Accuracy assessment of the nationwide forest attribute map of Norway constructed by using airborne laser scanning data and field data from the national forest inventory

Ana de Lera Garrido[a]*, Terje Gobakken[a], Marius Hauglin[b], Erik Næsset[a], Ole Martin Bollandsås[a]

[a] Faculty of Environmental Sciences and Natural Resource Management, Norwegian University of Life Sciences, Ås, Norway.

[b] Norwegian Institute of Bioeconomy Research, Ås, Norway

*Corresponding author:
Ana de Lera Garrido
E-mail address: ana.garrido@nibio.no

Abstract

The aim of this study was to analyze the accuracy of predictions of dominant height, mean height, basal area, and volume from the nationwide forest attribute map (SR16). The analysis took advantage of field observations from 33 different forest inventory projects across Norway used for validation. Forest attributes for more than 5000 plots were predicted using non-stratified and stratified models of SR16 and the predictions were compared against corresponding ground reference values. Finally, the effect of different factors that might have influenced the prediction errors were analyzed using partial least squared regression (PLSR) to determine under which conditions the SR16 is less apt. The overall results across all plots were adequate (RMSE of 10%, MD of 2% for dominant and mean height; RMSE of 28%, MD of 4% for basal area; RMSE of 31%, MD of 5% for volume). However, when the accuracy was assessed locally for each inventory project, large differences in accuracy were observed. The MD% values for some inventory projects were substantial (>30% for basal area and volume). The results showed that stratification did not necessarily improve the results and that factors related to the forest structure had the greatest impact on the PLSR analysis.

Keywords: Forest resource map, Forestry, Lidar, NFI, remote sensing

1. Introduction

Forest inventory information is collected at different geographical scales for different purposes. A broad range of applications such as international reporting, biodiversity and restoration programs, or disturbance assessments, require national and international statistics collected by national forest inventories (NFIs). Although there are differences, NFI sampling designs are often based on a network of permanent plots that are systematically distributed over the entire county and revisited periodically (Tomppo et al., 2010). The purpose of an NFI is to provide nationwide and regional statistics about forest resources, their changes, and monitoring of forest conditions (e.g., standing volume, increment, and carbon storage) (Tomppo et al., 2010, McRoberts et al., 2010).

Forest management decisions related to harvesting or other silvicultural activities are often made at stand level. For this purpose, forest information is typically acquired by means of forest management inventories (FMIs). The methods used in FMIs have changed with time and technological development (Maltamo et al., 2021). Nowadays, in the Nordic countries, stand-wise forest management plans usually originate from area-based inventories employing wall-to-wall data from airborne laser scanning (ALS) and a sample of field reference plots (Nilsson et al., 2017, Waser et al., 2017). The ALS data acquired for the entire area of interest (AOI) are tessellated into grid cells that serve as the primary prediction units. The field reference plots are distributed over the AOI, in some countries typically according to a stratified sampling design (Næsset, 2014). The plots are georeferenced, field measurements of diameter at breast height and tree heights are carried out to enable calculations of forest attributes, and metrics representing the properties of the ALS point cloud are calculated for each plot. ALS metrics are also extracted for each grid cell whose size is equal to the sample plot size. Then, prediction models of the relationships between the ALS metrics and the forest attributes are constructed. Finally, the models are used to predict the forest attributes for every grid cell, and the individual cell predictions are aggregated to stand level estimates (Næsset, 2002, White et al., 2013).

NFIs and FMIs are carried out independently, following separate methodologies. While most NFIs are designed to produce estimates of forest attributes from field data only, a modern FMI project often covers the forest area in a municipality or a single or a few larger forest properties and requires wall-to-wall remotely sensed data. However, if NFI plots are georeferenced, they can be an excellent source of reference data for model construction to create wall-to-wall predictions and produce the same type of stand-wise information as a traditional FMI (Chirici et al., 2020, Vega et al., 2021, Guerra-Hernández et al., 2022). Models could be constructed for an AOI using the local NFI plots as reference data. However, for smaller AOIs, the size of the NFI plot sample might in many cases be too small. Alternatively, now that nationwide ALS data have been collected for topographic mapping and other purposes in several countries, NFI field plots can be used as reference data to construct regional or nationwide models and produce forest resources maps (Nord-Larsen and Schumacher, 2012, Monnet et al., 2016, Nilsson et al., 2017, Hauglin et al., 2021)

Compared to traditional FMIs, the main advantage of an inventory system where NFI plots are used as field reference is that the costs of the NFI plots are already covered by other budgets. Additionally, NFI data are in many countries collected continuously (Gschwantner et al., 2022). For example, the Norwegian NFI has a five-year rotation period, so that one-fifth of the plots are measured every year, permitting annual updates of forest statistics for the entire country (Breidenbach et al., 2020). While the intervals between FMIs have traditionally been 10-20 years, an advantage of using NFI data for model calibration is that wall-to-wall prediction maps of forest attributes can be updated as frequently as the appearance of new ALS data permit. However, local map predictions based on NFI plots as field reference data are often considered to be less accurate than those of a local FMI (Kangas et al., 2018).

