RESEARCH ARTICLE

Estimating stem diameter distributions from airborne laser scanning data and their effects on long term forest management planning

Rami Saad [a*], Jörgen Wallerman [a], Tomas Lämås [a]

[*]Corresponding author. Email: rami.saad@slu.se

[a] *Department of Forest Resource Management, Swedish University of Agricultural Sciences, Umeå, Sweden*

Abstract

Data obtained from airborne laser scanning (ALS) are frequently used for acquiring forest data. Using a relatively low number of laser pulses per unit area (≤ 5 pulses per m^2), this technique is typically used to estimate stand mean values. In this study stand diameter distributions were also estimated, with the aim of improving the information available for effective forest management and planning. Plot level forest data, such as stem number and mean height, together with diameter distributions in the form of Weibull distributions, were estimated using ALS data. Stand-wise tree lists were then estimated. These estimations were compared to data obtained from a field survey of 124 stands in northern Sweden. In each stand an average of seven sample plots (radius 5-10 m) were systematically sampled. The ALS approach was then compared to a mean value approach where only mean values are estimated and tree lists are simulated using a forest decision support system (DSS). The ALS approach provided a better match to observed diameter distributions: ca. 35% lower error indices used as a measure

of accuracy and these results are in line with the previous studies. Moreover – which is unique compared to earlier studies – suboptimal losses were assessed. Using the Heureka DSS the suboptimal losses in terms of net present value due to erroneous decisions were compared. Although no large difference was found, the ALS approach showed smaller suboptimal loss than the mean value approach.

Keywords: forest management planning, suboptimal loss, Weibull distribution, Airborne Laser Scanning, Heureka, decision support system

Introduction

In forest planning, different potential management actions are analyzed and the actions best fulfilling stated goals are chosen by the forest owner or a decision maker. The analyses and decisions are based upon various characteristics of the particular stands within a forest property such as timber volume, basal area and mean tree height. These forest variables are used as inputs in decision support systems (DSS), such as the Swedish Heureka system (Wikström et al. 2011), to simulate and evaluate different possible treatments. The outcome from these systems is a management proposal for each individual forest stand, which aims to maximize the utility of the forest holding. Utility is often expressed as an economic yield, typically in terms of net present value (NPV) within a set of constraints based on, e.g., timber flows and environmental factors.

Naturally the accuracy of forestland data affects the scope for efficient management planning, therefore evaluating the quality of the available information is a critical step in forest management (Kangas 2010). In general statistical terms the quality of the data is defined as how far the available data are from the true value (accuracy). The forest information is usually gathered by sample-based surveying, visual estimations (ocular standwise field inventory) or remote sensing techniques such as airborne laser scanning (ALS) (McRoberts *et al.* 2010). Estimates gathered by visual estimation tends to include both random and systematic errors, while estimates from sample based surveys remote sensing can be expected to contain random errors only (estimates based on remote sensing data may contain systematic errors from different factors such as model lack of fit). Loss occurring from suboptimal decisions due to erroneous estimates is defined as the difference between NPV based on accurate data and that based on erroneous estimates on the same forest (Holmström et al. 2003). A method for maximizing the utility of available data is cost-plus-loss analysis, in which the accuracy level is chosen such that it minimizes the sum of direct inventory costs and the losses resulting from inaccurate data (Kangas 2010).

Forest information compiled in stand register databases tends to consist of stand-level values such as stem number, mean age and mean tree size. Given that DSSs typically use individual tree models in their calculations, models are required to simulate tree lists from the stand mean values contained in the register databases, as with the Heureka system. It is of interest to use directly estimated tree list data, such as those obtained from sample plot surveys, in order to avoid the inherent approximations involved in simulating tree lists from stand mean values.

The development of forest DSSs is an active research area, one example being the Heureka system (Borges et al. 2014; Gordon et al. 2013), which was developed at the Swedish University of Agricultural Sciences (SLU). It enables long term planning, analysis and management of forestland, and is used in this study. In the planning procedure Heureka is used to maximize a goal stated by the user, such as maximum NPV, subject to economic and environmental restrictions. Forest information (forest variables), either in terms of stand mean values (basal area, number of stems, mean diameter and height etc.), or as individual tree data, needs to be imported into the Heureka system in order to compute the NPV of different treatments.

The topic forest information quality was studied in recent papers and found to be essential in the process of forest management decision making. Inaccurate estimates lead to wrong management actions and timing of actions, which will lead to economic losses. Nevertheless, Duvemo & Lämås (2006) found that the quality of forest information had received relatively little attention, compared to other aspects of forest planning, owing to the complexity of the associated problems. They also found that evaluations of forest information quality are typically based on overly simplistic assumptions. Kangas (2010) emphasize the complexity of the subject and suggests methods, such as Bayesian decision theory, to improve the use of the available forest information.

ALS is presently widely used to capture high-quality information for forest management planning (Gobakken & Næsset 2004; Næsset *et al.* 2004; McRoberts *et al.* 2010). This is generally found to outperform traditional sources of information for management planning.

Today, nation-wide ALS campaigns have been conducted or are about to be initiated in countries such as Denmark, Switzerland, the Netherlands, Finland, and Sweden. The Swedish government decided in 2008 to finance the production of a new and highly accurate national Digital Elevation Model. The production is carried out between 2009 and 2013 by the Swedish National Land Survey (Lantmäteriet), using ALS operated by several private sub-contractors using various scanning systems. This will provide ALS data for all forested parts of Sweden at a low cost. ALS data can be used to estimate stand variables, both as stand mean values (area based method) and data for individual trees. In general the area based method uses a low number of laser pulses per area unit (≤5 pulses per m^2 (Næsset 2002)) and in the case of individual trees a higher number of laser pulses per area unit (typically >5 pulses per m^2 are used to detect individual trees and for estimating individual tree variables (e.g. Solberg et al. 2006; Breidenbach et al. 2010).

Besides estimating stand mean values using area based method there have been attempts to estimate stand diameter distributions, for example by Næsset (2004) and Gobakken & Næsset (2004). Gobakken & Næsset (2004) divided the forest area into strata according to age class and site quality. Weibull diameter distribution was estimated for each stratum. The area based method was used to relate the ALS information to the Weibull distribution parameters. Gobakken & Næsset (2005) used ALS information in order to compare the accuracy of estimating basal area that was assessed by parameter recovery of a two parameter Weibull distribution and a system of 10 percentiles of the observed diameter range, the latter approach being a non parametric method. Non parametric methods have also been used by, e.g., Gobakken & Næsset (2005) and Maltamo et al. (2009). Using this approach no assumptions are made regarding the diameter distribution. Imputation techniques such as the kMSN method are considered to be non parametric method for estimating diameter distributions (Maltamo et al. 2009).

In order to analyze the usefulness of diameter distributions estimated from ALS data three alternatives were used in this study. The first alternative was acquired through a sample plot field survey of 124 stands. The second alternative contained estimates based on ALS information. Using the area based method both a set of mean values, such as basal area and stem number, and diameter distributions, were estimated per plot. Based on the second alternative stand mean values were estimated to correspond to data in a traditional stand register and made up the third alternative. Both the first and second alternatives contained tree lists per plot, which were used in the subsequent DSS calculations. From the mean values in the third alternative tree lists were simulated in the DSS using built in functions. Suboptimal losses due to non-perfect data in the second and third alternatives were then estimated.

