# Predicting stand-thinning maturity from airborne laser scanning data

**Article** *in* Scandinavian Journal of Forest Research · April 2011

DOI: 10.1080/02827581.2010.547870



Mikko Vastaranta
University of Eastern Finland


Markus Holopainen
University of Helsinki


Xiaowei Yu
Shanghai Jiao Tong University affiliated Sixth People's Hospital


Juha Hyyppä
Finnish Geospatial Research Institute at National Land Survey

# PREDICTING STAND-THINNING MATURITY FROM AIRBORNE LASER SCANNING DATA

Mikko Vastaranta[1], Markus Holopainen[1], Xiaowei Yu[2], Juha Hyyppä[2], Hannu Hyyppä[3], Risto Viitala[4]

[1] University of Helsinki, Department of Forest Sciences, Latokartanonkaari 7, 00014 Helsinki, Finland, (mikko.vastaranta, markus.holopainen)@helsinki.fi

[2] Finnish Geodetic Institute, Geodeetinrinne 2, PL 15, 02431 Masala, Finland. (yu.xiaowei, juha.hyyppa)@fgi.fi,

[3]Helsinki University of Technology, Research Institute of Modelling and Measuring for the Built Environment, Finland, hannu.hyyppa@tkk.fi

[4] HAMK University of Applied Sciences, risto.viitala@hamk.fi

**Pre-print of published version. Reference: Vastaranta, M., Holopainen, M., Yu, X., Hyyppä, J., Hyyppä, H. and Viitala, R., 2011. Predicting stand-thinning maturity from airborne laser scanning data. Scandinavian Journal of Forest Research, 26(2), pp.187-196.**

**DOI: 10.1080/02827581.2010.547870**



## Abstract

Airborne laser scanning (ALS) has been used in recent years to acquire accurate remote-sensing material for carring out practical forest inventories. Still, much of the information needed in forest management planning must be collected in the field. For example, forest management proposals are often determined in the field by an expert. In the present paper, statistical features extracted from ALS data were used in logistic regression models and in nonparametric *k*-MSN estimation to predict the thinning maturity of stands. The research material consisted of 381 treewise measured circular plots in young and advanced thinning stands from the vicinity of Evo, in southern Finland. Timing of thinning was determined in the field by an expert and coded as a binary variable. Models were developed (1) to locate stands that will reach thinning maturity within the next 10-year period, and separately (2) for stands in which commercial thinning should be done immediately. For comparison purposes, logistic regression models were formulated from accurately measured stand characteristics. Logistic regression models based on ALS features predicted the thinning maturity with a classification accuracy of 79% (1) and 83% (2). The respective percentages were 66% and 83% with models based on stand characteristics and 70% and 86% with *k*-MSN. The study showed that ALS data can be used to predict stand-thinning maturity in a practical way.

**Key words:** *airborne laser scanning, forest management, k-MSN, logistic regression, thinning*

## 1. Introduction

In Finland, information for forest management planning is collected in two phases. In the first phase, inventory of the stand characteristics is carried out, and in the second, the data are augmented with information collected in the field. Information collected in the field includes forest site types, biodiversity targets and forest management proposals. An area-based airborne laser-scanning (ALS) inventory method (Naesset 2002) has been introduced for practical forest stand characteristic inventories in Finland. The method uses statistical features extracted from the ALS data in addition to aerial photographs, and the estimation of stand characteristics is based on nonparametric estimation methods (Maltamo et al., 2006). ALS is the most accurate remote-sensing (RS) technique for forest inventory, providing relative accuracies ranging between 10% and 20% at the stand level for mean volume (e.g. Hyyppä & Hyyppä, 1999; Næsset, 2002; Maltamo et al., 2006). The current data acquisition cost is comparable to that of the traditionally used inventory method – standwise field inventory (SWFI) – which is now being replaced. ALS devices providing small-footprint diameters (10–30 cm) allow accurate height determination of the forest canopy (e.g. Næsset, 1997; Magnussen & Boudewyn, 1998; Magnussen et al., 1999; Means et al., 1999). Although ALS inventory provides accurate estimates for stand characteristics, much of the additional information needed in forest management planning must be collected in the field, e.g. forest management proposals. Forest management planning has yet to make the most of the information obtained by ALS and better linking methods should be studied.

Commercial thinnings are management practices in which both the silvicultural and economic aspects are taken into account. From the silvicultural point of view, the goal of thinning is to provide enough growing space and thus improve the vitality of the future crop trees. The timing and intensity of thinning are always affected by the previous management of the stand. Harvest schedules and timing of forest operations are selected so that the utility of the decision maker is maximized. The decision maker is usually assumed to maximize the net present value (NPV) of the forest area or stand. General guidelines for timing of thinning and clear-cuts are presented as recommendations for good silviculture (Recommendations of Tapio…, 2006), which are based on growth and yield studies. In practical forest management planning, stand-thinning maturity is determined in the field by a forest manager and decisions are based on the above-mentioned guidelines, spatial distributions of trees and vigour of the tree crowns. When forest management calculation systems are applied, the thinning maturity is determined by the basal area (BA) or stem number and dominant height (HDOM) via thinning curves as the clear-cutting maturity is determined by the age or mean diameter (Recommendations of Tapio…, 2006). When timing of harvest is proposed computationally, these proposals include a degree of uncertainty and are often replaced by proposals done in the field. This is mainly caused by the incapability of forest management calculation systems to predict stand spatial distributions of trees and vigour of the living crowns, i.e. the silvicultural aspects of the thinning are ignored. Delayed thinning or incorrect harvest decisions may result in growth and income losses for forest owners (Haara & Korhonen, 2004).