The use of NFI plots as calibration data for prediction and subsequent stand-level estimation has some challenges. The sampling intensity of an NFI is small, with the consequence that prediction models will have to be calibrated with field plots collected over a large spatial domain, typically tens of thousands of square kilometers (Gschwantner et al., 2022). The stand structure, for example, as expressed by the three-dimensional distribution of biological material in the crowns and captured in the ALS data, varies relative to stem properties and thus attributes such as tree height, stem diameter, and volume, according to factors such as latitude, elevation, soil properties, and other factors with a distinct geographical pattern (Næsset, 2014). Therefore, the field reference plots from larger geographical regions will likely represent relationships between field reference data and the ALS metrics that are not necessarily representative for smaller geographical domains (Nilsson et al., 2017). A model calibrated on data for a larger region (e.g. using NFI plots) may not be correctly specified for a small AOI with its peculiarities, potentially leading to systematic errors in the model predictions (Guerra-Hernández et al., 2022). Also, the systematic sample typically acquired in NFIs is unlikely to capture the entire range of variability of the forests and "extreme" cases of particular AOIs (Kangas et al., 2018). Therefore, it is important to evaluate the predictions across different AOIs whose forest conditions and corresponding variability do not necessarily match those of the model construction region.

Another challenge with model calibration and prediction based on NFI data is the temporal differences that might occur between different ALS acquisitions and between different parts of the NFI plot dataset within a region. For a larger region, field data will be acquired over many years, and they will have to be projected to a common date, which may introduce errors. A larger region will typically be covered by different non-overlapping ALS acquisitions from different points in time using different acquisitions parameters and instruments that may affect the point clouds and the derived metrics (Næsset, 2005, Goodwin et al., 2006, Næsset, 2009). As opposed to the field data, ALS data cannot be prorated or back-casted, so the field and ALS data will simply reflect different forest conditions, which will affect the models (Hill et al., 2018). Also, the temporal gap between ALS and field data acquisitions might not be constant but could be accounted for in the model along with e.g., sensor effects by using, for example, mixed-effect models (Hauglin et al., 2021).

Beyond temporal inconsistencies among data collected on field plots and among ALS acquisitions, there may also be temporal inconsistencies between field data and ALS data for the same geographical area, resulting in different states of the forest when the two types of data are collected. Causes of state differences could be tree growth, tree recruitment, harvests, and natural disturbances. For example, the ALS data might be collected over plots that have been thinned while the field data were collected before thinning, and vice versa. Failure to detect even a small number of plots with disturbances has the potential to dramatically inflate the model uncertainty (Massey and Mandallaz, 2015). Thus, detection of, for example, disturbances using procedures based on e.g., satellite data (Huang et al., 2010, Verbesselt et al., 2012, Hansen et al., 2013, Jutras-Perreault et al., 2021) might be necessary to discard such observations from the model calibration data. However, plots subject to thinning operations are not easily recognized and may introduce errors in models and, therefore, also in the final predictions.

Given the obvious potential for cost-savings but also the risk of less accurate predictions when adopting regional or nationwide prediction maps of forest attributes based on NFI data and large-area ALS campaigns, it remains an open question if such large-scale prediction maps can be useful for operational forest management planning, and thus substitute local FMIs. In Norway, such nationwide prediction maps were constructed and made publicly available some years ago, known as the Norwegian SR16 forest resource map. "16" refers to the size of the map pixels (16 m × 16 m) (Astrup et al., 2019). Such map products have the potential to bridge the gap between NFI statistics and the need for local forest information (Astrup et al., 2019). However, more empirical research is needed to evaluate the prediction maps locally across a broad range of differences in local forest conditions.

The main objective of the study was to calculate, identify and assess any potential systematic errors of the Norwegian SR16 forest resource map (Hauglin et al., 2021), which is available on the web (http://kilden.nibio.no, Skog, SR16). The analysis included calculations of the root mean squared error of the differences between SR16 predictions and ground reference plot values. The study was based on observations of more than 5000 field plots distributed over 33 different local FMI projects across the country. This study addressed the four important attributes dominant height, mean height, basal area, and volume. A secondary objective was to identify forest properties that might influence the differences between the SR16 predictions, and the field reference values. Such insight could provide guidance on when SR16 can be expected to perform well and when the prediction maps are less apt to inform forest management decisions, as well as to improve the SR16 maps in the future.

2. Materials and methods

2.1. Study area

The study region comprises most of southern Norway, for which the forest resource map SR16 was available (Figure 1). The area represents 2/3 of the productive forest area in Norway and covers different vegetation zones (nemoral, boreonemoral, and boreal) (Moen, 1999), for which forest and growing conditions vary considerably with latitude, altitude, and climatic conditions. The dominant tree species in the study region are Norway spruce (*Picea abies* (L.) Karst.), Scots pine (*Pinus sylvestris* L.), and deciduous species, mainly birch (*Betula pubescens* Ehrh.).