The purpose of the study was to estimate diameter distributions using ALS information and – which is unique compared to earlier studies – to determine if these distributions notably improved decision making in terms of reduced suboptimal losses compared to traditional methods of simulating tree lists from stand mean values. As ALS information can now be acquired cheaply and highly accurately for some stand level variables, such as tree height, basal area and timber volume, ALS approaches are often preferable to traditional ocular data acquisition methods. Use of ALS should therefore reduce losses from suboptimal decisions, since the quality of information is critical for good decision making. The results of the study indicate that ALS-based estimates of diameter distributions have the potential to further improve the process, although the gain in NPV was not very high. The study focused on long-term (strategic) planning, hence details such as distributions of timber assortments in the near future, which are typically of interest in tactical planning and also affected by diameter distribution estimations, are not considered.

Material and methods

Forest area and field survey

The study was performed in a managed boreal forest landscape in northern Sweden (64°06’N, 19°10’E, 245 – 320 m.a.s.l. owned by the state owned forest company Sveaskog. The forest landscape is dominated by Norway spruce (*Picea abies* (L.) Karst.) and Scots pine (*Pinus silvestris* (L.)), birch (*Betula* spp) being the most frequent broad-leaved species. A field survey was performed in 2008 and 2009 in which all stands where surveyed using 2 - 15 (mean 7.33) circular sample plots in each stand (except of one stand that was represented by one plot). The sample plots were located in a systematic grid in each stand. Geographic position of each plot was determined using post-processed differential GPS with an expected accuracy of less than 1 m. Sapling and young stands were also inventoried, however not used in this study. Plots that did not include any trees were removed. Plot radii for the stands included were 10 m (117 stands) and 5 m (7 stands). On the plots stem diameter at breast height (1.3 m above the ground) and species were registered for all trees. The stem diameter at breast height and species of all trees on the plots were registered. The height and age of at least three trees on each plot (typically the two largest diameter trees and one randomly selected tree) were also registered.

“<Table 1 here>”

Airborne laser scanning

Strömsjöliden was scanned using the ALS system TopEye (S/N 425) carried by a helicopter in the 3rd and 5th of August 2008, operated by the contractor Blom Sweden AB. Flying height was 500 m above ground and the mission measured approximately 5 pulses per m^2. The point data were classified using a progressive Triangular Irregular Network (TIN) algorithm (Axelsson 1999) and (Axelsson 2000) to estimate which returns are measurements of the ground level. Following this, the height above ground was determined for all returns, using a digital elevation model produced from the classified ALS data. A set of fundamental ALS metrics were then computed from the ALS data in accordance to the area based method (Næsset 2002); metrics corresponding to the elevation information, as well as the density of the vegetation, see Table 2. A cut-off value of 1.0 m was applied for calculation of metrics.

“<Table 2 here>”

Three studied alternatives

Three alternatives were used in the study. The first alternative was comprised of the field survey observations. The second alternative was based on the ALS metrics. Stand mean values estimated from the second alternative that corresponds to traditional stand register information made up the third alternative, termed later as the mean values alternative, see Fig. 1. Tree lists estimated from the ALS alternative and simulated in the DSS in the mean values alternative were assumed to have diameter distributions that could be described by a two parameter Weibull function for each plot in the ALS case and per stand in the mean values case. In the ALS case each plot was tested according to Kolmogorov-Smirnoff test to measure the goodness of fit of the estimated Weibull distribution and approximately 96% (869 out of 909) of the null hypothesizes were not rejected, meaning that the diameter distributions are likely to follow the Weibull distribution assumption, see appendix 1. That is, in the ALS alternative the stand level tree list when aggregated over plots did not necessarily follow a Weibull distribution. As the mean values alternative were estimated from the ALS alternative, these two alternatives were in many parts comparable, that is, the study is not aiming at comparison of the accuracy of different forest information acquisition methods. The elaborations of the three data sets are described below, see also Fig. 1.

Observed alternative

The data acquired in the field survey of the case study area made up the observed alternative. As all trees on sample plots within the stands were callipered tree lists were available.

ALS alternative

Based on the observed alternative and the ALS data functions estimating plot level forest variables including diameter distribution were elaborated. Along with the ALS metrics also the proportion basal area of pine was used as it turned out to be an important variable. This information is typically available in stand registers.

The diameter distribution of each plot was modeled as a two parameter Weibull distribution using the following steps:

1- A Weibull distribution was fitted to the stem diameter measurements for each plot in the observed (field survey) alternative to estimate the two parameters of the distribution, namely scale and shape.

2- Multiple linear regression was used, after stepwise regression, to relate the ALS metrics and the proportion of pine from the plot sampling alternative to the scale and shape parameters estimated from the field survey alternative in step 1. In this process the scale and shape were the dependent variables, and the ALS metrics and proportion of pine were the independent variables.

3- Scale and shape parameter estimates were predicted for each plot using the regression estimation for the ALS independent variables and the proportion of pine estimated from step 2.

Expected diameter (ALS estimation) of each plot was compared with the mean diameter of the sample field survey of each plot in order to validate the estimation. Expected diameter, $\mathrm{E(D)}$, of the fitted two parameters Weibull distribution was computed as follows: D describe the diameter and it is a Weibull distributed (Hogg & Tanis 2010, page 170) random variable $\mathrm{D} \sim \mathrm{Weibull}(\lambda, \kappa)$, where λ and κ are the two parameter of Weibull distribution. Expected value of D is given by Equation (1):

(1) $E(D) = \lambda \cdot \Gamma\left(1 + \frac{1}{\kappa}\right)$,

where λ is the distribution scale, κ is the distribution shape and Γ is the gamma function $\Gamma(\mathrm{z}) = (\mathrm{z} - 1)!$, where z is a integer and the sign ! is factorial.

Values for the basal area per hectare, the number of stems per hectare, the basal area weighted mean height and the quadratic mean diameter were estimated using the ALS independent variables and the proportion pine from the observed alternative, in the same way as the scale and shape were estimated in step 3. In order to estimate these variables linear regression was employed (after applying the stepwise regression) where the dependent variables were the variables in the observed alternative and the independent variables were the ALS independent variables and the proportion pine. The variables mentioned above were predicted for each plot using the regression estimates for the ALS independent variables and the proportion pine as it was done for scale and shape in step 3. Tree species proportions per plot and site variables from the observed alternative were used when the different alternatives were imported to the Heureka DSS.

An essential step in the processing of the ALS data was the generation of tree lists. This was achieved by using the fitted Weibull distribution parameters to generate a diameter distribution for each plot, incorporating the fitted number of stems per hectare (estimated for each plot separately). One diameter value was assigned to each 10th percentile of the diameter distribution. Each percentile represented a diameter class boundary. First the basal areas corresponding to the upper and lower diameter class boundary were calculated. The diameter corresponding to the mean of the upper and lower basal area was then the diameter representing the diameter class. Each diameter that representing the diameter class, was replicated by the number of trees of each diameter class. The sum of trees over the diameter classes then made up the total number of trees on the plot.

Mean values alternative

The mean values alternative (corresponding to stand register mean values) of each stand was simply averaged from the ALS alternative. That is, the mean value alternative was derived from the ALS alternative and not the observed alternative.

“<Figure 1 here>”

Software used for calculations and handling of the different alternatives

The R Program, the free software programming language and a software environment for statistical computing and graphics, was used for calculations (regression analysis etc.) and handling of the three alternatives.