A few studies (e.g. Hyvönen, 2002; Vastaranta, 2006) have addressed the convergence of thinning decisions determined in the field and produced computationally. More studies have been done to clarify the reliability of computationally determined thinning proposals when these are made from forest stand characteristics that include uncertainty (e.g. Ojansuu et al., 2002; Haara & Korhonen, 2004; Vanhatalo 2010). The effects of ALS inventory and SWFI errors have previously been examined (e.g. by Eid et al. 2004, Holopainen & Talvitie, 2006; Holopainen et al., 2010) with respect to the timing of harvests and the NPV of harvest outturns. In these studies the basic assumption was that erroneous inventory data result in less than optimal timing of harvests, which leads to losses in harvest revenues. Holopainen et al. (2010) showed that input data accuracy significantly affects both harvest timing and harvest revenue NPV. With respect to harvest timing, the inaccurate inventory data resulted in an error in thinning or clear-cutting timing, ranging from 6.5 to 10.3 years, depending on input data source and simulation methodology. With respect to simulated harvest revenue NPV, the inaccurate inventory data resulted in a relative error ranging from 28.2% to 57%. Ojansuu et al. (2002) and Vastaranta (2006) studied the effects of erroneous input data and errors in growth models on thinning proposals made by the forest management calculation system. Haara and Korhonen (2004) studied the effects of various error sources on forest management decisions in updated SWFI data. In the above-mentioned studies, errors in input data have had the most severe effects. Vanhatalo (2010) searched for tolerable error intervals for input stand characteristics, for which the next simulated forest operation would still be correct.

In many previous studies RS material has been used in the detection of forest operations, such as thinning or clear-cuttings (e.g.Varjo 1996, Hyvönen & Anttila 2006, Yu et al. 2004, Hyvönen et al. 2010). Timing of forest management operations with RS material has been examined in only a few studies. Landsat TM satellite images and stand register data have been used in nonparametric *k*–nearest neighbour (*k*-NN) estimation of forest stand characteristics and forest management actions (Hyvönen, 2002). Root-mean-squared errors (RMSEs) achieved for stand characteristics, e.g. 42.1% for mean volume, were much higher than RMSEs achieved in several ALS studies (e.g. Naesset, 2002; Maltamo et al., 2006; Holopainen et al., 2008). The classification accuracies for forest management operations were 61.3% for thinnings and 64.1% for clear-cuttings (Hyvönen, 2002). Pesonen et al. (2007) used national forest inventory data and Landsat TM satellite images to locate stands needing precommercial thinning. In their study, stand precommercial thinning need was correctively classified with timing in 58% and without timing in 74% of stands. Precommercial thinning need has also been estimated with ALS data combined with information from forest management plans (Närhi et al. 2008). Närhi et al. used three urgency classes for forest management operations and the classification accuracy varied from 63.6% to 85.7% with overall accuracy of 71.8%.

Information provided by ALS offers new perspectives for assessing stand-thinning maturity, and since ALS is becoming a widely popular method for forest inventory, these tools could be adopted for practical forest management planning. As a method, ALS provides direct information from the highest parts of the tree crown and structure of the forest canopy. ALS-based estimates for crown closure, leaf-area index (LAI) or crown coverage have been utilized in many studies (Means et al., 1999; Holmgren et al., 2003; Hopkinson & Chasmer, 2009). Basically, variation in the penetration of laser pulses describes the variation in density of the crowns (Hirata et al., 2009). For practical forest inventory, statistical features are extracted from the ALS data. These include penetration as vegetation laser pulse returns *versus* total returns, height percentiles of the distribution of canopy heights, and canopy cover percentile as a proportion of laser returns below a given percentage of total height. From these, features such as penetration and canopy cover percentiles are used to predict LAI and crown closure with relatively good accuracy (e.g. Hopkinson & Chasmer, 2009; Jensen et al., 2008). These features probably correlate well with stand-thinning maturity.

The objective of the present study was to test the accuracy of the thinning maturity predictions determined from ALS data with canopy-based statistical features. Logistic regression models and nonparametric *k*-most similar neighbour (*k*-MSN) estimation were used. Reference suggestions for stand-thinning maturity and the timing of thinning were determined in the field. For comparison, logistic regression models were also formulated using accurate field-measured characteristics such as HDOM and BA, on which thinning proposals are usually based.

## 2. Material and methods

### Study area

The study area is in an app. 2000-ha managed forest area located in the vicinity of Evo, Finland (61.19° N, 25.11° E, Fig 1.). The area is dominated by coniferous tree species, namely Scots pine (*Pinus sylvestris* L.) (52%) and Norway Spruce (*Picea abies* (L.) H. Karst.) (31%). Classified by stand development class, the area consists mainly of young thinning stands (26%), advanced thinning stands (40%) and mature stands (23%). The corresponding proportions of forest site type are as follows: grass-herb sites (8%), moist sites (71%), dry sites (19%) and poor sites (2%).


Figure 1. Location of the study area.

### Field measurements and determination of forest management actions

The research material consisted of 381 tree-level measured fixed-radius (10 m) plots located in young and advanced thinning stands. The plots were located with a Trimble's GEOXM 2005 Global Positioning System (GPS) device (Trimble Navigation Ltd., Sunnyvale, CA, USA), and the locations were postprocessed with local base station data, resulting in an average error of app. 0.6 m. Tree-level field measurement data from these fixed-radius (10 m) field plots were collected in 2007, 2008 and 2009. The following variables were measured from trees having a diameter-at-breast height (dbh) of over 5 cm: location, tree species, dbh and height. BA and HDOM were computed based on accurate field measurements. The field plots are henceforth referred to as stands.

