Modelling of Aboveground Biomass Change Using LiDAR Metrics and NFI Field Data: A Case Study of Southern Sweden

Ritwika Mukhopadhyay[1], Indu Indirabai[1], Mats Nilsson[1], Mikael Egberth[1], Emma Holmström[1], Magnus Ekström[1,2]

[1]Department of Forest Resource Management, Swedish University of Agricultural Sciences, Umeå, Sweden
Email: {ritwika.mukhopadhyay; indu.indirabai; mats.nilsson; mikael.egberth; emma.holmstrom; magnus.ekstrom}@slu.se

[2]Department of Statistics, USBE, Umeå University, Sweden
Email: magnus.ekstrom@umu.se

Keywords: Aboveground biomass, Airborne LiDAR, Categorical variables, Change estimation, GLS, Heteroscedasticity, Random forest.

1. Introduction

The estimation of forest state and change is crucial for the assessment of the changes in the carbon stocks which are necessary as per the requirements of the Kyoto Protocol (UNFCCC 2021). The estimation of the dynamics of forest aboveground biomass (AGB) is also important to study the impact of forest management towards climate change mitigation (e.g., Eggleston et al. 2006; Puliti and Astrup 2020). The field data from the National Forest Inventories (NFIs) are commonly used to estimate the state and change of forest attributes such as AGB and growing stock volume for countries and regions within countries (e.g., Tomppo et al. 2010). However the collection of the data is expensive, time consuming and sometimes also impossible in inaccessible areas (e.g., Saarela et al. 2020). To improve cost-efficiency of the forest attribute estimation, remotely sensed (RS) data are incorporated along with the field data. The use of RS data enables mapping of parameters across the landscape it covers as well as estimation of the target population mean and total (e.g., Saarela et al. 2020). Previous studies have been conducted combining field data with Light Detection and Ranging (LiDAR) data that proved to be efficient in monitoring AGB changes (e.g., Dubayah et al. 2010, Bollandsås et al. 2013, Næsset et al. 2013, Skowronski et al. 2014, McRoberts et al. 2015, Magnussen et al. 2015, Hopkinson et al. 2016, Ene et al. 2017, Puliti and Astrup 2020). Hudak et al. (2012) estimated the change in above ground biomass (ΔAGB, where Δ represents the change) through direct and indirect approaches from the changes in the predictor variables retrieved from the airborne laser scanning (ALS) data. McRoberts et al. (2015) presented direct and indirect estimation methods for ΔAGB using ALS data along with forest inventory data for a boreal forest in Våler Municipality, Norway.

Categorical variables have been implemented for the estimation of AGB using RS data and field inventory data in a number of studies (e.g., Ou et al. 2019, Li et al. 2019 and 2020). In Li et al. (2019) and Li et al. (2020) the categorical variables were formed based on the available field data and Landsat 8 data for different classes of forest crown densities. The categorical variables were used in the parametric models for the estimation of AGB. A comparative analysis between the models with and without categorical variables was performed proving the efficiency of the inclusion of categorical variables in modelling. In Ou et al. (2019), a comparative analysis of parametric models (linear model (LM) and LM with combined variables) and non-parametric methods (random forest (RF) and artificial neural network (ANN)) was conducted based on the inclusion of categorical variables for different age classes of *Pinus densata* forests. The models included categorical variables were observed to improve the overall accuracy of estimation by 14-42% and 32-44% for the training and testing plots based on the root mean-squared error (RMSE) values.

The objective of this study was to incorporate parametric models (LMs) and non-parametric (RF) methods along with categorical variables and using NFI field data and auxiliary LiDAR data for the estimation of ΔAGB. The study is mainly focused to observe the ability of LiDAR for ΔAGB estimation when different management practices of the forests are taken into account. The categorical variables were grouped based on the management practices such as, thinning and felling operations conducted in the plots.