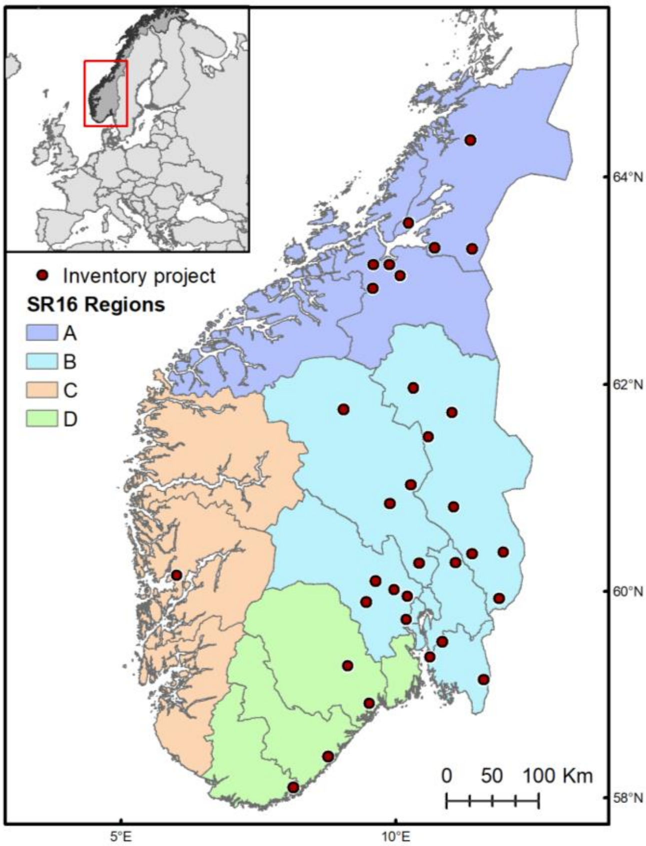

Figure 1. Map of the study region showing the location of the FMI projects (circles) used for validation and SR16 regions displayed in different colors, for which unique prediction models were constructed and applied.

2.2. Field data.

We used sample plot data from 33 FMI projects across Norway provided by four private forest inventory companies (Figure 1). Across all FMIs, 5167 sample plots were available for the current study. The field plot data were collected as parts of ALS-based FMIs during the years 2012 - 2019. For each FMI, forest stands were delineated based on age, site index, and development class obtained by photointerpretation. Development class was defined according to a national system of maturity classification described in Anon. (1987). Classes 1 to 5 represent development stages from clear-felled stands to mature stands ready for harvest. We excluded plots where dominant heights were <8 m, so only classes 3 to 5 were included in this study. These classes are potentially subject to treatments such as thinning and various forms of final felling and therefore require information on mean height, basal area, and volume.

The field measurement protocols were slightly different among the projects because of the different companies involved. Thus, there were some differences in plot sizes (232 m^2 or 250 m^2 circular plots), sampling designs (systematic cluster sampling or stratified systematic sampling), and different lower caliper limits (4, 6, or 10 cm). Planimetric coordinates of plot centers were recorded using observations of global navigation satellite systems (GNSS) which were post-processed and corrected against observations from official base stations of the Norwegian Mapping Authority.

On each plot, all trees above the lower caliper limit were registered; diameter at breast height (dbh) was measured, and tree species were recorded. The procedures for height measurements varied among the measurement protocols of the projects. While heights of all trees were measured for one of the projects, only heights of sample trees were measured for all the other projects. The sample trees were selected with a probability proportional to stem basal area using a relascope aiming for 10 sample trees per plot.

Based on the field plot registrations, we calculated ground reference values for dominant height (Hd, m), mean height (Hm, m), basal area (G, $m^2\ ha^{-1}$), and volume (V, $m^3\ ha^{-1}$). To obtain data consistent with those of SR16, trees with dbh < 5 cm were discarded from the analysis. Since heights were only measured for sample trees, we applied the following procedure to obtain values for our selected forest attributes. First, single tree volumes ($\hat{v}$) were obtained by calculating a reference-level volume (rlv) for all trees, multiplied with a correction factor (cf).

$$\hat{\mathrm{v}} = \mathrm{rlv} \times \mathrm{cf} \tag{1}$$

rlv was obtained by first applying the diameter-heigh model of (Fitje and Vestjordet, 1977) to predict a reference-level height which was used as input to the national single-tree volume models (Braastad, 1966, Brantseg, 1967, Vestjordet, 1967) together with dbh. Plot-wise values of cf were obtained as the ratio between the volumes (v) obtained by using the measured height and dbh, and the corresponding values of rlv for the sample trees:

$$\mathrm{cf} = \frac{\sum_{\mathrm{i}=1}^{\mathrm{st}} \mathrm{v_i} \times \mathrm{w_i}}{\sum_{\mathrm{i}=1}^{\mathrm{st}} \mathrm{rlv_i} \times \mathrm{w_i}} \tag{2}$$

where st is the number of sample trees on a given plot. A weight (w) was given to each tree to adjust for unequal inclusion probability of the sample trees.

Single-tree heights were then predicted by using $\hat{v}$ and dbh as fixed values in the single-tree volume model and solving the equation (model) for height. Hd was then calculated as the mean height of the two or three trees with the largest dbh for the 232 and 250 m^2 plots, respectively. Hm was computed as the mean height weighted by basal area. In this estimation, a model-assisted estimator was adopted by which the heights were adjusted for prediction bias observed on the height sample trees. G for each plot was calculated as the sum of individual tree basal areas and scaled to m^2 per ha. Similarly, plot values of V were estimated as the sum of individual tree volumes and scaled to m^3 per ha. The means and standard deviations of these plot-level attributes for all the FMIs are presented in Table 1.