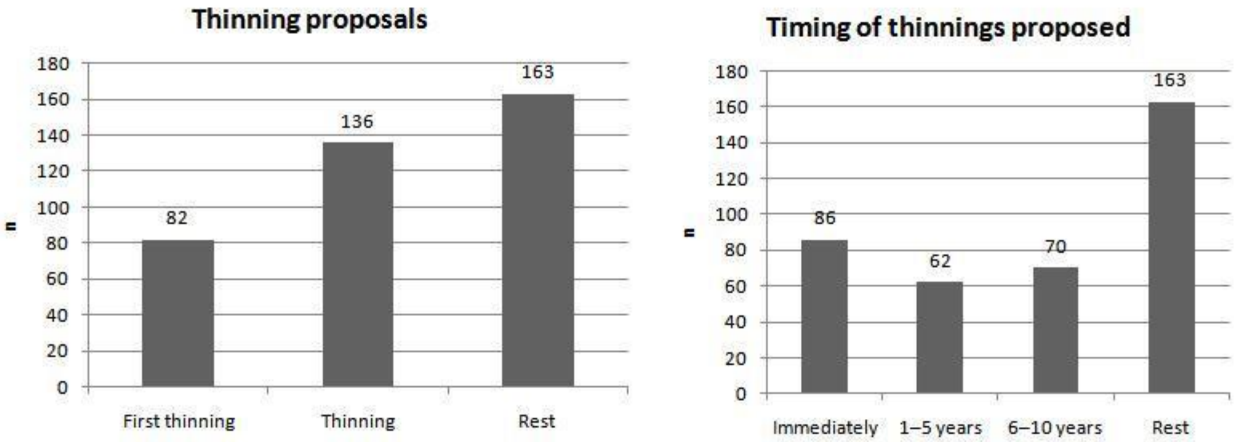


Figure 2. Thinning proposals and timings determined in the field.

Proposals for forest management actions were determined in the field for all 381 stands in winter 2009 (Fig 2.). The studied proposals included first commercial thinning and later thinning. Timing of thinning is dependent on both the silvicultural and economic aspects and affects the entire forest management chain. Timing of the determined thinning was classified as immediate, 1-5 years or 6-10 years. If there was no thinning proposal, it was interpreted as a rest. Thinnings were proposed for 228 stands, as the next forest management action for a 10-year planning period. These proposals included 89 first thinnings and 139 later thinnings. Immediate thinning proposals included 35 first and 53 later thinnings. Stands were randomly divided for modelling (281) and validation (100). Statistics of stand characteristics describing thinning maturity are presented in Table 1.

Table 1. Statistics of stand characteristics describing thinning maturity in test and modelling data.

| | | n | MIN | MEAN | MAX | SD |
|---|---|---|---|---|---|---|
| TEST | BA, $m^2ha^{-1}$ | 100 | 3.2 | 18.8 | 36.2 | 8.1 |
| TEST | HDOM, m | 100 | 8.5 | 17.9 | 27.8 | 4.1 |
| MODELLING | BA, $m^2ha^{-1}$ | 281 | 2.5 | 19.6 | 45.6 | 7.7 |
| MODELLING | HDOM, m | 281 | 7.4 | 17.9 | 44.6 | 4.3 |

## Acquisition and Processing of ALS Data

The ALS data were acquired on July 2009 with an Optech 3100 laser scanner (Optech Inc., Vaughan, Ontario, Canada). The flying altitude was 400 m. The density of the returned pulses within the plot was approximately 10 points per $m^2$. The ALS data were first classified into ground and nonground points, using the standard approach of the TerraScan-based method explained in Axelsson (2000). A digital terrain model (DTM) was then developed, using classified ground points, and laser heights above the ground (normalized height or canopy height) were calculated by subtracting the ground elevation from the laser measurements. Canopy heights close to zero were considered as ground returns and those greater than 2 m as vegetation returns. The intermediate data between them were considered as returns from ground vegetation or bushes. Only vegetation returns were used for ALS feature extraction. Several features were extracted from the vegetation returns for plot. They included the maximum laser hit of the plot, mean, standard deviation and coefficient of variation of the canopy heights, penetration as vegetation returns *versus* total returns, height percentiles of the distribution of canopy heights from 10% to 100% at intervals of 10%, and canopy cover percentile as a proportion of laser returns below a given percentage (from 10% to 100% at 10% intervals) of the total height (Table 2.).

Table 2. ALS features used.

| Feature | Description |
|---|---|
| Hmax | Maximum laser height |
| Hmean | Arithmetic mean of laser heights |
| penetration | Proportion of vegetation hits |
| CV | Coefficient of variation |
| h10-90 | Percentiles of canopy height distribution |
| p10-90 | Canopy cover percentile as proportion returns below certain percentages of total height |

**Logistic regression**

The probability of the stand-thinning maturity was modelled with multible logistic regression, using the function glm in the R statistical package (R Development Core Team, 2007). Logistic regression is commonly used in modelling the probability of an event's occurrence. In logistic regression, logit transformation is used to make the relationship between the response probability and the explanatory variables linear. The multiple logistic regression model is expressed as follows:

$$\mathrm{logit}(p) = \ln[p/(1-p)] = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n \quad (1)$$

where p is the probability that an event will occur and $x_1 \ldots x_n$ are the variables explaining the probability. The predicted probabilities are calculated by transforming back to the original scale:

$$p = \exp(\mathrm{logit}(p))/[1 + \exp(\mathrm{logit}(p))]. \quad (2)$$

For selecting the independent variables in the models, stepwise logistic regression was applied with both forward and backward directions. The maximum number of steps to be considered was 1000 and the used multiple of the number of degrees of freedom for the penalty was log (n).