Accuracy measurement

To assess the accuracy of the estimated diameter distributions, the tree lists for each plot were first scaled, using the plot area, to obtain the number of trees per hectare in each stand separately. This was done for all three alternatives, and subsequently the estimated diameter distribution accuracy was determined using two error indices, computed for each stand separately using the diameter classes' absolute differences.

The first error index (e, Equation 2) gives one measure of the degree of the diameter distribution errors, in which the total number of the trees is taken into account. Its value can range between 0 to 200, where 0 represents a perfect match between two compared distributions.

$$(2)\ e = \sum_{j=1}^{15} e_j = 100 \cdot \sum_{j=1}^{15} \frac{|n_{oj} - n_{pj}|}{N},$$

Here, e_j is the error in diameter class j (of 15 classes from 0 to 30 cm with 2 cm increments), n_{oj} is the number of observed trees in diameter class j and n_{pj} is the number of predicted trees in diameter class j, N is the observed total number of trees. The stand level error is the sum of the diameter class errors e_j. This error index, which was first proposed by Reynolds et al. (1988), has been widely used in previous studies, e.g. Gobakken & Næsset (2004) and Gobakken & Næsset (2005).

The second error index (δ, Equation 3), termed the total variation distance index (Levin et al. 2009), measures a degree of the diameter distribution errors that is independent of the total number of trees. Each diameter class in each stand was divided by the total number of stand trees in order to obtain a diameter probability distribution. The value of index δ can range between 0 to 1, where 0 represents a perfect match of two compared distributions.

$$(3)\ \delta = \sum_{j=1}^{15} \delta_j = \frac{1}{2} \cdot \sum_{j=1}^{15} \left| P(x_j) - Q(x_j) \right|,$$

where δ_j is the error in diameter class j, $P(x_j)$ is the observed relative frequency of diameter class j, and $Q(x_j)$ is the relative frequency of diameter class j in the diameter distribution predicted by either the ALS or mean values alternatives. The error index is multiplied by ½ to scale the error between 0 and 1. $P(x_j)$ is calculated by dividing the observed number of trees in each class by the observed total number of trees in the stand. $Q(x_j)$ is calculated by dividing the number of predicted trees in each class by the predicted total number of trees in the stand. The stand level error is the sum of the diameter class errors δ_j.

Calculation of suboptimal losses

Each of the three alternatives was imported into the Heureka system (see Fig. 1). The observed alternative and ALS alternative were imported as tree lists, while Heureka simulated tree lists in the mean value alternative. This was done using functions implemented in the software that estimate the scale and shape of stands by taking into account tree species, mean stand age, tree age uniformity and quadratic mean diameter. The Heureka system simulates tree list in a similar way as the simulation tree list was done for the ALS alternative with two main differences. The first difference is that Heureka uses stand level estimated scale and shape where in the ALS alternative the estimated and fitted scale and shape were used (changed from plot to plot). The second notable difference is that Heureka takes equal diameter class intervals containing different tree numbers, while the ALS simulation uses unequal diameter classes containing equal numbers of trees.

In Heureka, a set of potential management alternatives is generated. A management alternative is a sequence over time of management actions such as regeneration, thinning and final felling. Each action has a calculated net cost or income, and a NPV is calculated for each potential management alternative. Then for each stand the alternative providing the highest NPV is selected. The optimal management strategies selected for the ALS and mean values alternatives were then applied to the forest information in the observed alternative. The differences between the NPV of the observed alternative to the NPV of the applied programs on the forest information in the observed alternative were considered to be the suboptimal losses. The applied treatment programs were fixed only for the two first periods (10 years) since it is expected that in the future new and better information is probable after a period of time (Holmström et al. 2003). The aim was to determine if losses from suboptimal decision can be decreased by using ALS estimations rather than the mean values alternative which is traditionally used in forest planning.

Results

The estimated scale and shape in the ALS alternative were used to estimate the expected diameter of trees in each plot. This was then compared with the mean diameters obtained from the field survey data to validate the ALS estimation. Figure 2 shows mean diameters and quadratic mean diameters from the survey data compared to the expected values estimated in the ALS alternative (Equation 1). Figure 2 also shows the Weibull distribution scale and shape parameters compared to the estimated values in the ALS alternative.

“<Figure 2 here>”

The regression results for six dependent forest variables, with 15 independent variables, are summarized in Table 3. The independent variables are the ALS variables as described in the Methods section and the proportion of pine from the plot sampling alternative. The independent variable Percentile70 was not included since it was found to have insignificant effects (at a significant level of 5%) on the dependent variables.

“<Table 3 here>

Calculated error indices, indicating the closeness of the estimated diameter distributions to the measured stand level diameter distributions, are summarized in Table 4.

“<Table 4 here>”

Table 4 shows that the ALS information yields smaller error indices than the mean values.

NPV results

The NPV calculated in the three alternatives and the suboptimal losses are presented in Table 5. Two different price lists were used for sensitivity analysis.

“<Table 5 here>”

NPVs were calculated using a 3% real interest rate and two different price lists. The effects of interest rate (3% vs 10%) and the growth model used (a stand growth model vs individual tree growth model (Fahlvik *et al.* 2014)) were also checked but were found to have little impact on suboptimal losses. The default price list used by Heureka, based on pulpwood and sawn timber pricings in mid-Sweden for 2013 (see Appendix 1), resulted in small suboptimal losses (see Table 5). However, as can be seen in Appendix 1, this default price list is not very sensitive to log diameters. This necessitated the construction of a hypothetical price list in which sawn timber prices increased with log diameter, following the curve for the highest log quality, and pulpwood prices were decreased by 50 percent of the mid-Sweden prices for 2013 (see Appendix 1). Use of this hypothetical pricelist increased the estimated difference in suboptimal losses, the ALS alternative yielding 111 SEK ha^{-1} smaller suboptimal losses than the mean value alternative (Table 5).

Discussion

In this study diameter distributions of stems on plots within stands were estimated from ALS information, assuming that they followed Weibull distributions, and the two parameters – scale and shape – of the distribution for each plot were estimated. Stand level tree lists were then simulated based on the plotwise diameter distributions and then imported to the Heureka forest DSS. This approach was compared to an approach were estimated stand mean values only were used and imported to Heureka. In Heureka tree lists were then simulated using inbuilt default Weibull distribution parameters corresponding to a single plot per stand but different parameters for different species. The ALS-derived tree lists yielded smaller suboptimal losses than the lists generated from stand mean values. Thus, in addition to providing robust estimates of stand characteristics such as tree height and basal area, ALS can provide valuable estimates of diameter distributions, thereby improving forest planning. Furthermore the use of error indexes also showed that the stand level ALS based tree lists was closer to the observed diameter distributions than the Heureka derived tree lists.

The use of ALS information resulted in up to 111 SEK ha^{-1} smaller suboptimal losses (using the hypothetical price list) than the mean values approach. As ALS information is already available for estimating mean values of stand characteristics, the only additional costs are in estimating the diameter distribution, thus the marginal profit can be increased by a similar amount to the suboptimal loss reduction. These results also reveal that long-term NPV calculations are substantially less sensitive to estimated diameter distributions than other factors such as volume, age, height and site index. However, diameter distributions have potentially greater impacts on short-term NPVs, for instance those related to the dimensional demands of sawmills.