Published in: Markus Hollaus, Norbert Pfeifer (Eds.): Proceedings of the SilviLaser Conference 2021, Vienna, Austria, 28–30 September 2021. Technische Universität Wien, 2021. DOI: 10.34726/wim.1861
This paper was peer-reviewed. DOI of this paper: 10.34726/wim.1982

2. Material and Methods

The study area is located in south of Sweden with a forest cover of 332171.8 ha and species composition with proportions such as, 24.6% Pine (*Pinus sylvestris*), 53.8% Spruce (*Picea abies*), 11.1% Birch (*Betula spp.*) and 24.5% of other broadleaved tree species. The Swedish NFI field data were available for 218 plots for two time periods, 2010-2014 and 2015-2019. The plots were circular with 10m radius sampled using the systematic cluster sampling method.

For each corresponding field plot the LiDAR metrics were retrieved using the Fusion software (McGaughey 2020). Laser returns above 1.5m height were retained in order to eliminate the non-vegetation returns. The LiDAR metrics used for the regression modelling were 80% height percentile (h_{p80}) and the vegetation ratio (vr) based on the previous studies (Nilsson et al. 2017, Saarela et al. 2020).

The ΔAGB was estimated directly from the plot-level NFI data available for time period 2 (2015-2019) and 1 (2010-2014). The data were grouped based on the silvicultural operations into three categories namely, plots with thinning operation, plots with clear felling operations and plots with no activity. The plot-level values for ΔAGB and the change in LiDAR metrics (ΔLiDAR metrics) were used for developing the relationship between the response variable (ΔAGB) and the predictor variables (Δh_{p80} and Δvr) along with the categorical variables (indicators I_1 and I_2). I_1 and I_2 represent the categories of plots with no activity and with thinning operation, respectively. Figure 1 presents an overview of the modelling workflow.

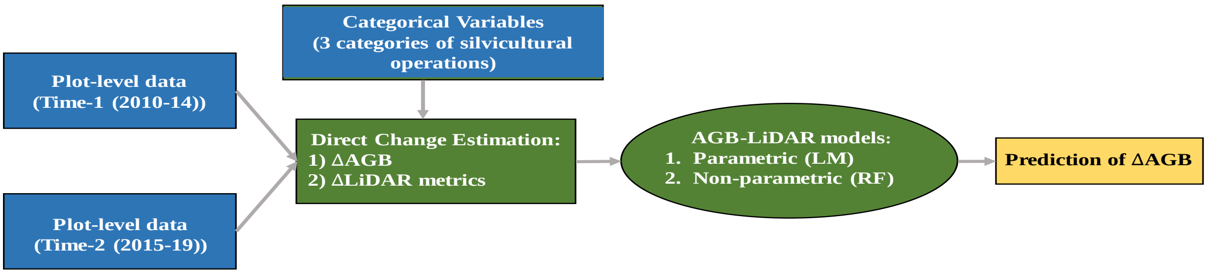

Figure 1: The methodological overview of this study.

For the parametric modelling, the LMs with and without accounting for heteroscedasticity were used along with the LiDAR metrics and categorical variables. For the model accounting for heteroscedasticity the *nlme* package in R was used (Pinheiro et al. 2021). The first model selected for the study was the LM with no intercept and with the LiDAR metrics and the categorical variables assuming that the random errors are homoscedastic. The second parametric model accounted for heteroscedasticity in random errors. To calculate weights, the variance function with the exponential form of the random error variance was selected from the *nlme* package in R (Pinheiro et al. 2021). The selection was based on the Akaike information criterion (AIC). The models have been represented in Table 1.

Table 1. The model forms of the parametric LMs, where, β and γ represent the coefficients and ε represents the random error for the first model. α and δ represent the coefficients and υ represents the random error for the second model. And, I_1 and I_2 represent the first and the second group of the categorical variables.