***k*-MSN**

Nonparametric estimation methods are one alternative for predicting stand-thinning maturity if there is thinning maturity determined in the ground truth. The nonparametric estimation method applied in practical forest stand characteristic inventory in Finland is *k*-MSN (Maltamo et al. 2006). In *k*-MSN, the similarity is based on canonical correlations and the Mahalanobis distance (Moeur & Stage, 1995). The benefit of the MSN method is that the similarity measure can be solved analytically. The *k*-MSN method is the same as MSN except that it takes the *k* nearest observations into account. The R yaImpute library (Crookston & Finley, 2007) was applied in the *k*-MSN estimations.

Before *k*-MSN estimation, linear transformations were done for all the ALS features; the transformations included $x^2$, sqrt(x), 1/x and log(x). Automatic feature selection was then carried out, using the simple genetic algorithm (GA) presented by Goldberg (1989) and implemented in the R GALGO library (Trevino & Falciani, 2006). The GA process begins by generating an initial population of strings (chromosomes or genomes) that consists of separate features (genes). The strings evolve during a user-defined number of iterations (generations). The evolution includes the following operations: selecting strings for mating using a user-defined objective criterion (better if more copies are in the mating pool), letting the strings in the mating pool swap parts (crossing over), causing random noise (mutations) in the offspring (children) and passing the resulting strings to the next generation. GA was used previously in ALS feature selection for nonparametric estimation with promising results (Holopainen et al., 2008).

**Prediction of stand-thinning maturity**

Predictions were carried out for two different practical scenarios:

1) to separate stands in which thinning was proposed for the next 10-year interval from stands without a thinning proposal.
2) to separate stands in which thinning should be done immediately from stands with nonurgent or without any thinning proposal in the next 10-year interval

Stand-thinning proposals were coded as a binary variable. In the first case, the binary variable was defined as one (1) if the thinning was proposed to be done during the next 10 years in the field or zero (0) if no thinning was proposed for the next 10 years. In the second case, the binary variable was defined as one (1) if thinning was proposed as immediate and (0) if thinning was proposed for the next 10 years, but was not urgent or there were no proposed thinnings for the next 10 years. If the predicted probability of the thinnings was over 0.5 it was interpreted as being necessary. Classification accuracy was determined with a test dataset, and the thinning proposals determined in the field were used as references. The accuracy of the predictions was evaluated by calculating the classification accuracy percentage and kappa value.

## 3. Results

The logistic regression models were constructed and *k*-MSN estimation was performed to predict the probability of the stand-thinning maturity and to classify the stand-thinning maturity phase. The proposed stand-thinning need during the next 10 years was predicted correctly with an overall accuracy of 79% (kappa = 0.58) with logistic regression. In the model, the ALS-derived statistical features Hmax, Hmean, h50, h60, p20 and p70 best accounted for the probability of thinning (Table 3). The importance of stand's main tree species and site-class information acquired from the field measurements was tested by including them in the preliminary variables from the field measurements. Still, they were not selected for the final model. The GA selected eight features for *k*-MSN estimation. The *k*-MSN classified 70% (kappa = 0.40) of the stands in this case correctly.

Computationally determined thinning maturity is based on stand HDOM and BA. The logistic regression model based on treewise-measured BA, HDOM, site type and tree species as independent variables was constructed for comparison with ALS-feature-based estimations (Table 3). With the model based on accurate characteristics measured in the field, the proposed stand-thinning need during the next 10 years was predicted correctly with an overall accuracy of 66% (kappa = 0.30). The ALS-feature-based logistic regression model worked significantly better.

The need for immediate stand thinning was detected correctly with an overall accuracy of 83% (kappa = 0.4) using a logistic regression model based on ALS features. In the logistic regression model the ALS-derived features penetration and h90 best accounted for the probability of immediate thinning (Table 4). Slighlty better results were achieved with *k*-MSN estimation. It classified 86% (kappa = 0.55) of the immediate thinning stands correctly. It should be noted that in this case, results with the same accuracy level as the ALS feature-based estimations were obtained with a logistic model using field-measured BA, HDOM and tree species as independent variables (Table 4). If these variables are modelled first as in practical forest inventory, slightly inaccurate results would probably be obtained.

Table 3. Predictors used in logistic regression models and *k*-MSN estimation for the probability that the stand would need thinning during the next 10 years (n = 100). The symbol "•" means that the feature is used in *k*-MSN estimation.

