area. *Remote Sensing of Environment*, *115*(12), 3599-3614.

Stephens, P. R., Kimberley, M. O., Beets, P. N., Paul, T. S. H., Searles, N., Bell, A., Brack, C., & Broadley, J. (2012). Airborne scanning LiDAR in a double sampling forest carbon inventory. Remote Sensing of Environment, 117, 348-357.