Table 1. Summary of the field plot attributes by forest management inventory project (FMI, official name of the municipality), corresponding field plot numbers (n), dominant height (Hd), mean height (Hm), basal area (G), and volume (V). Code is an abbreviation for each inventory project where the letters represent the SR16 regions shown in Figure 1 and the numbers are running numbers of FMIs within each SR16 region.

				Hd (m)		Hm (m)		G ($m^2 \cdot ha^{-1}$)		V ($m^3 \cdot ha^{-1}$)	
Code	**FMI**	**year**	**n**	**mean**	**sd**	**mean**	**sd**	**mean**	**sd**	**mean**	**sd**
A1	Leksvik	2016	152	16.86	4.75	14.01	4.39	22.76	12.18	159.31	128.82
A2	Meldal	2018	127	18.85	4.11	15.62	3.55	30.43	13.71	242.55	147.33
A3	Melhus	2013	85	17.17	3.85	14.86	3.40	25.29	10.15	188.11	99.97
A4	Meråker	2019	95	17.05	3.14	13.17	3.05	27.45	10.31	182.87	95.80
A5	Orkdal	2018	60	18.62	3.85	15.23	3.78	31.85	11.17	237.91	116.07
A6	Overhalla	2019	79	18.72	5.10	15.24	4.65	29.10	13.78	234.02	157.73
A7	Skaun	2015	107	18.27	4.08	14.83	3.85	34.52	14.57	250.74	146.89
A8	Stjørdal	2019	296	19.37	4.43	15.70	3.92	33.56	17.65	271.73	189.83
B1	Alvdal	2017	130	15.98	3.45	13.48	3.19	19.79	9.92	138.97	86.43
B2	Aremark, Idd	2018	306	18.62	4.74	15.51	4.36	24.78	11.21	204.41	136.62
B3	Dovre, Lesja, Vågå	2014	112	15.89	2.89	13.37	2.98	27.76	13.44	188.05	110.27
B4	Eidskog	2018	240	20.61	4.27	17.18	4.34	23.59	9.47	210.44	117.08
B5	Eidsvoll	2018	291	20.24	5.02	16.78	4.72	28.54	13.13	252.38	174.78
B6	Grue	2016	129	19.26	3.95	15.57	3.70	26.13	10.96	211.36	121.24
B7	Hadeland	2016	295	19.68	4.64	16.16	4.41	28.63	12.39	242.19	153.14
B8	Hamar, Løten	2019	99	21.05	5.19	17.61	4.88	32.47	14.67	300.65	188.44
B9	Hobøl	2019	88	20.19	4.58	16.63	4.15	27.59	11.55	241.01	156.17
B10	Hole	2017	81	19.62	4.09	16.50	3.69	30.84	11.63	245.78	121.42
B11	Krødsherad	2016	103	19.68	4.06	16.85	3.85	29.91	12.67	253.07	150.49
B12	Lillehammer	2015	127	18.77	4.19	15.40	4.06	29.56	12.24	229.44	136.79
B13	Modum, Lier, Røyken, Hurum	2019	174	20.86	4.29	18.04	3.92	29.48	11.24	260.09	144.43
B14	Moss	2019	82	19.87	5.08	16.47	4.57	29.00	12.96	250.48	169.66
B15	Nord-Odal	2016	143	20.46	4.44	16.98	4.09	26.33	10.82	228.91	126.55
B16	Nordre Land	2017	169	20.62	4.27	17.55	3.91	28.81	12.25	253.57	144.79
B17	Rendalen	2019	228	16.86	4.21	13.94	3.76	19.45	10.89	143.89	106.78
B18	Sigdal, Flesberg	2019	275	20.00	4.06	17.39	3.52	24.90	11.56	216.91	131.47
B19	Stor-Elvdal	2017	223	16.88	4.09	14.05	3.51	18.79	10.39	137.97	97.47

Code	FMI	year	n	Hd (m) mean	Hd (m) sd	Hm (m) mean	Hm (m) sd	G ($m^2 \cdot ha^{-1}$) mean	G ($m^2 \cdot ha^{-1}$) sd	V ($m^3 \cdot ha^{-1}$) mean	V ($m^3 \cdot ha^{-1}$) sd
B20	Tyristrand	2017	103	19.54	3.29	17.56	2.91	24.03	9.36	203.92	96.43
C1	Fusa	2012	113	18.14	5.12	15.21	4.42	30.49	16.10	246.10	181.20
D1	Arendal	2018	141	19.29	4.40	16.37	4.07	32.68	11.87	264.30	144.86
D2	Bamle	2017	122	19.41	4.24	16.36	3.61	32.92	12.83	266.89	155.42
D3	Bø, Nome, Sauherad	2019	266	18.72	4.01	15.84	3.61	27.44	10.61	216.76	114.58
D4	Kristiansand	2017	126	17.28	3.98	14.60	3.59	30.55	11.29	220.34	113.56

2.3. Norwegian forest resource map SR16.

The forest resource map SR16 provides raster-based predictions of the forest attributes in 16 m x 16 m resolution that can be directly used to estimate means and totals of forest attributes within a defined area of interest, for example, individual forest stands. Due to the substantial amounts of data involved, the production of SR16 was carried out within several individual regions which were processed separately. Four primary SR16 regions were used in the present study, hereafter referred to as A, B, C, and D (see Figure 1).

