Model-Assisted Estimation of Timber Volume by Means of Harvester and ALS Data

J. Räty[1*], R. Astrup[1], J. Breidenbach[1*]

[1]Norwegian Institute of Bioeconomy Research (NIBIO), Høgskoleveien 8, 1433 Ås
Email: {janne.raty; rasmus.astrup; johannes.breidenbach}@nibio.no
*Corresponding author

1. Introduction

Cut-to-length harvesters can provide ground-truth data for predicting forest attributes using airborne laser scanning (ALS) data (Söderberg et al. 2021). Harvester datasets are, however, not a representative sample, which may cause limitations regarding the applicability of the harvester datasets in forest inventories. Harvests are typically carried out in actively managed and harvest-ready forests, which means that the use of harvester-collected ground-truth data as reference data may lead to systematic errors in maps and estimates of forest attributes. It is not fully understood if systematic errors can be of relevance in the estimation of forest attributes when training models with harvester data.

Our objective was to study the applicability of harvester data in the model-assisted (MA) estimation (e.g. Räty et al. 2021) of timber volume in a 250,000 ha study area in Norway. We predicted timber volume for National Forest Inventory (NFI) plots using harvester and ALS data and evaluated systematic errors using correction factors associated with the MA estimates. We also compared the efficiencies of the direct (field data-based) and MA estimators.

2. Data and Methods

2.1 Study area and data

The 250,000 ha study area is located in the Innlandet county in Norway and comprises seven municipalities: Etnedal, Gausdal, Nordre Land, Nordre-Aurdal, Vang, Vestre Slidre, and Øystre-Slidre. In the study area, forests cover 215,000 ha of which 65% are dominated by Norway spruce (*Picea abies* [L.] Karst.).

The NFI data utilized here were collected between 2014 and 2018. We used plots in the lowland stratum (Breidenbach et al. 2020) where sample plots are located on a 3×3 km systematic grid. The field plots are circular plots (250 m²) and each tree with a DBH ≥ 5 cm was measured. There were 157 spruce-dominated field plots (248 field plots in forest, 277 field plots in total) within our study area.

The harvester data used for the fitting of a volume model were collected from spruce-dominated clear-cut areas using a Komatsu 931XC harvester in 2020 and 2021. DBH was registered for each harvested tree and the tree heights were predicted using taper curves. Treetop volumes missing from harvester data were predicted using tree-level volume functions. The harvester registered the XY position of the harvester head for each harvested tree with a positioning accuracy of approximately 5–10 m. The trees were linked to the stand-like segments of Norwegian Forest Resource Map SR16 (Astrup et al. 2019). The SR16 segments were further cropped using alpha-shapes (α-shapes) around the XY positions of the harvested trees. The resulting segments are called harvested segments.

The harvested trees were also linked to the SR16 grid cells (256 m²). We omitted grid cells with obvious discrepancies between the mean height of harvested trees and the 95% ALS height percentile. We also omitted grid cells with less than 66% of cell area in the harvested segments. In total, 166 harvested segments (minimum 0.1 ha, mean 0.7 ha) and 2,953 grid cells comprising 80,099 harvested trees were used for modelling. The mean timber volume associated with the harvested segments and grid cells were 220 and 225 $m^3 \cdot ha^{-1}$.

Low-density (< 5 pulses per m²) ALS data were collected from the study area between 2011 and 2017. Standard ALS data processing steps of the area-based approach were carried out, and a set of ALS predictor variables (features) like mean height, height percentiles and density metrics were calculated

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.2006

for the harvested segments, harvested SR16 grid cells, and NFI plots. The set of ALS features comprised 28 features computed from the ALS point cloud.

2.2 Prediction of volume

We fitted k nearest neighbour (NN) models (k = 5) using both the harvested SR16 grid cells and the harvested segments. Three predictor variables were selected by means of an optimization algorithm.

2.3 Estimation methodology

We estimated the mean volume for spruce-dominated forests in the study area from the NFI data (direct estimation). The variances were estimated for both direct and MA estimates. In the MA estimation, the NN models were applied to the NFI plots, and the prediction residuals associated with the NFI plots were utilized in the variance estimation.

The performances of the MA and direct estimators were compared using the relative efficiency (RE, ratio of variances), and the half width of 95 % confidence intervals (CI). Our main interest was on the estimates of correction factors ($\hat{\mu}_{cor}$) that indicate the magnitude of systematic errors in the synthetic ("pixel counting") estimate. The CIs and correction factors are presented as a percentage value in terms of direct estimates (Räty et al. 2021).

3. Results and Discussion

Despite the good performance of the NN models (Figure 1), the non-zero correction factors of MA estimates showed that the synthetic ("pixel-counting") estimate of timber volume resulted in systematic errors (Table 1). The NN model fitted using the harvested segments produced a negative correction factor (overestimation). This indicates that the NN model was not capable of extrapolating outside the training data which was mainly collected from mature forests. The NFI data also comprised plots from younger forests for which volume was often overestimated. The NN model fitted using the SR16 grid cells as modeling units resulted in a positive correction factor (underestimation). The positioning errors of trees were non-negligible and may have negatively affected the predictive performance of the NN model fitted using the harvested SR16 grid cells (Figure 1). It is also worth noting that the time lag between the ALS data acquisition and NFI data differs from the time lag between the ALS data acquisition and harvester data, which can also be a minor source of systematic errors.