In most cases the Weibull scale parameter was estimated notably more accurately than the shape parameter. This is to be expected as the area-based ALS approach will provide a low number of measurements for individual trees. It provides accurate information on the height and density of trees, but is less able to distinguish whether a forest consists of numerous thin trees, or fewer thicker trees. Estimates of the shape parameter could also be improved by higher density ALS sampling and use of larger sample plots, which would provide more accurate reference data for the subsequent modeling of diameter distributions.

In the regression modeling of diameter distribution parameters from ALS information the proportion of pine trees in each plot was used as an independent variable as well as height percentiles. The proportion of pine trees was needed as the relationship between diameter distributions and ALS data is different for different tree species. In this study, the diameter distribution of all species in each plot was modeled; in order to take the species variations into account the proportion of tree pine was included as an independent variable. In operational practice, this information cannot be estimated directly from ALS information but can be acquired by aerial photo interpretation and potentially also by computerized algorithms using aerial laser scanning data and digital aerial photos (Packalén & Maltamo 2007). A proxy for plot level pine proportion is also readily available in existing stand registers.

A potential way to further improve the approach is to use non-parametric methods to estimate plot level diameter distributions, as described by Gobakken (2005) and Maltamo et al. (2009). In such a case no parametric diameter distribution is assumed (in contrast to our assumption of Weibull distributions), and in operational applications today imputation techniques, based for instance on kMSN methods (Maltamo et al. 2009), are usually applied. In this approach, predictions are made using the actual diameter measurements in the reference data and no smoothing or distribution assumptions are needed. Such methods can be further evaluated in future studies to assess their potential for improving data to be used in forest DSSs.

In conclusion, the results of the study indicate that ALS-based estimates of diameter distributions have the potential to further improve the planning process, although in this study the gain in NPV was not very high. Use of ALS data should reduce losses from suboptimal decisions, but the level of reduction depends on, e.g., the design of timber price list.

References

Axelsson P. 1999. Processing of laser scanner data—algorithms and applications. ISPRS Journal of Photogrammetry & Remote Sensing. 54:138–147.

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

Borges JG, Nordström E.M., Garcia-Gonzalo J, Hujala T., Trasobares A (Eds.). 2014. Computer-based tools for supporting forest management. The experience and the expertise world-wide. Umeå (Sweden): Swedish University of Agricultural Sciences, Department of Forest Resource Management.

Breidenbach J, Næsset E, Lien V, Gobakken T, Solberg S. 2010. Prediction of species specific forest inventory attributes using a nonparametric semi-individual tree crown approach based on fused airborne laser scanning and multispectral data. Remote Sens. Environ. 114:911–924.

Duvemo K, Lämås T. 2006. The influence of forest data quality on planning processes in forestry. Scand. J. For. Res. 21:327–339.

Fahlvik N, Wikström P, Elfving B. 2014. Evaluation of growth models used in the Swedish Forest Planning System Heureka. Silva Fennica. 48:2.

Gobakken T, Næsset E. 2004. Estimation of diameter and basal area distributions in coniferous forest by means of airborne laser scanner data. Scand. J. For. Res. 19:529–542.

Gobakken T, Næsset E. 2005. Weibull and percentile models for lidar-based estimation of basal area distribution. Scand. J. For. Res. 20:490–502.

Gordon SN, Floris A, Boerboom L, Lämås T, Eriksson LO, Nieuwenhuis M, Garcia J, Rodriguez L. 2013. Studying the use of forest management decision support systems: An initial synthesis of lessons learned from case studies compiled using a semantic wiki. Scand. J. For. Res. DOI:10.1080/02827581.2013.856463.

Hogg RV, Tanis EA. 2010. Probability and Statistical Inference. New Jersey: Pearson.

Holmström H, Kallur H, Ståhl G. 2003. Cost-plus-loss analyses of forest inventory strategies based on kNN assigned reference sample plot data. Silva Fennica. 37:381–398.

Kangas AS. 2010. Value of forest information. Eur J Forest Res. 129:863–874.

Levin DA, Peres Y, Wilmer EL. 2009. Markov Chains and Mixing Times. American Mathematical Society. Page 48.

Maltamo M, Næsset E, Bollandsås OM, Gobakken T, Packalén P. 2009. Non-parametric prediction of diameter distributions using airborne laser scanner data. Scand. J. For. Res. 24:541–553.

McRoberts RE, Cohen WB, Næsset E, Stehman SV, Tomppo EO. 2010. Using remotely sensed data to construct and assess forest attribute maps and related spatial products. Scand. J. For. Res. 25:340–367.

Næsset E. 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. Remote Sens. Environ. 80:88–99.

Næsset E, Gobakken T, Holmgren J, Hyyppä H, Hyyppä J, Maltamo M, Nilsson M, Olsson H, Persson Å, Söderman U. 2004. Laser scanning of forest resources: the nordic experience. Scand. J. For. Res. 19:482–499.

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

Reynolds MR, Burk TE, Huang WC. 1988. Goodness-of-FiT tests and model selection procedures for diameter distribution models. For. Sci. 34:373–399.

Solberg S, Næsset E, Bollandsas OM. 2006. Single tree segmentation using airborne laser scanner data in a structurally heterogeneous spruce forest. Photogramm. Eng. Remote Sensing. 72:1369–1378.

Wikström P, Edenius L, Elfving B, Eriksson LO, Lämås T, Sonesson J, Öhman K, Wallerman J, Waller C, Klintebäck F. 2011. The Heureka forestry decision support system: an overview. Mathematical and Computational Forestry&Natural-Resource Sciences. 3:87–94.

Appendix 1.

“<Figure 1 pine default prices here>”