| **Thinnings in next 10 years** | Logistic regression (ALS) | | | | Logistic regression (Field) | | | | *k*-MSN |
|---|---|---|---|---|---|---|---|---|---|
| **Predictor** | **Estimate** | **SE** | **z value** | **Pr(> \|z\|)** | **Estimate** | **SE** | **z value** | **Pr(> \|z\|)** | |
| intercept | 13.84 | 2.86 | 4.83 | 0.00 | 1.49 | 0.71 | 2.10 | 0.04 | |
| Hmax | 0.67 | 0.12 | 5.48 | 0.00 | | | | | |
| Hmean | -1.86 | 0.40 | -4.67 | 0.00 | | | | | |
| h50 | 2.19 | 0.61 | 3.61 | 0.00 | | | | | |
| h60 | -1.54 | 0.55 | -2.79 | 0.01 | | | | | |
| h20 | -5.83 | 1.25 | -4.66 | 0.00 | | | | | |
| h70 | -11.76 | 3.00 | -3.92 | 0.00 | | | | | |
| pine | | | | | -1.70 | 0.37 | -4.57 | 0.00 | |
| spruce | | | | | -0.55 | 0.47 | -1.18 | 0.24 | |
| BA | | | | | 0.12 | 0.02 | 5.05 | 0.00 | |
| HDOM | | | | | -0.13 | 0.04 | -3.18 | 0.00 | |
| penetration | | | | | | | | | • |
| p10 | | | | | | | | | • |
| 1/CV | | | | | | | | | • |
| ln h80 | | | | | | | | | • |
| 1/h80 | | | | | | | | | • |
| √h80 | | | | | | | | | • |
| √p40 | | | | | | | | | • |
| 1/p90 | | | | | | | | | • |
| **Classification accuracy** | 79 % | | | | 66 % | | | | 70 % |
| **Kappa-value** | 0.58 | | | | 0.30 | | | | 0.40 |

Computational thinning proposals are determined based on stand BA and HDOM. Figure 3 shows that even some stands that had immediate thinning proposed in the field are far from the thinning curve, mostly due to the uneven spatial distribution of trees. Logistic regression based on ALS statistical features provided more accurate predictions in these cases.

Table 4. Predictors used in logistic regression models and *k*-MSN estimation of the probability that a stand needs immediate thinning (n = 100). Symbol "•" means that the feature is used in *k*-MSN estimation.

| **Immediate thinnings** | Logistic regression (ALS) | | | | Logistic regression (Field) | | | | *k*-MSN |
|---|---|---|---|---|---|---|---|---|---|
| **Predictor** | **Estimate** | **SE** | **z value** | **Pr(> \|z\|)** | **Estimate** | **SE** | **z value** | **Pr(> \|z\|)** | |
| intercept | -0.57 | 1.01 | -0.56 | 0.57 | -3.75 | 1.00 | -3.76 | 0.00 | |
| penetration | -10.31 | 1.53 | -6.73 | 0.00 | | | | | • |
| h90 | 0.15 | 0.05 | 2.67 | 0.01 | | | | | |
| pine | | | | | -1.99 | 0.44 | -4.49 | 0.00 | |
| spruce | | | | | -1.13 | 0.51 | -2.21 | 0.03 | |
| BA | | | | | 0.24 | 0.04 | 6.57 | 0.00 | |
| HDOM | | | | | -0.08 | 0.06 | -1.43 | 0.15 | |
| h30 | | | | | | | | | • |
| p10 | | | | | | | | | • |
| p60 | | | | | | | | | • |
| √penetration | | | | | | | | | • |
| $h30^2$ | | | | | | | | | • |
| ln h30 | | | | | | | | | • |
| √h30 | | | | | | | | | • |
| √p10 | | | | | | | | | • |
| $p50^2$ | | | | | | | | | • |
| $p60^2$ | | | | | | | | | • |
| **Classification accuracy** | 83 % | | | | 83 % | | | | 86 % |
| **Kappa-value** | 0.40 | | | | 0.40 | | | | 0.55 |

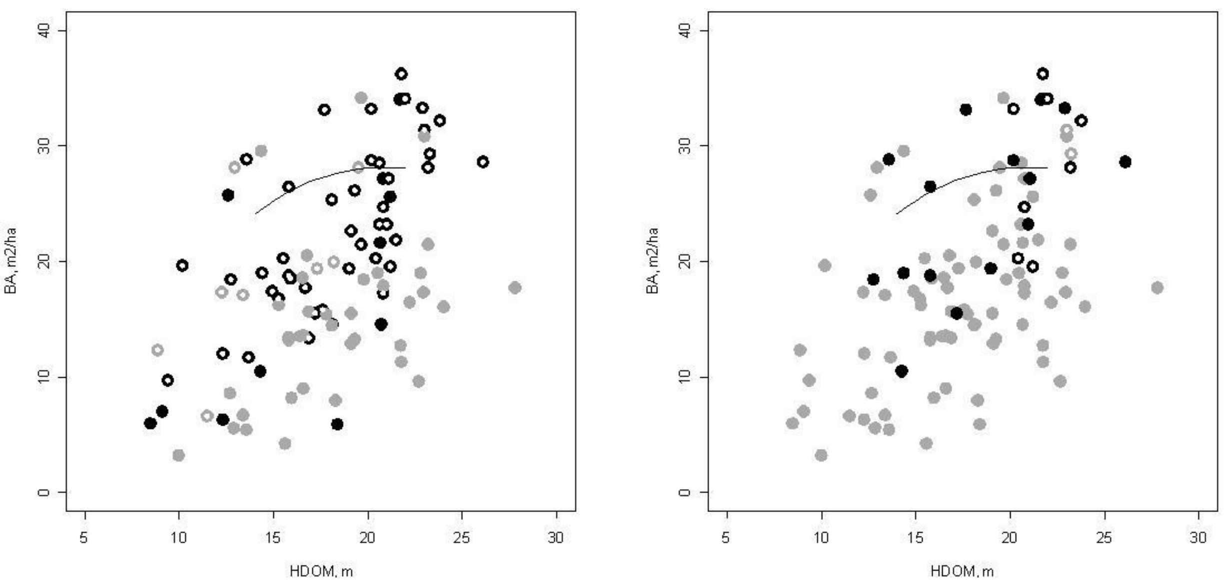


Figure 3. Left: Black circles describe test stands (n=100) in which thinning was proposed for the next 10 years. Grey circles are plots with no thinning need. If a white dot is inside the circle, thinning was predicted with logistic regression. Right: Stands that need thinning immediately. Symbol logic is the same as in left. The lower limit of thinning based on BA and HDOM (Recommendations of Tapio…, 2006) for moist sites (Pine and Spruce) is also plotted.