Model type	Model form
LM (no account for heteroscedasticity)	$\Delta AGB = \beta_1 \Delta hp_{80} + \beta_2 \Delta vr + \gamma_1(\Delta hp_{80} \cdot I_1) + \gamma_2(\Delta vr \cdot I_1) + \gamma_3(\Delta hp_{80} \cdot I_2) + \gamma_4(\Delta vr \cdot I_2) + \varepsilon$
LM (account for heteroscedasticity)	$\Delta AGB = \alpha_1 \Delta hp_{80} + \alpha_2 \Delta vr + \delta_1(\Delta hp_{80} \cdot I_1) + \delta_2(\Delta vr \cdot I_1) + \delta_3(\Delta hp_{80} \cdot I_2) + \delta_4(\Delta vr \cdot I_2) + \upsilon$

For the non-parametric modelling, the RF method with and without categorical variables were used. The RF methods were formed using the *randomForest* package in R (Liaw and Wiener 2002). For this study, the default value of 'n_{tree}' = 500 trees was used and the same dataset was used to fit the parametric models and the non-parametric methods for ΔAGB prediction.

3. Results and Discussion

The predicted ΔAGBs versus the field ΔAGBs were plotted for the four models, as seen in Figure 2. In case of the parametric models the under estimation of ΔAGB values for the plots for category 1 (plots with no activity) is higher compared to that of the non-parametric models where the underestimation is observed to be mostly in the positive range of the predicted ΔAGB values. The overestimation of the ΔAGB values for category 1 is observed to be clustered around 0 in case of the parametric models whereas, the RF model with categorical variables has a lower range of overestimated ΔAGB values. For category 2 (plots with thinning operation) the range of overestimation of ΔAGB values is observed to be lower and more spread out in case of the RF model with categorical variables compared to the other three models. The RMSE values of the four models have been listed in Table 2. The models with interactions with categorical variables have lower RMSE values. Out of the three models with categorical variable interactions, the LM model (with no account of heteroscedasticity) is observed to have a wider range of predicted ΔAGB values and the lowest RMSE value of 32.269 $Mgha^{-1}$ followed by the RF method with RMSE value of 34.608 $Mgha^{-1}$.

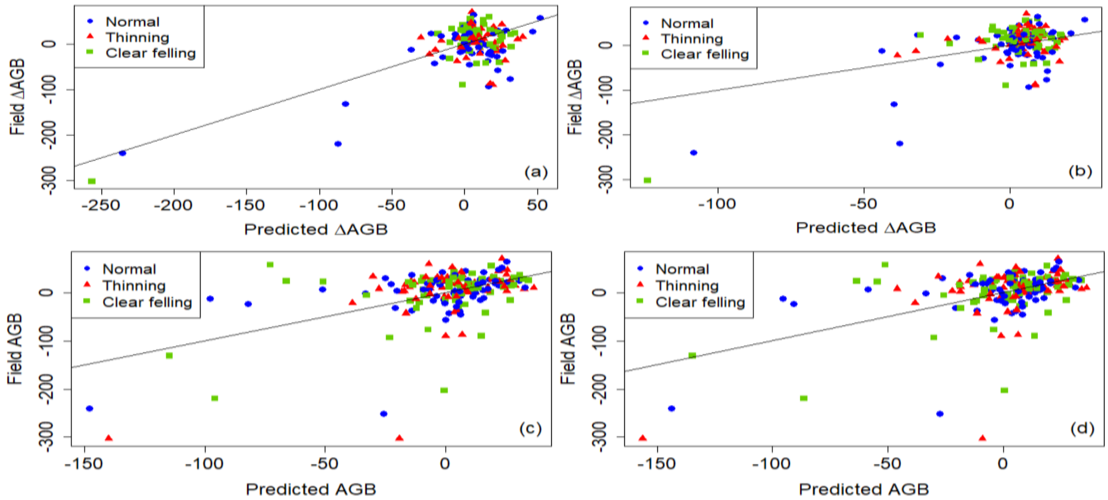

Figure 2: Predicted ΔAGB versus measured ΔAGB for: (a) LM without account for heteroscedasticity, (b) LM with account for heteroscedasticity, (c) RF without categorical variables, and, (d) RF with categorical variables.