In the SR16, non-stratified models for the different forest attributes were constructed using linear mixed-effect regression, accounting for data from multiple ALS acquisitions carried out between 2009 and 2020. The models were based on the field observations from the Norwegian NFI forecasted to a common reference date as field reference data, and metrics derived from the ALS data were adopted as explanatory variables. Then, the models were used to predict the forest attributes for 16 m x 16 m cells tessellating all areas where ALS data were available. More specific models and predictions were also available for specific forest types defined according to a stratification of the NFI plots. For a detailed description of the SR16 models and products, see Astrup et al. (2019) and Hauglin et al. (2021). In the current study, 12 stratified SR16 models based in different development classes, site indices, and dominant tree species were available for each forest attribute within each region.

For the current study, we have used both non-stratified and stratified SR16 predictions that were available for two different reference years: 2019 in regions A, B, and C, and 2020 for region D. The only changes taken into consideration between the ALS acquisition and the reference year were harvested areas detected by the global forest watch (Hansen et al., 2013). The forest attributes Hd, Hm, G, and V were provided by both the non-stratified and stratified SR16 models, except for Hd in region D, for which the stratified SR16 predictions were not available when the analyses were carried out.

The accuracy of the SR16 predictions was evaluated using the ground reference values of the FMI sample plots. SR16 predictions of the forest attributes were extracted for each circular FMI plot by weighting the individual cell predictions for cells intersecting the plot by the individual cell's area included within the plot. The basis for choosing a particular stratified model used for predictions for a certain FMI plot was the stratification following the delineated stands in the FMI.

Since ground reference values and predictions were related to different points in time, the SR16 predictions were back-casted to the date of the FMI field plot acquisition using growth models for Hd (Sharma et al., 2011) and V (Delbeck, 1965, Blingsmo, 1988). HL was corrected by keeping the relative difference between Hd and Hm fixed. G was corrected by the mean ratio between the back-casted and the initial volume prediction.

Since the FMI projects used as field reference were carried out in the period between 2012 and 2019, the maximum time difference accounted for by the correction outlined above was seven years. The largest differences between the corrected and the initial SR16 predictions were 5 m for Hd and Hm, 18 m^2 ha^{-1} for G, and 196 m^3 ha^{-1} for V. The mean values of the corrections were 0.4 m, 0.3 m, 2 m^2 ha^{-1}, and 13 m^3 ha^{-1} for Hd, Hm, G, and V, respectively.

2.4. Accuracy assessment

To assess the systematic differences, SR16 predictions of the forest attributes of interest were compared against the field plot ground reference values at plot level by computing the difference between ground reference and the predicted value (D_i), mean difference (MD%), and relative root mean squared error (RMSE%), calculated as:

$$D_i = y_i - \hat{y}_i \tag{3}$$

$$MD\ (\%) = \frac{\frac{1}{n}\sum_{i=1}^{n}(D_i)}{\bar{y}} \cdot 100 \tag{4}$$

$$\mathrm{RMSE\%} = \frac{\sqrt{\frac{1}{\mathrm{n}}\sum_{\mathrm{i}=1}^{\mathrm{n}}(\mathrm{D_i})^2}}{\bar{\mathrm{y}}} \cdot 100 \tag{5}$$

where n is the number of plots, y_i is the ground reference value for the forest attribute in plot i, $\hat{y}_i$ is the corresponding predicted forest attribute from SR16, $\bar{y}$ is the mean ground reference value for the forest attribute.

For some plots, the state at the time of ALS acquisition and the time of FMI field inventory could be different as a result of, for example, undetected disturbances in SR16 or incorrectly classified harvests in SR16 for which the FMI field inventory acquired field plot data representing full stocking. Such discrepancies would result in large outliers in the analysis. However, it was not feasible to check every individual among the more than 5000 plots. Instead, we applied Rosner's test (Rosner, 1983) to each FMI separately to automatically detect outliers that differed significantly from the rest of the observations (plots) after calculating the difference between ground reference and SR16 predictions. In the test, the number of observations that are considered outliers in a distribution was limited to 10. Plots considered as outliers for any forest attribute were removed. The non-stratified and stratified SR16 predictions were analyzed separately, potentially resulting in the identification of different outliers for the same attribute. The accuracy and outlier detection were assessed separately for each FMI project, and paired t-tests were carried out to estimate the statistical significance of the differences.

2.5. Other auxiliary plot data

Additional local back-ground factors that may capture differences between local FMI plot data and SR16 predictions were extracted to analyze if the SR16 predictions were equally accurate over a greater range of conditions. The local factors were divided into four categories: climate, topography, forest conditions, and other factors (Table 2).

Mean monthly precipitation and temperature predictions (Tveito et al., 2005) from the Norwegian Meteorological Institute for the period between 1989 and 2018 were used as proxies to describe climate at each plot. The predictions were processed and adjusted according to local elevation (Skaugen et al., 2003). Means for the summer (June, July, August) and winter (December, January, February) months were calculated for each plot. The topographic factors were elevation, slope, and topographic heterogeneity calculated as the standard deviation of the elevation and designed to represent the topographic variation inside the plot. All were derived from the national detailed elevation model (10 m resolution) created by the Norwegian Mapping Authority. The forest condition category included stand age, height-diameter ratio (HDR), development class, tree species composition, and site index. HDR represents the allometric relationship between height and diameter and was calculated as the mean height-diameter ratio of all trees in the plot. The standard deviation of the height-diameter ratio (HDR.sd) was calculated to represent the variation in tree allometry inside the plot. Tree species compositions were represented as the proportions of deciduous, pine, and spruce species in each plot according to stem basal area.