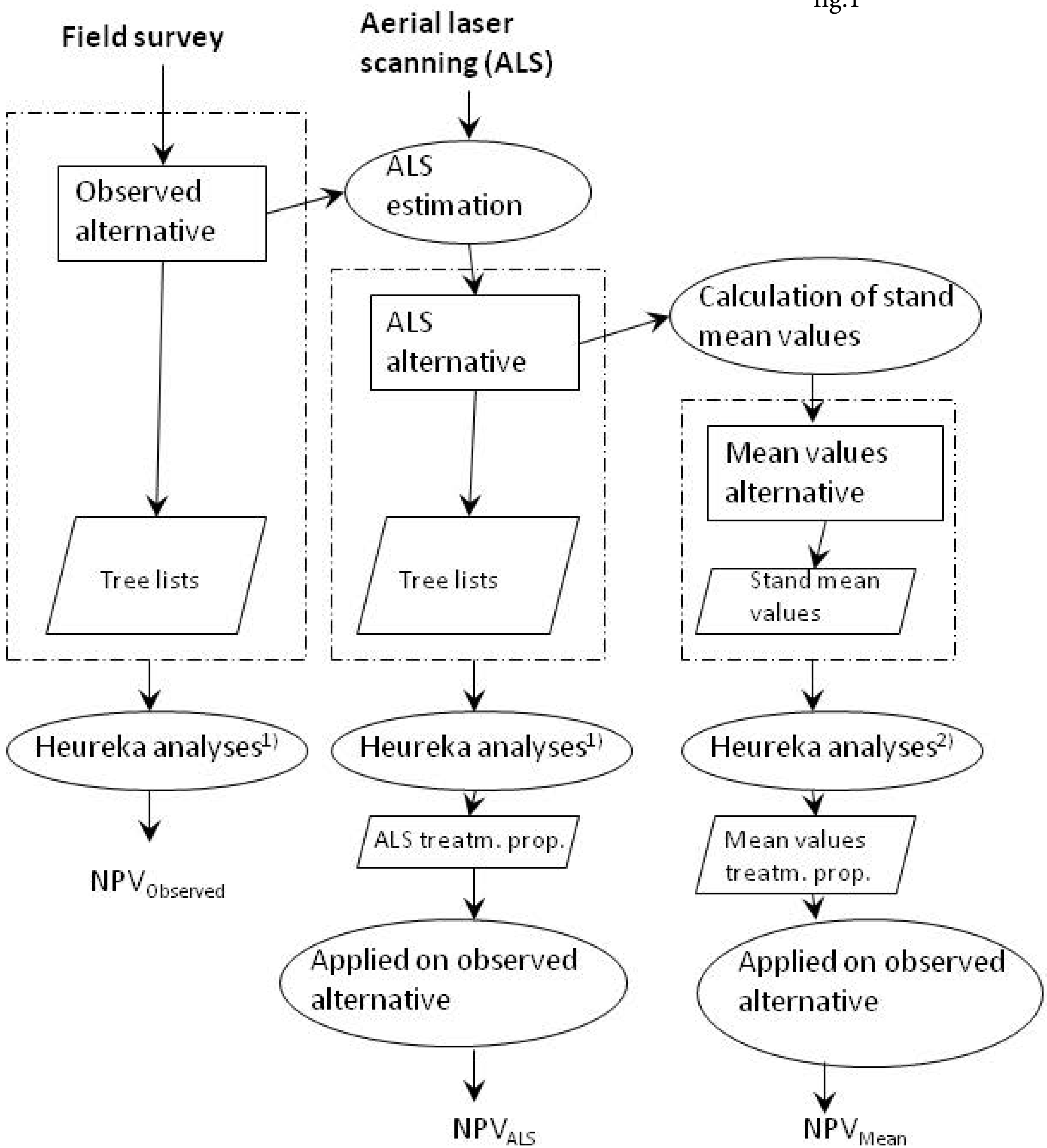
“<Figure 2 spruce default prices here>”

“<Figure 3 pine hypothetical prices here>”

“<Figure 4 spruce hypothetical prices here>”

“<Figure 5 histogram of Kolmogorov-Smirnoff statistics values here >”

fig.1



1) Includes generating treatment proposals and selection of the treatment giving the highest NPV
2) Includes simulation of tree lists with inbuilt functions, generating treatment proposals and selection of the treatment giving the highest NPV

fig.2

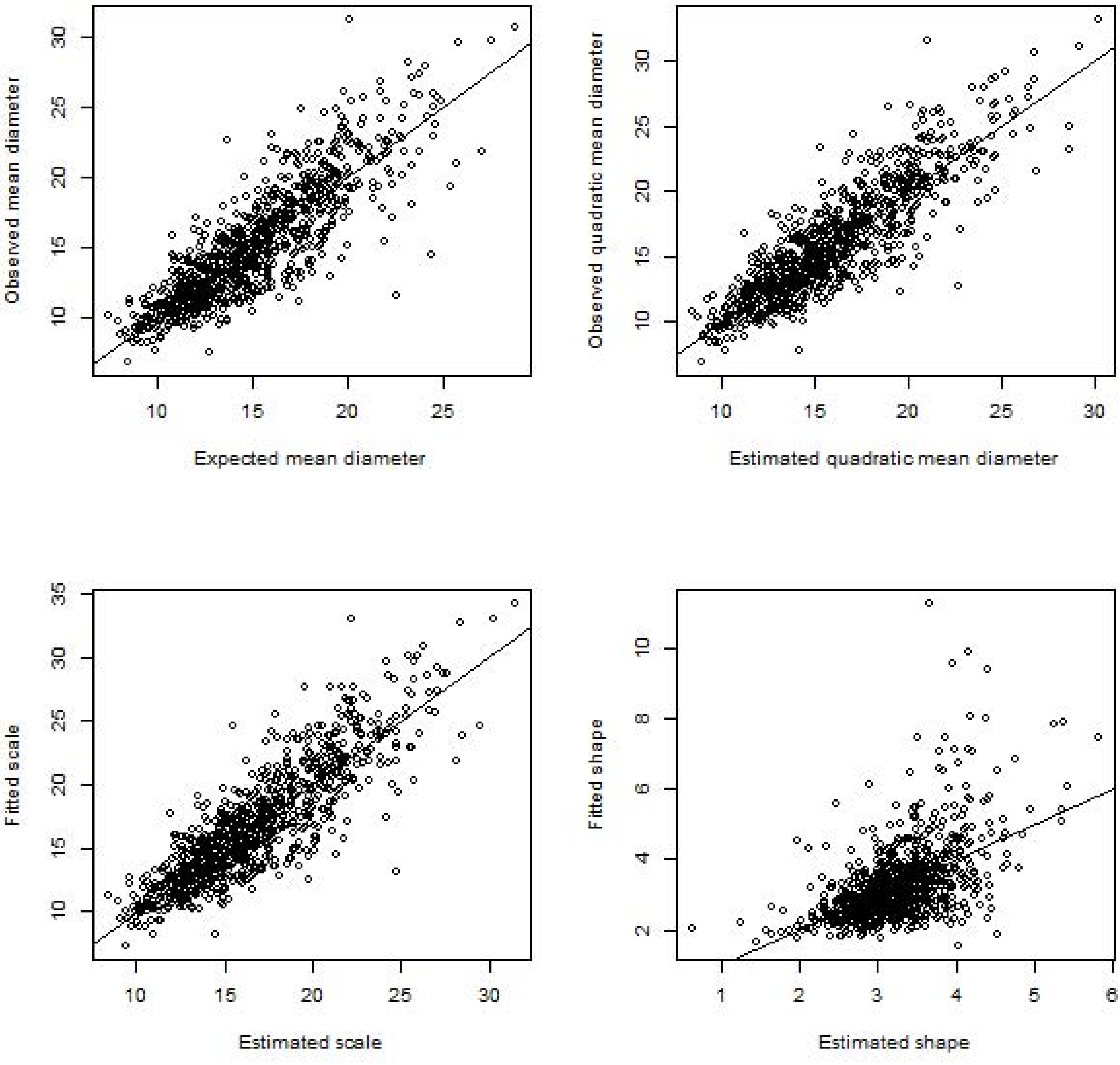

Observed mean diameter
Expected mean diameter
Observed quadratic mean diameter
Estimated quadratic mean diameter
Fitted scale
Estimated scale
Fitted shape
Estimated shape

Table 1. Characteristics of the stands used in the study according to the field survey (124 stands, total area 1,135 hectares).

Variable	Mean	Minimum	Maximum
Area (ha)	9	0.14	66.7
Age (year) [2)]	59[1)]	20	169
Stem volume (m3 ha^{-1})	146[1)]	24	569
Stem diameter [2)] (cm)	19.72[1)]	11.27	34.2

1) Area weighted mean, stand area as the weight.

2) Basal area weighted within stand.

Table 2. ALS metrics extracted for the field sampled plots.