The ΔAGB values can be predicted for the entire study area on the availability of the raster maps for the categorical variables based on the silvicultural operations conducted. From the trend in the above plots it can be expected to have a more heterogeneous map of predicted values of ΔAGB in case of the non-parametric RF method as the overestimation and underestimation of the smaller and larger ΔAGB values, respectively, is lesser compared to that of the parametric models.

Table 2. The models with their respective RMSE values in $Mgha^{-1}$.

Model/ Method	RMSE ($Mgha^{-1}$)
LM (no account for heteroscedasticity)	32.269
LM (account for heteroscedasticity)	35.882
RF (without categorical variables)	43.708
RF (with categorical variables)	34.608

4. Conclusions

In this study, we incorporated the parametric and non-parametric models with categorical variables based on the different silvicultural operations conducted in the sample plots. The incorporation of the categorical variables along with LiDAR metrics was seen to improve the accuracy of ΔAGB prediction.

It was observed from this study that the models with interactions with categorical variables perform better, out of which the LM assuming the random errors are homoscedastic was observed to perform the best in terms of yielding the lowest RMSE value of 32.269 $Mgha^{-1}$. And, also the overestimation of lower ΔAGB values and underestimation of higher ΔAGB values was seen to improve in case of the non-parametric RF model along with the grouped factor of silvicultural operations.

Acknowledgements

Funding was provided by grants from the Swedish Research Council for Sustainable Development (FORMAS: FR-2019/0007), the Swedish Kempe Foundation (Kempestiftelserna: SMK-1847) and the Swedish National Space Agency (SNSA-171/19). We acknowledge the Swedish NFI for providing the field data used in the study.

References

Bollandsås O M, Gregoire T G, Næsset E and Øyen B H, 2013, Detection of Biomass Change in a Norwegian Mountain Forest Area Using Small Footprint Airborne Laser Scanner Data. *Statistical Methods and Applications*, 22(1):113–29. doi: 10.1007/s10260-012-0220-5.

Dubayah R O, Sheldon S L, Clark D B, Hofton M A, Blair J B, Hurtt G C and Chazdon R L, 2010, Estimation of Tropical Forest Height and Biomass Dynamics Using Lidar Remote Sensing at La Selva, Costa Rica. *Journal of Geophysical Research: Biogeosciences*, 115(G2):n/a-n/a. doi: 10.1029/2009JG000933.

Eggleston H S, Buendia L, Miwa K, Ngara T and Tanabe K, 2006, *2006 IPCC Guidelines for National Greenhouse Gas Inventories.* Institute for Global Environmental Strategies (IGES), Hayama, Japan.

Ene L T, Næsset E, Gobakken T, Bollandsås O M, Mauya E W and Zahabu E, 2017, Large-Scale Estimation of Change in Aboveground Biomass in Miombo Woodlands Using Airborne Laser Scanning and National Forest Inventory Data. *Remote Sensing of Environment*, 188:106–17. doi: 10.1016/j.rse.2016.10.046.

Hopkinson C, Chasmer L, Barr A G, Kljun N, Black T A and McCaughey J H, 2016, Monitoring Boreal Forest Biomass and Carbon Storage Change by Integrating Airborne Laser Scanning, Biometry and Eddy Covariance Data. *Remote Sensing of Environment*, 181:82–95. doi: 10.1016/j.rse.2016.04.010.

Hudak A T, Strand E K, Vierling L A, Byrne J C, Eitel J U H, Martinuzzi S and Falkowski M J, 2012, Quantifying Aboveground Forest Carbon Pools and Fluxes from Repeat LiDAR Surveys. *Remote Sensing of Environment*, 123:25–40. doi: 10.1016/j.rse.2012.02.023.

Li C, Li M and Li Y, 2020, Improving Estimation of Forest Aboveground Biomass Using Landsat 8 Imagery by Incorporating Forest Crown Density as a Dummy Variable. *Canadian Journal of Forest Research*, 50(4):390–98. doi: 10.1139/cjfr-2019-0216.