The category denoted as "other factors" included factors related to the study design that might have affected the results. These factors were: SR16 region, if there were differences in lower caliper limits between the SR16 and the ground reference plots, the number of years between SR16 predictions (i.e., 2019 or 2020) and ground references, and the number of years between SR16 predictions and the ALS acquisition. The latter is of particular interest since, as mentioned above, the only change on the ground accounted for between acquisitions were harvest. A summary of the factors' mean values by FMI project is shown in appendix A.

Table 2. Factors that represent the differences between forest inventory projects. Group numbers denote the main categories: 1) climate, 2) topography, 3) forest conditions, 4) other factors.

Group	Factor	Description
1	P.s	Mean of the monthly mean precipitation in the summer months
1	P.w	Mean of the monthly mean precipitation in the winter months
1	T.s	Mean of the monthly mean temperature in the summer months
1	T.w	Mean of the monthly mean temperature in the winter months
2	elev	Elevation from the sea level
2	slope	Mean slope
2	T.H	Topographic heterogeneity
3	Age	Stand age
3	HDR	Mean height-diameter ratio
3	HDR.sd	Standard deviation of the height diameter ratio
3	HKL	Development class (classes 3 - 5)
3	p.D	Proportion of deciduous trees according to stem basal area
3	p.P	Proportion of pine trees according to stem basal area
3	p.S	Proportion of spruce trees according to stem basal area
3	SI	Site index
4	Area	SR16 region (classes A – D)
4	c.d	Difference between the lower caliper limit of SR16 and the FMI project (Yes - No)
4	Year.d	Number of years between SR16 predictions and ground reference
4	ALS.d	Number of years between SR16 predictions and ALS acquisition

2.6. Variable importance analysis.

To estimate and evaluate how strongly the factors detailed in Table 2 influenced the differences (eq. 1) between the SR16 predictions and the ground reference values, a partial least square (PLS) analysis was carried out. We performed a PLS regression (PLSR) with the Kernel algorithm (Dayal and MacGregor, 1997).

PLSR is a multivariate linear regression method widely used in chemometrics to analyze data with numerous predictor variables that might be strongly collinear and noisy (Wold et al., 2001). The method has also been adopted in studies related to use of ALS data in forestry (Næsset et al., 2005). It finds independent latent variables that explain as much of the covariance as possible between the predictors and the response variables. Prior to the analysis, we standardized the predictor variables and made the distributions symmetrical to avoid a possible bias towards numerically larger values caused by different units in the factors. PLSR was computed with the pls package in R, and the optimal number of components was selected automatically with the randomization strategy (van der Voet, 1994).

To summarize the global contribution of each predictor variable (the factors in Table 2) to the complete PLSR model, we calculated the variable importance on projection (VIP). As a rule-of-thumb, predictors with VIP values larger than 1 are considered important and highly influential predictors of the model (Eriksson et al., 2013). This threshold comes from the fact that the average of the squared values of the VIPs is equal to 1.

3. Results

3.1. Overall accuracy of the SR16

We compared both the non-stratified and the stratified SR16 predictions against the ground reference values. Figure 2 shows 2d-histograms of both the non-stratified and stratified SR16 predictions versus the ground reference values for the different forest attributes. Values of MD% and RMSE% for the respective SR16 predictions and forest attributes are displayed for each plot. The total numbers of plots used to evaluate the non-stratified and stratified SR16 predictions were different because different numbers of plots were removed in the automatic outlier detection (see details above). MD% values ranged from -0.9% to 4.5%, whereas RMSE% values were 10%, 10%, 28%, and 31% for Hd, Hm, G, and V, respectively, for both the non-stratified and stratified predictions. The comparison of results between the non-stratified and stratified predictions revealed only minor differences, where the largest difference was found for Hm (MD% difference of 0.66 and RMSE% difference of 0.37).