Figure 4 shows the plotted ALS features that describe the proportional canopy covers at 20%, 40%, 60% and 80% relative heights (lower values mean denser canopy). Clear negative correlations (cor = -0.78, cor = -0.55) between the predicted probability of the thinning and the canopy density can be seen, especially at the relative heights of 20% and 40%.

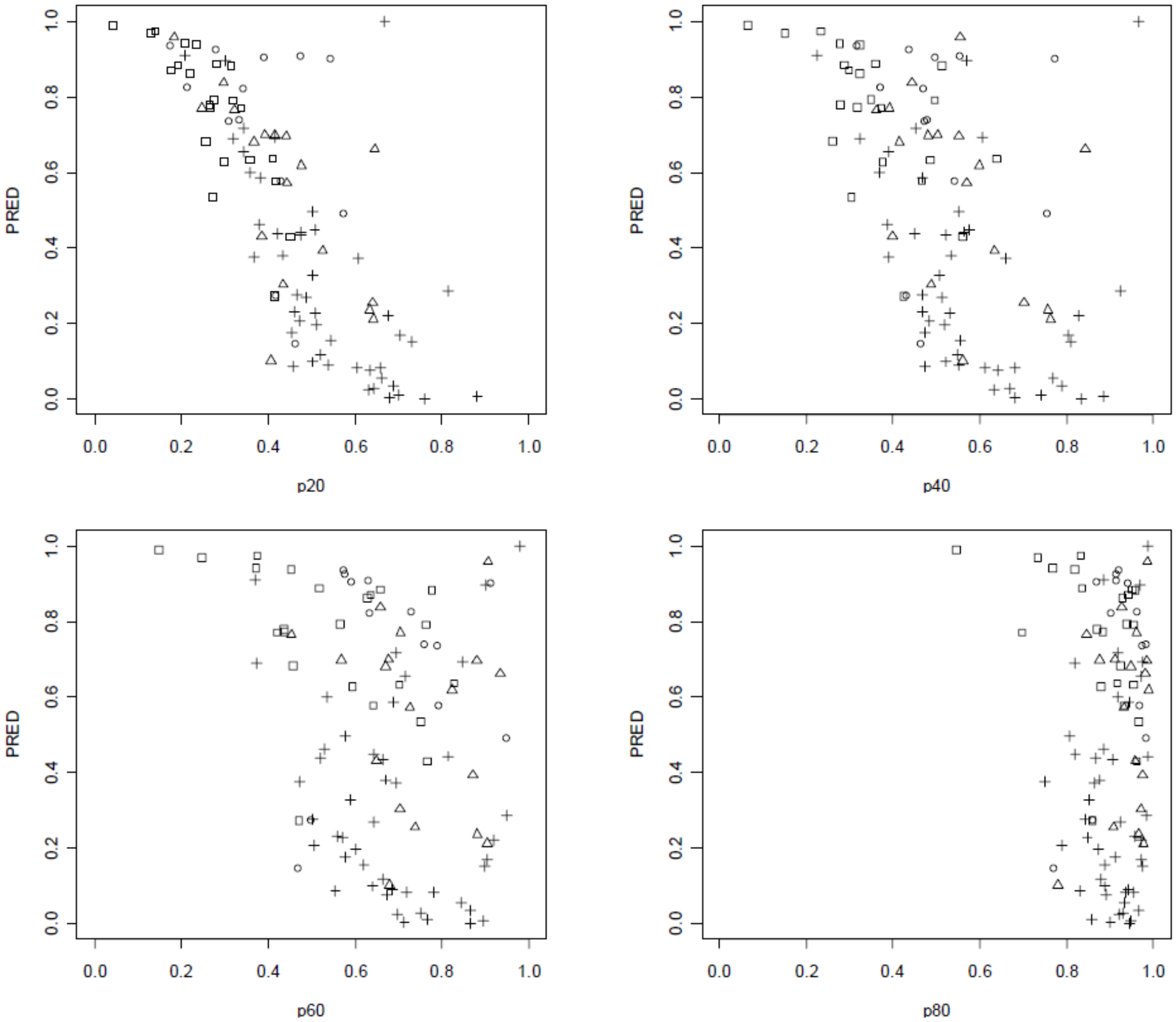


Figure 4. Proportional canopy covers in test stands (n=100) plotted against the predicted probability of thinning (PRED) during the next 10 years. Logistic regression was used for predictions. Stand's field thinning proposals were immediate (□), 1-5 years (○), 6-10 years (Δ) or rest (+) stands.

## 4. Discussion

In the present study, logistic regression and *k*-MSN estimation were used to predict stand-thinning maturity. The constructed logistic regression models resulted in classification accuracies ranging from 79% to 83% for predicting timing of next thinning, while the accuracies with *k*-MSN ranged from 70% to 86%. ALS data collected for practical forest resource inventory could be used in the manner presented to enhance the quality of computationally produced thinning proposals or to locate stands that should be checked in the field. If this type of procedure is used in practice, proposals for thinnings or other forest management actions should be collected in the field at the same time as the ground truth data for estimation of stand characteristics.