Li C, Li Y and Li M, 2019, Improving Forest Aboveground Biomass (AGB) Estimation by Incorporating Crown Density and Using Landsat 8 OLI Images of a Subtropical Forest in Western Hunan in Central China. *Forests*, 10(2):104. doi: 10.3390/f10020104.

Liaw A and Wiener M, 2002, *Classification and Regression by RandomForest*. Vol. 2.

Magnussen S, Næsset E and Gobakken T, 2015, Lidar-Supported Estimation of Change in Forest Biomass with Time-Invariant Regression Models. *Canadian Journal of Forest Research*, 45(11):1514–23. doi: 10.1139/cjfr-2015-0084.

McGaughey B, 2020, FUSION/LDV LIDAR Analysis and Visualization Software. *Pacific Northwest Research Station USDA Forest Service*, Retrieved May 14, 2021. (http://forsys.cfr.washington.edu/fusion/fusion_overview.html).

McRoberts R E, Næsset E, Gobakken T and Bollandsås O M, 2015, Indirect and Direct Estimation of Forest Biomass Change Using Forest Inventory and Airborne Laser Scanning Data. *Remote Sensing of Environment*, 164:36–42. doi: 10.1016/j.rse.2015.02.018.

Næsset E, Bollandsås O M, Gobakken T, Gregoire T G and Ståhl G, 2013, Model-Assisted Estimation of Change in Forest Biomass over an 11year Period in a Sample Survey Supported by Airborne LiDAR: A Case Study with Post-Stratification to Provide ‘Activity Data. *Remote Sensing of Environment*, 128:299–314. doi: 10.1016/j.rse.2012.10.008.

Nilsson M, Nordkvist K, Jonzén J, Lindgren N, Axensten P, Wallerman J, Egberth M, Larsson S, Nilsson L, Eriksson J and Olsson H, 2017, A Nationwide Forest Attribute Map of Sweden Predicted Using Airborne Laser Scanning Data and Field Data from the National Forest Inventory. *Remote Sensing of Environment*, 194:447–54. doi: 10.1016/j.rse.2016.10.022.

Ou G, Li C, Lv Y, Wei A, Xiong H, Xu H and Wang G, 2019, Improving Aboveground Biomass Estimation of Pinus Densata Forests in Yunnan Using Landsat 8 Imagery by Incorporating Age Dummy Variable and Method Comparison. *Remote Sensing*, 11(7):738. doi: 10.3390/rs11070738.

Pinheiro J, Bates D, DebRoy S, Sarkar D, Heisterkamp S, Willigen B V and Ranke J, 2021, *Package "nlme" Linear and Nonlinear Mixed Effects Models*.

Puliti S and Astrup T F B, 2020, *Above-Ground Biomass Change Estimation Using National Forest Inventory Data with Sentinel-2 and Landsat 8*.

Saarela S, Wästlund A, Holmström E, Mensah A A, Holm S, Nilsson M, Fridman J and Ståhl G, 2020, Mapping Aboveground Biomass and Its Prediction Uncertainty Using LiDAR and Field Data, Accounting for Tree-Level Allometric and LiDAR Model Errors. *Forest Ecosystems*, 7(1). doi: 10.1186/s40663-020-00245-0.

Skowronski N S, Clark K L, Gallagher M, Birdsey R A and Hom J L, 2014, Airborne Laser Scanner-Assisted

Estimation of Aboveground Biomass Change in a Temperate Oak-Pine Forest. *Remote Sensing of Environment*. doi: 10.1016/j.rse.2013.12.015.

Tomppo E, Schadauer K, McRoberts R E, Gschwantner T, Gabler K and Ståhl G, 2010, *National Forest Inventories*. Springer, Netherlands.

UNFCCC, 2021, Kyoto Protocol - Targets for the First Commitment Period | UNFCCC. Retrieved May 12, 2021. (https://unfccc.int/process-and-meetings/the-kyoto-protocol/what-is-the-kyoto-protocol/kyoto-protocol-targets-for-the-first-commitment-period).