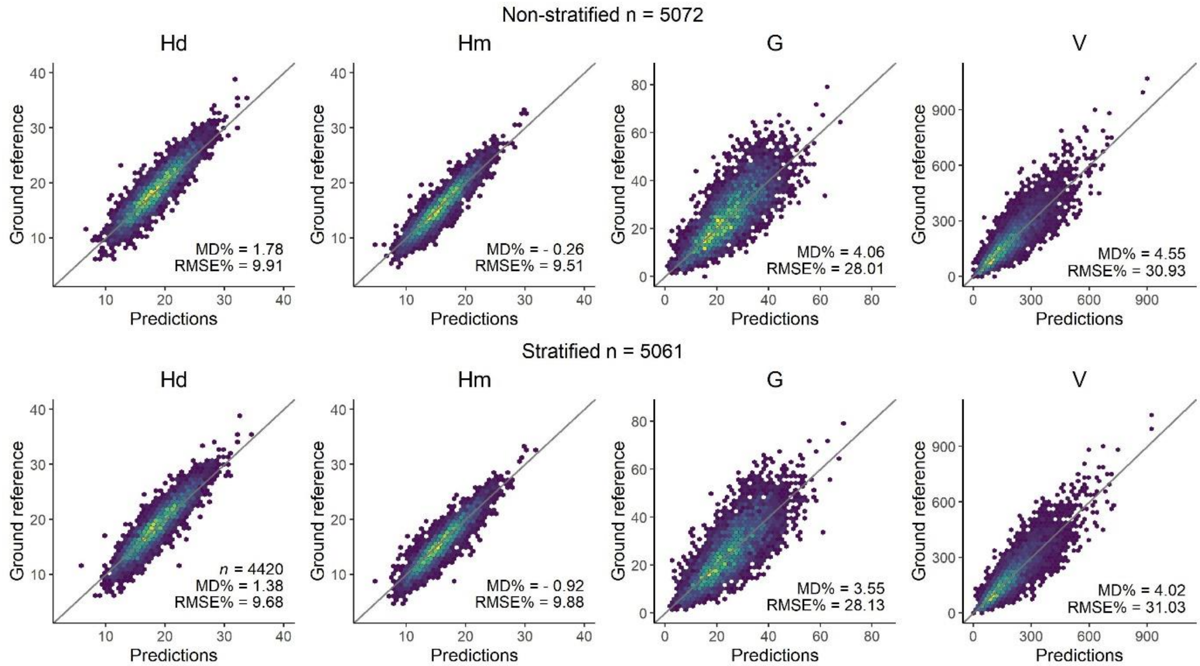

Figure 2. Ground reference data versus SR16 predictions (non-stratified and stratified) for the forest attributes dominant height (Hd), mean height (Hm), basal area (G), and Volume (V). The 1:1 line is presented in grey.

3.2. Local accuracy of the SR16 by FMI project

The MD% results by FMI for both the non-stratified and stratified SR16 predictions are shown in Figure 3. The results showed systematic differences that were not evident when the overall accuracy was analyzed across all plots and FMI projects. Locally, MD% results showed a wide range of values from -7 to 6% for Hd, -10 to 3% for Hm, -17 to 41% for G, and from - 15 to 39% for V. Differences in MD% between non-stratified and stratified SR16 predictions were particularly evident in some of the FMI projects (e.g., A1 for Hd, B3 for Hm, A4 for G and V and C1 for V) but were mostly small. However, when differences in MD% between non-stratified and stratified predictions occurred, MD% was sometimes smallest for the non-stratified predictions (e.g., for A4) and sometimes smallest for the stratified predictions (e.g., for C1). The results showed that the MD% for regions C and D tended to be positive for all forest attributes, whereas in regions A and B the values were more evenly distributed, especially for the Hd and Hm predictions.