All the thinning proposals used as references were determined in the field at plot level by a single expert and the decisions were based on silvicultural recommendations. In practice, forest management actions are always carried out at the stand (compartment) level. In forest inventory, plot-level field measurements are generalized to a grid or segments, and stand-level characteristics are calculated from grid cells (segments) that are within stand boundaries. Forest management actions, such as thinning proposals, could be generalized in the same way. The accuracy and usability of this procedure need further investigation.

The ALS features are highly correlated with the forest canopy structure and thus provide information from both the silvicultural and economic points of view, when assessing the thinning maturity. In this study, correlations between the predicted probability of the thinning and the canopy density can be seen, especially of the relative heights of 20% and 40%. This strengthens our hypothesis that ALS provides information that can be utilized in the prediction of forest operations. Therefore direct models could also be developed to predict thinning maturity. A better understanding of the capabilities of ALS could provide means of removing subjective field references or oversimplified computational thinning proposals.

The results here are comparable with the study by Hyvönen (2002) and particularly when a stand’s operational need during the next 10 years is predicted. In the present study the classification accuracies were 79%, 70% and 66% with the logistic ALS model, *k*-MSN and logistic field model, respectively, while Hyvönen achieved an accuracy of 64.1%. It should be noted that Hyvönen used satellite images as auxiliary data, operated at the stand level, and that the reference and test sites were located in different areas. Our study has demonstrated the feasibility of utilizing ALS data for predicting stand-thinning maturity. Although ALS data are far more expensive auxiliary data than satellite images, they are beginning to be widely available, at least in Finland, when ALS-based inventory is applied. Närhi et al. (2008) also used ALS-features in classifying stand’s precommercial thinning maturity with an overall accuracy of 71.8%. The results of their study are in line with those achieved here. However, precommercial thinnings were not examined in this study. In general, ALS-based prediction of forest management proposals could provide a practical future means of locating stands having some operational need.

Forest stand operational need is always derived from decision maker’s objectives. Society’s needs are described in the Recommendations of Tapio for good silviculture (2006). The lower limits of thinning curves maintain growing stock at sustainable levels since self-thinning is avoided with upper limits and saw wood acquisition is ensured by regeneration diameter limits (Haara & Korhonen, 2004). In the present study, reference thinning proposals were determined in a practical way, based on the Recommendations of Tapio for good silviculture (2006). In this procedure, the forest owner is assumed to maximize the NPV of the forest. The decision maker’s alternative objectives can be taken into account when field proposals are determined.

Although tree species and site classes were tested as preliminary variables for ALS-feature-based logistic regression models, they were not included in the final models. This may have resulted from the dataset used. The majority of the thinning or advanced thinning stands used in this study were spruce- or pine-dominant and located in the same site type (moist site); e.g. thinning curves are similar for pine and spruce stands in moist sites.

The ALS-based predictions outperformed models that only included variables from the field measurements in locating thinning stands, although BA and HDOM are measured without significant errors in these kinds of data. When the immediate thinning maturity was modelled, both methods provided accurate results. In logistic regression, the result is a predicted probability of an event to occur. In the present study this probability was simplified into two classes. If the probability was over 0.5 the event (thinning) occurred, and vice versa. We noted that stands in which the probability was close to 0 or 1 were classified more accurately than the overall classification accuracy. The predicted probability of the thinning to occur could be used in locating thinning stands with the required confidence level.

A new practical procedure for predicting stand-thinning maturity was presented in this study with promising results. Further research is needed to test the method at stand level and also to develop models that could utilize single-tree-level information. Our results can be used in linking ALS-based forest inventory with practical forest management planning.

**References**

Axelsson, P. (2000). DEM generation from laser scanner data using adaptive TIN models. *International Archives of Photogrammetry and Remote Sensing, 33*(B4), 110–117.

Crookston, N. & Finley, A. (2007). yaImpute: An R Package for k-NN Imputation. *Journal of Statistical Software, 23*(10), 1–16.

Eid, T., Gobakken, T. & Næsset, E. (2004). Comparing stand inventories for large areas based on photo interpretation and laser scanning by means of cost-plus-loss analyses. *Scandinavian Journal of Forest Research* 19: 512–523.

Goldberg, D.E. (1989). Genetic algorithms in search, optimization, and machine learning. Addison-Wesley Publishing Company, Reading, Massachusetts, 412 p.

Haara, A. & Korhonen, K.T. (2004). Toimenpide-ehdotusten tuottaminen laskennallisesti ajantasaistetusta kuvioaineistosta [Determination of forest management actions from updated standwise field inventory data]. *Metsätieteen aikakauskirja,* 2/2004, 157–173 (in Finnish).

Hirata, Y., Furuya, N., Suzuki, M. & Yamamoto, H. (2009). Airborne laser scanning in forest management: Individual tree identification and laser pulse penetration in a stand with different levels of thinning. *Forest Ecology and Management, 258,* 752–760.

Holmgren, J., Nilsson, M. & Olsson, H. (2003). Estimation of tree height and stem volume on plots using airborne laser scanning. *Forest Science, 49,* 419–428.

Holopainen, M. & Talvitie, T. (2006). Effects of data acquisition accuracy on timing of stand harvests and expected net present value. *Silva Fennica, 40*(3), 531–543.