Metric	Variable names
Height above ground values corresponding to the 10th, 20th, …, 90th, 95th and 100th percentiles	h10, h20, …, h90, h95, h100
Mean height above ground	hmean
Standard deviation of height above ground	hs
Proportion of returns from the vegetation layer	d

Table 3: Regression coefficients for six plot-level variables versus 14 independent variables obtained from the ALS information and the proportion of pine (from the plot sampling data). All presented coefficients are statistically significant at the 5% level. Intercepts and F statistics for each dependent variable are also shown.

		Regression coefficients of the independent variables																
Dependent variables	Intercept	Perc10	Perc20	Perc30	Perc40	Perc50	Perc60	Perc80	Perc90	Perc95	*Perc952	perc100	h_{mean}	h_s	d	proportion Pine	R^2	F statistic
Shape	6.004	-0.493	-0.496		-0.589		-0.354		-0.462			-0.103	2.782	-1.374	-3.066	-.593	0.26	34.87
Scale	11.195	-0.502				2.216		2.433		1.093	0.015		-4.348	-5.150	-7.704		0.74	347.1
Basal area per hectare	-20.684	0.854	-1.269	1.603					1.209		-0.022				34.547	1.756	0.69	303.9
Number of stems per hectare	-427.386	73.319		155.038					121.281		-3.434	30.432	-332.325		2804.327		0.55	165.7
Basal area weighted mean height	0.716	-0.031		0.129						0.649		0.088		0.367	-1.078	-0.536	0.81	564.1
Quadratic mean diameter	10.635	-0.647		-0.817		1.687	-1.017	1.532			0.018			-2.829	-7.024		0.76	376

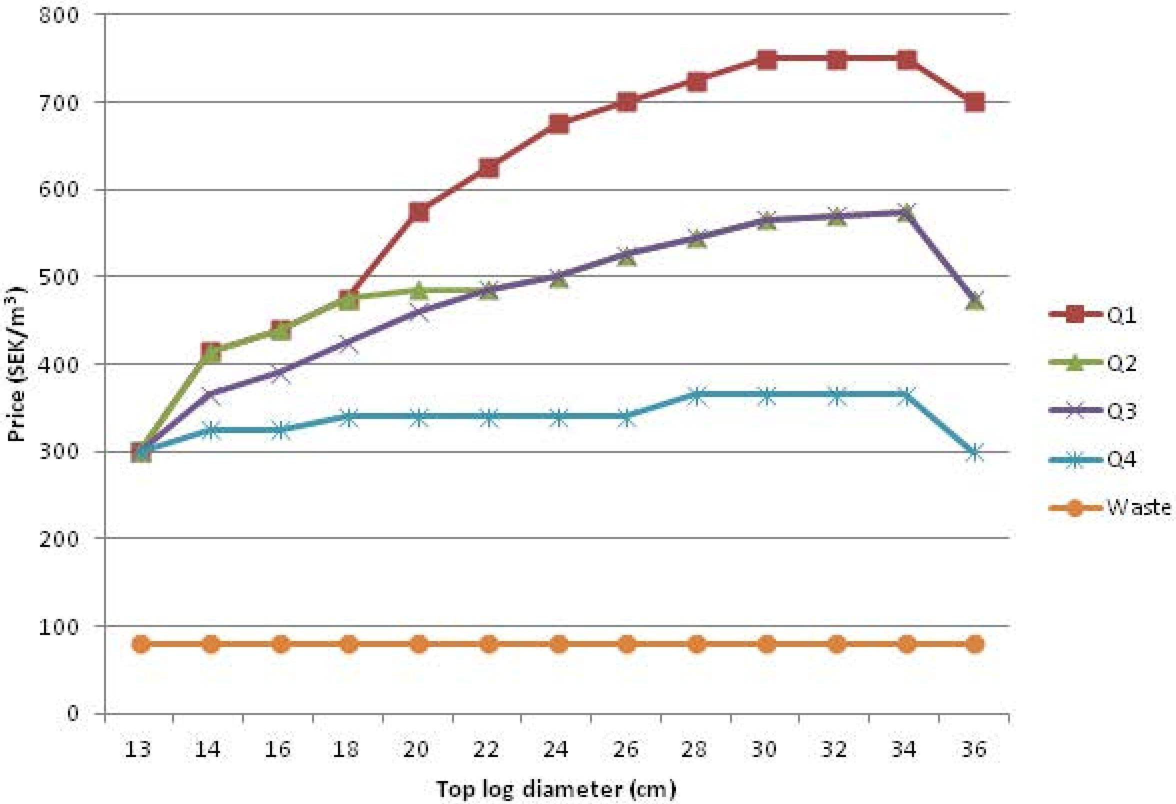
*Perc952 is the Perc95 rise to the power 2.

Table 4: Summary of error indices indicating the accuracy of diameter distributions estimated using the ALS and mean values approaches compared to the measured diameter distributions. e_{ALS} and $e_{Heureka}$ are Reynold indices (range 0 – 200), while δ_{ALS} and $\delta_{Heureka}$ are total variation distance indices (range 0 -1) for the ALS and mean values approaches, respectively. The index value 0 in both indices present perfect matches of the compared distributions.

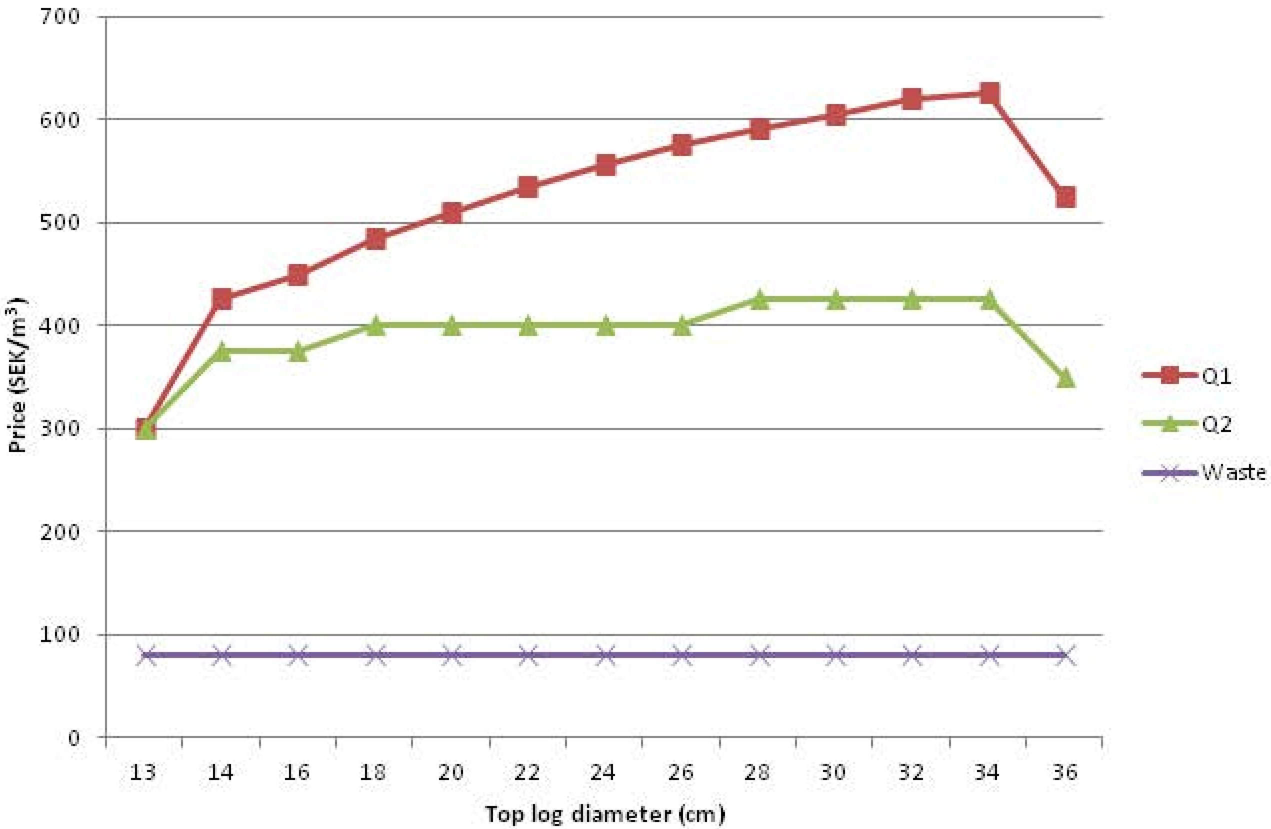
	Error indices			
	Reynolds index		Total variation distances index	
	e_{ALS}	$e_{Heureka}$	δ_{ALS}	$\delta_{Heureka}$
Mean	50.896	79.160	0.251	0.388
Maximum	123.529	159.191	0.542	0.777
Minimum	23.348	39.021	0.090	0.145
Standard deviation	17.454	25.262	0.088	0.122