The use of the harvester-based models resulted in a considerable efficiency gain compared with the direct estimation regardless of the modeling unit (Table 1). The largest RE of 6.17 was achieved using the NN model fitted using harvested segments. The additional use of harvester and ALS data more than halved the CI of the estimate.

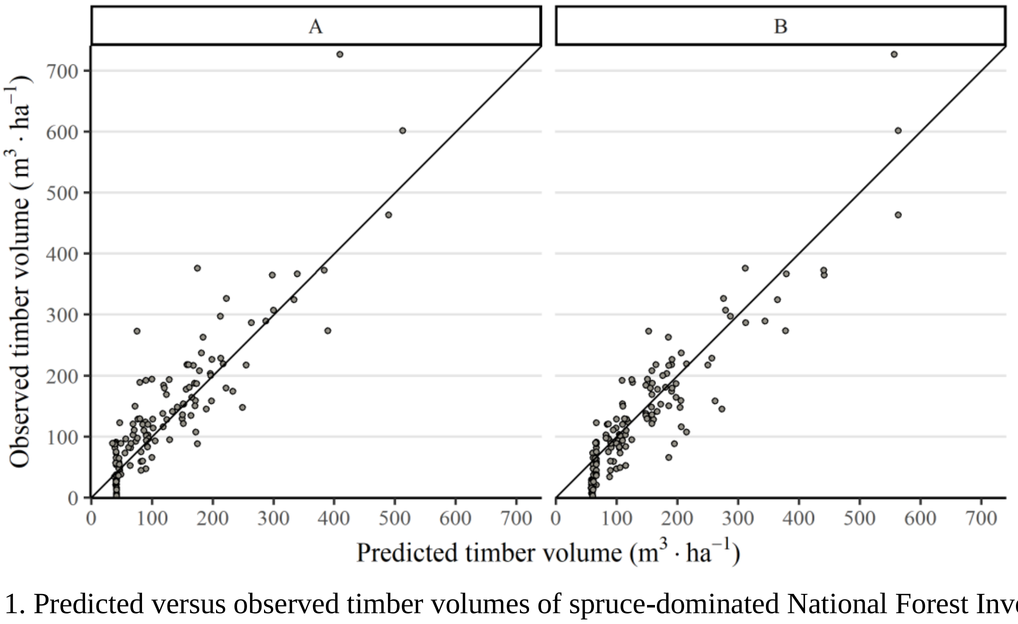

Figure 1. Predicted versus observed timber volumes of spruce-dominated National Forest Inventory field plots. A: model fit using harvested SR16 grid cells and B: model fit using harvested segments.

Table 1. Characteristics associated with the estimation of timber volume for spruce-dominated forests of the study area. RE – relative efficiency, MA – model-assisted.

Modeling unit	Direct estimate $\hat{\mu}$ ($m^3 \cdot ha^{-1}$)	Half 95% CI $\hat{\mu}$, (%)	Half 95% CI $\hat{\mu}_{MA}$, (%)	Correction factor $\hat{\mu}_{cor}$, (%)	RE
Grid cell	121.17	14.42	6.65	8.18	4.71
Segment			5.81	-11.39	6.17

4. Conclusions

We draw the following conclusions: (1) The use of a model fitted using cut-to-length harvester and ALS data results in considerable efficiency gains in the model-assisted estimation of timber volume. (2) Harvester data can be valuable for model fitting despite of non-negligible uncertainties in harvester-recorded stem positions. (3) Synthetic ("pixel-counting") estimates of harvester-based forest attribute maps can result in systematic errors that need to be corrected for to avoid wrong conclusions.

Acknowledgements

This study was supported by NIBIO (Norwegian Institute of Bioeconomy Research) and the PRECISION project (NFR# 11067). We acknowledge Simon Berg for technical support with the harvester data.

References

Astrup R, Rahlf J, Bjørkelo K, Debella-Gilo M, Gjertsen A-K, Breidenbach J, 2019. Forest information at multiple scales: development, evaluation and application of the Norwegian forest resources map SR16. *Scandinavian Journal of Forest Research*, 34: 484-496.

Breidenbach J, Granhus A, Hylen G, Eriksen R, Astrup R, 2020. A century of National Forest Inventory in Norway – informing past, present, and future decisions. *Forest Ecosystems*, 7:46.

Räty J, Astrup R, Breidenbach J, 2021. Prediction and model-assisted estimation of diameter distributions using Norwegian national forest inventory and airborne laser scanning data. *Canadian Journal of Forest Research*.

Söderberg J, Wallerman J, Almäng A, Möller JJ, Willén E, 2021. Operational prediction of forest attributes using standardised harvester data and airborne laser scanning data in Sweden. *Scandinavian Journal of Forest Research*, 36: 306-314.