Holopainen, M., Haapanen, R., Tuominen, S. & Viitala, R. (2008). Performance of airborne laser scanning- and aerial photograph-based statistical and textural features in forest variable estimation. In Hill, R., Rossette, J. and Suárez, J. 2008. Silvilaser (2008) proceedings:105–112.

Holopainen, M., Mäkinen, A., Rasinmäki, J., Hyyppä, J., Hyyppä, H., Kaartinen, H., et al. (2010). Effect of tree level airborne laser scanning accuracy on the timing and expected value of harvest decisions. *European Journal of Forest Research,* Doi 10.1007/s10342-009-0282-6.

Hopkinson, C. & Chasmer, L. (2009). Testing LiDAR models of fractional cover across multiple forest ecozones. *Remote Sensing of Environment, 113*, 275–288.

Hyvönen, P. (2002). Kuvioittaisten puustotunnusten ja toimenpide-ehdotusten estimointi *k*-lähimmän naapurin menetelmällä Landsat TM -satelliittikuvan, vanhan inventointitiedon ja kuviotason tukiaineiston avulla [Estimation of stand characteristics and forest management operations with *k*-NN using Landsat TM satellite images, stand register and standwise field inventory as auxiliary data]. *Metsätieteen aikakauskirja,* 3/2002: 363–379 (in Finnish).

Hyvönen, P. & Anttila, P. (2006). Change detection in boreal forests using bi-temporal aerial photographs. *Silva Fennica* 40(2): 303–314.

Hyvönen, P., Heinonen, J. & Haara, A. (2010). Detection of forest management operations using bi-temporal aerial photographs. In: Wagner, W & Székely, B. (eds.). ISPRS Technical Commission VII Symposium - 100 Years ISPRS Advancing Remote Sensing Science. Vienna, Austria, July 5-7, 2010. International Archives of Photogrammetry and Remote Sensing XXXVIII, part 7B: 309-313.

Hyyppä, H., & Hyyppä, J. (1999). Comparing the accuracy of laser scanner with other optical remote sensing data sources for stand attribute retrieval. *The Photogrammetric Journal of Finland,* 16, 5–15.

Jensen, J., Humes, K., Vierling, L. & Hudak, A.T. (2008). Discrete return lidar based prediction of leaf area index in two conifer forests. *Remote Sensing of Environment, 112*, 3947–3957.

Magnussen, S. & Boudewyn, P. (1998). Derivations of stand heights from airborne laser scanner data with canopy-based quantile estimators. *Canadian Journal of Forest Research*, 28, 1016–1031.

Magnussen, S., Eggermont, P. & LaRiccia, V. (1999). Recovering tree heights from airborne laser scanner data. *Forest Science,* 45, 407–422.

Maltamo, M., Malinen, J., Packalén, P., Suvanto, A. & Kangas, J. (2006). Nonparametric estimation of stem volume using airborne laser scanning, aerial photography, and stand-register data. *Canadian Journal of Forest Research, 36*, 426–436.

Means, J., Acker, S., Harding, D., Blair, J., Lefsky, M., Cohen, W. (1999). Use of large-footprint scanning airborne lidar to estimate forest stand characteristics in the western cascades of Oregon. *Remote Sensing of Environment*, 67, 298–308.

Moeur, M. & Stage, A.R. (1995). Most similar neighbor: an improved sampling inference procedure for natural resource planning. *Forest Science, 41*, 337–359.

Næsset, E. (1997). Determination of mean tree height of forest stands using airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing, 52*, 49–56.

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

Närhi, M., Maltamo, M., Packalén, P., Peltola, H. & Soimasuo, J. 2008. Kuusen taimikoiden inventointi ja taimikonhoidon kiireellisyyden määrittäminen laserkeilauksen ja metsäsuunnitelmatietojen avulla [Inventory and determination of forest management operations for spruce seedling stands using ALS and stand register data.] *Metsätieteen aikakauskirja* 1/2008, 5–15 (in Finnish).

Ojansuu, R., Halinen, M. & Härkönen, K. (2002). Metsätalouden suunnittelujärjestelmän virhelähteet männyn esiharvennuskypsyyden määrittämisessä [Error sources in determination of Scots pine stand first thinning maturity incurred from forest management calculations system]. *Metsätieteen aikakauskirja,* 3/2002, 441–457 (in Finnish).

Pesonen, A., Korhonen, K.T., Tuominen, S., Maltamo, M. & Lukkarinen, E. (2007). Taimikonhoitotarpeen arviointi valtakunnan metsien inventoinnin metsävarakartan pohjalta [Assessment of forest management operatonal need in seedling stands using thematic maps made from national forest inventory data]. *Metsätieteen aikakauskirja* 2/2007, 77–86 (in Finnish).

R Development Core Team. (2007). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL www.R-project.org

Recommendations of Tapio for good silviculture. (2006). Metsäkustannus 100 p. (in Finnish).

Trevino, V. & Falciani, F. (2006). GALGO: an R package for multivariate variable selection using genetic algorithms. *Bioinformatics, 22*(9), 1154–1156.

Vanhatalo, K. (2010). Modelling the probability of incorrect harvest decisions due to errors of stand characteristics. Manuscript for Master’s thesis. University of Helsinki, Department of Forest Sciences (in Finnish, abstract in English).

Varjo, J. (1996). Controlling continuously updated forest data by satellite remote sensing. *International Journal of Remote Sensing*,17(1), 43 – 67

Vastaranta, M. (2006). Accuracy in updating stand level data. Master’s thesis. University of Helsinki, Department of Forest Resource Management. 61 p. (in Finnish, abstract in English).