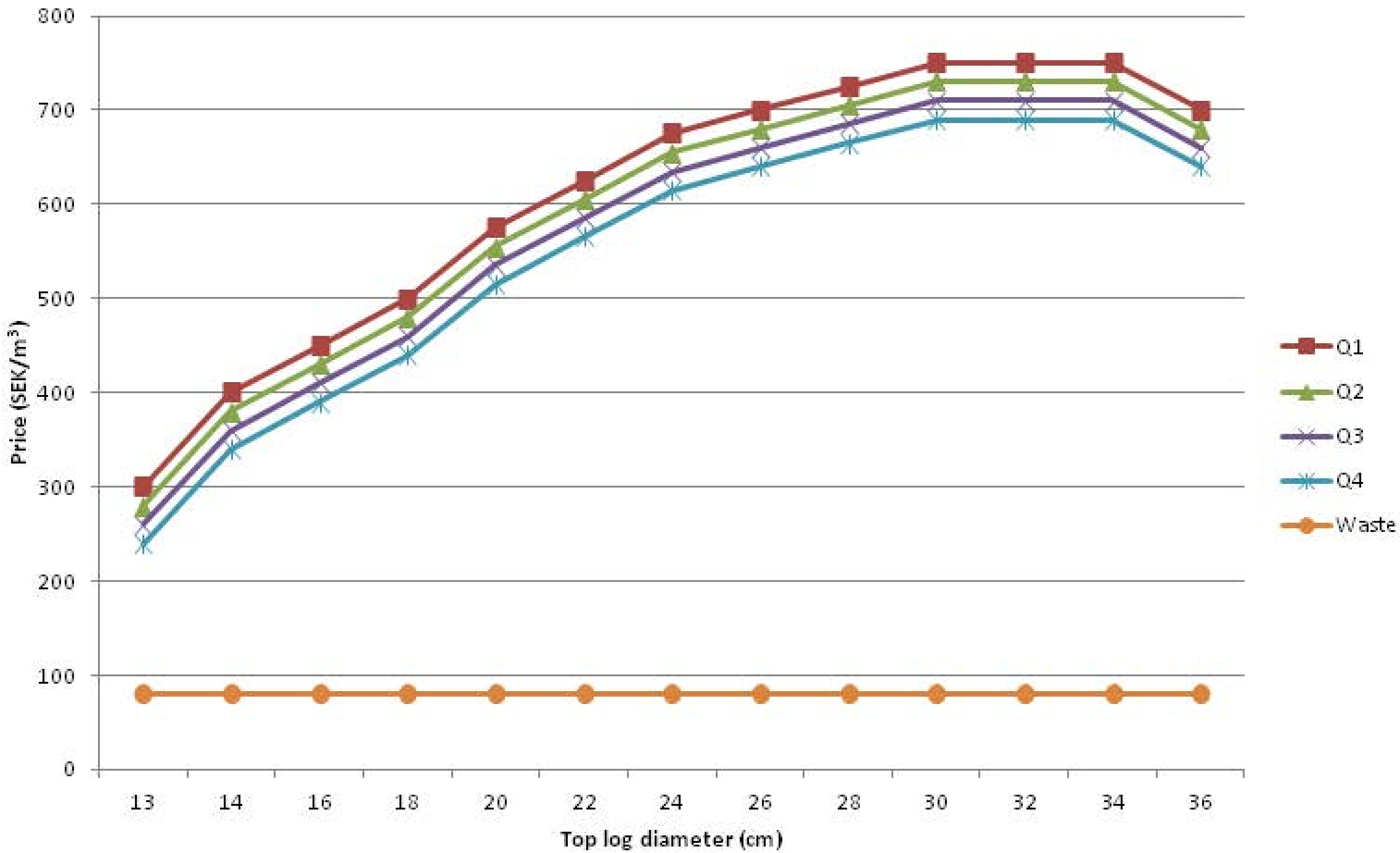
Table 5: Calculated NPVs. $NPV_{Observed}$ is the NPV of the observed alternative. NPV_{ALS} and NPV_{Mean} are the NPV based on the forest information in the observed alternative where the two first period's management alternatives from the ALS and mean values alternatives were applied on the observed alternative, respectively. The difference between NPV_{ALS} and NPV_{Mean} is considered to be the suboptimal loss when ALS information is utilized.

	NPV results (SEK ha^{-1})			
	$NPV_{Observed}$	NPV_{ALS}	NPV_{Mean}	Decrease in suboptimal loss utilizing the ALS information compared to the mean values alternative
Default price list	38,824	38,778	38,712	66
Hypothetical price list	34,139	34,090	33,979	111

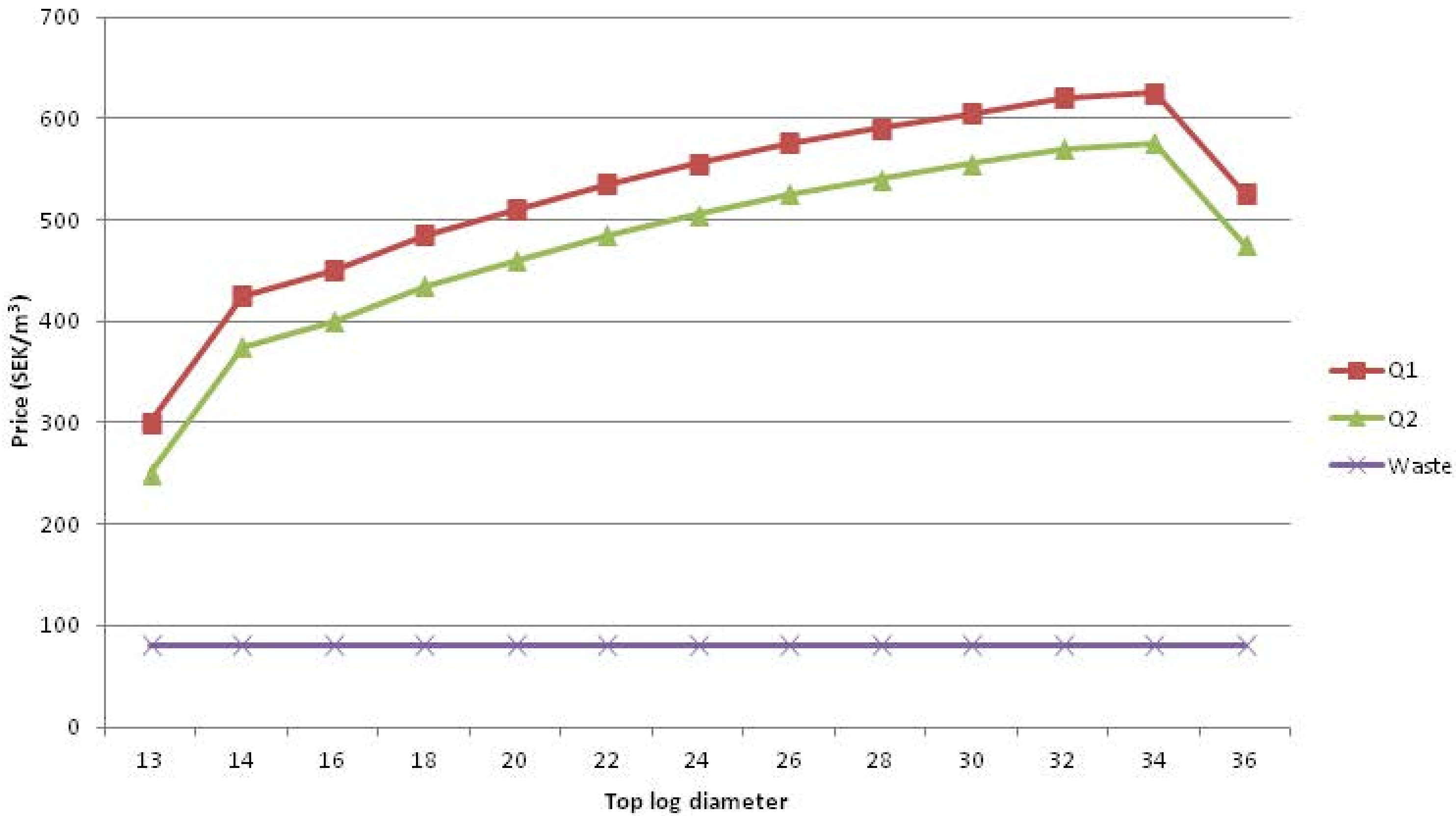


800
700
600
500
400
300
200
100
0
Price (SEK/m³)
13
14
16
18
20
22
24
26
28
30
32
34
36
Top log diameter (cm)
Q1
Q2
Q3
Q4
Waste

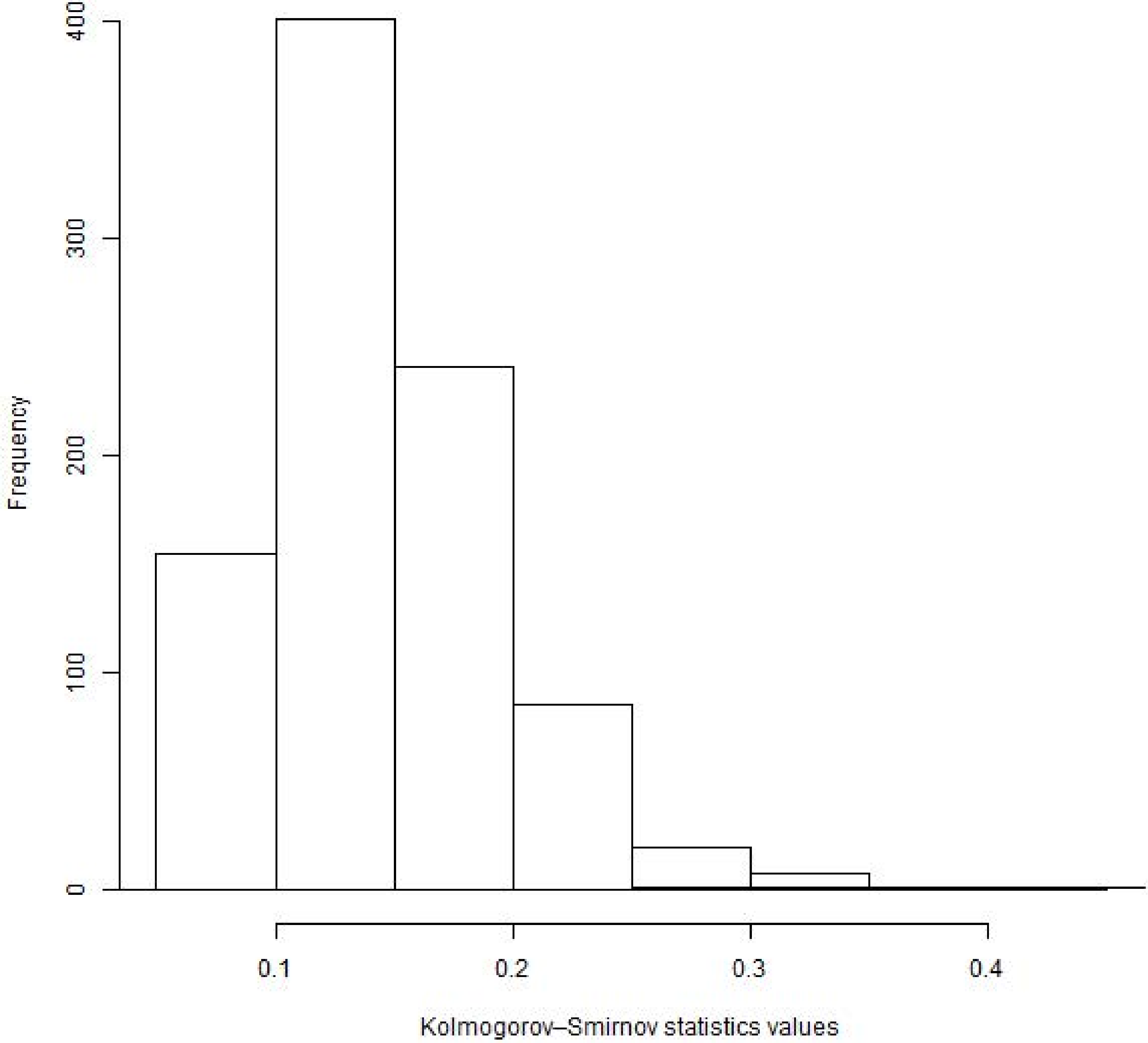

700
600
500
400
300
200
100
0
Price (SEK/m³)
13
14
16
18
20
22
24
26
28
30
32
34
36
Top log diameter (cm)
Q1
Q2
Waste



800
700
600
500
400
300
200
100
0
Price (SEK/m^3)
13
14
16
18
20
22
24
26
28
30
32
34
36
Top log diameter (cm)
Q1
Q2
Q3
Q4
Waste



700
600
500
400
300
200
100
0
Price (SEK/m3)
13
14
16
18
20
22
24
26
28
30
32
34
36
Top log diameter
Q1
Q2
Waste



Frequency
0
100
200
300
400
0.1
0.2
0.3
0.4
Kolmogorov–Smirnov statistics values