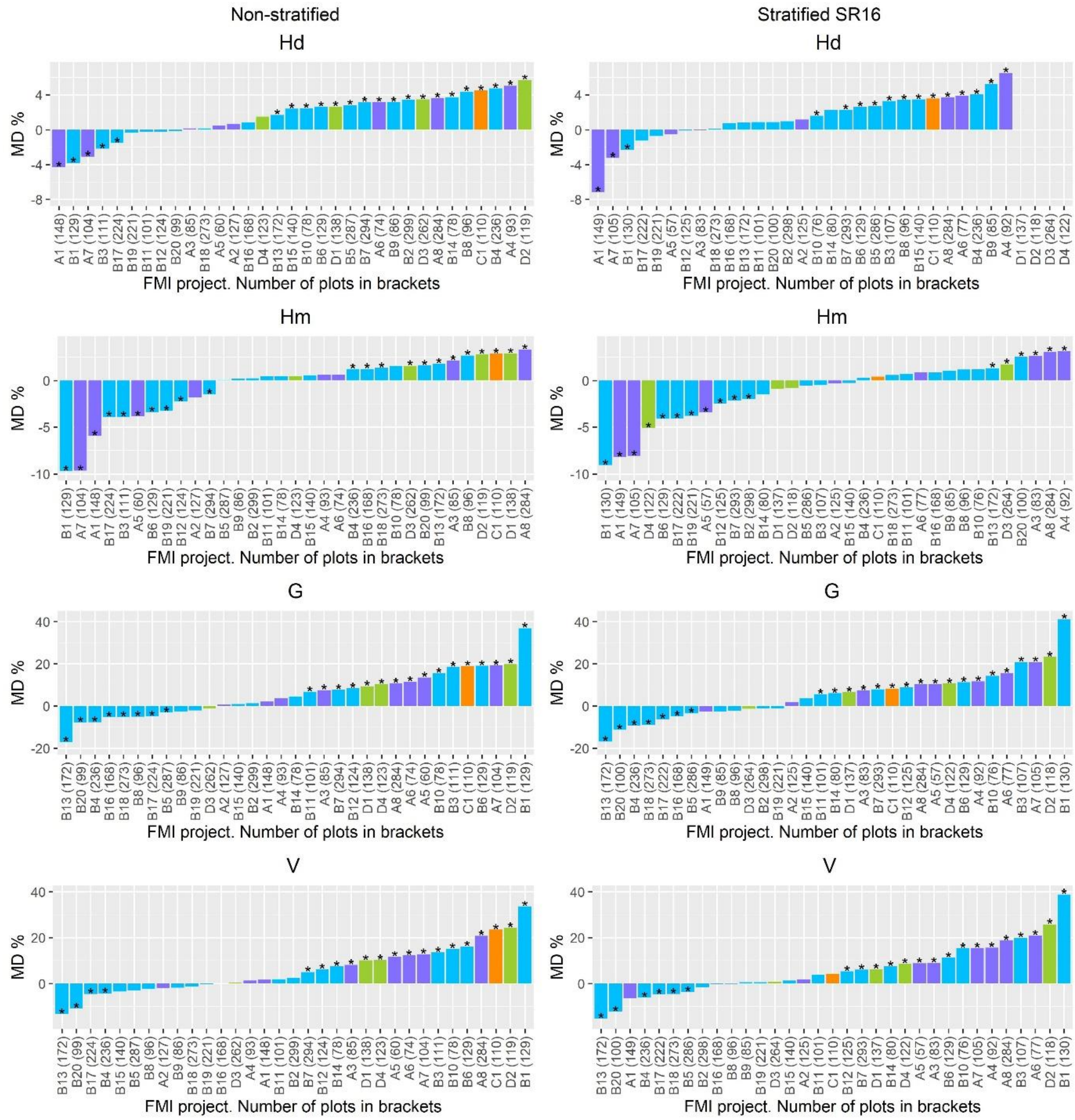

Figure 3. Relative mean differences (MD%) between non-stratified (left) and stratified (right) SR16 predictions and ground references by FMI project for the forest attributes dominant height (Hd), mean height (Hm), basal area (G), and Volume (V). The different colors represent the SR16 regions (A = purple, B = blue, C = orange, and D = green). * Represents the differences that are statistically significantly different from zero (p-value < 0.05).

The non-stratified RMSE% results are shown in Figure 4. As with the MD%, Hd and Hm were the attributes with the smallest RMSE% values, ranging from 6 to 17%. The results for G ranged from 22 to 46%, and V was the forest attribute with the largest values, ranging from 24 to 51%.

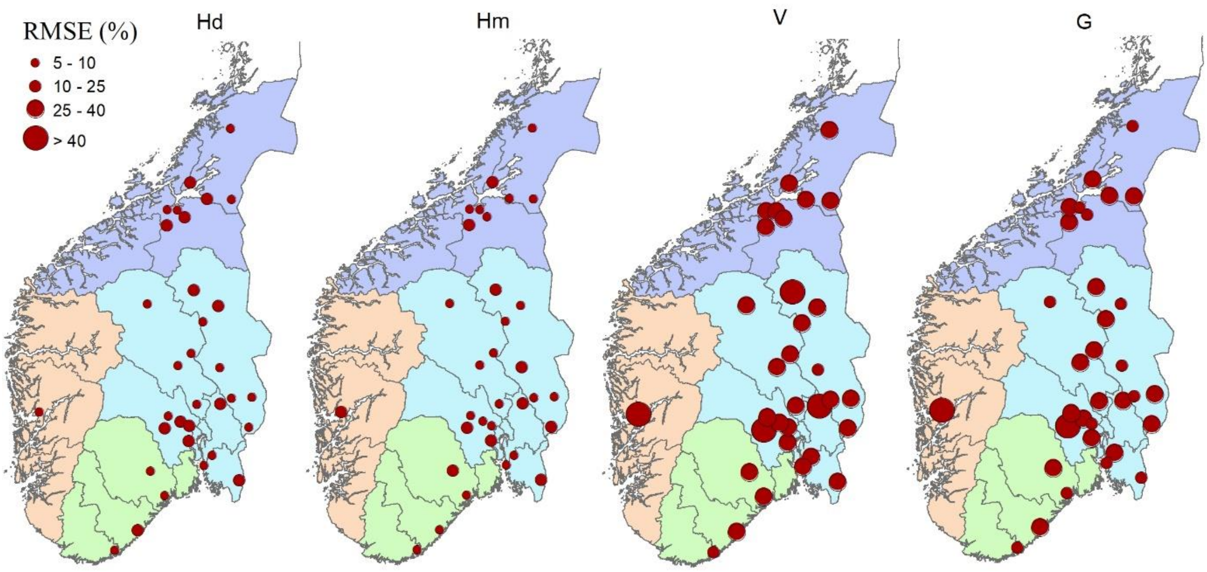

Figure 4. Geographic distribution of the relative root mean squared error (RMSE%) for the non-stratified SR16 predictions over the 33 FMI projects. The forest attributes are dominant height (Hd), mean height (Hm), basal area (G), and volume (V). The different colors represent the SR16 regions (A = purple, B = blue, C = orange, and D = green).

3.3. VIP analysis and factor importance

The VIP results for the different forest attributes are shown in Figure 5. The result showed that there were only slight differences in the VIP scores of the factors that were tested for correlation to the prediction errors, depending on if non-stratified or stratified SR16 prediction models were used. For the two height attributes (Hd and Hm), the factors with the largest VIP scores were in the forest conditions category, specifically height-diameter ratio (HDR), site index (SI), and the standard deviation of the HDR (HDR.sd). For HL, also the proportion of deciduous trees (p.D) was one of the most important factors.

Among the most important factors explaining the prediction errors of G and V, p.D was the only one considered important for both forest attributes and for both sets of predictions (non-stratified and stratified). Other important factors in common for G and V were the years of difference between the SR16 predictions and field reference value (Yeard.d), and the mean monthly temperature of the summer months (T.s).

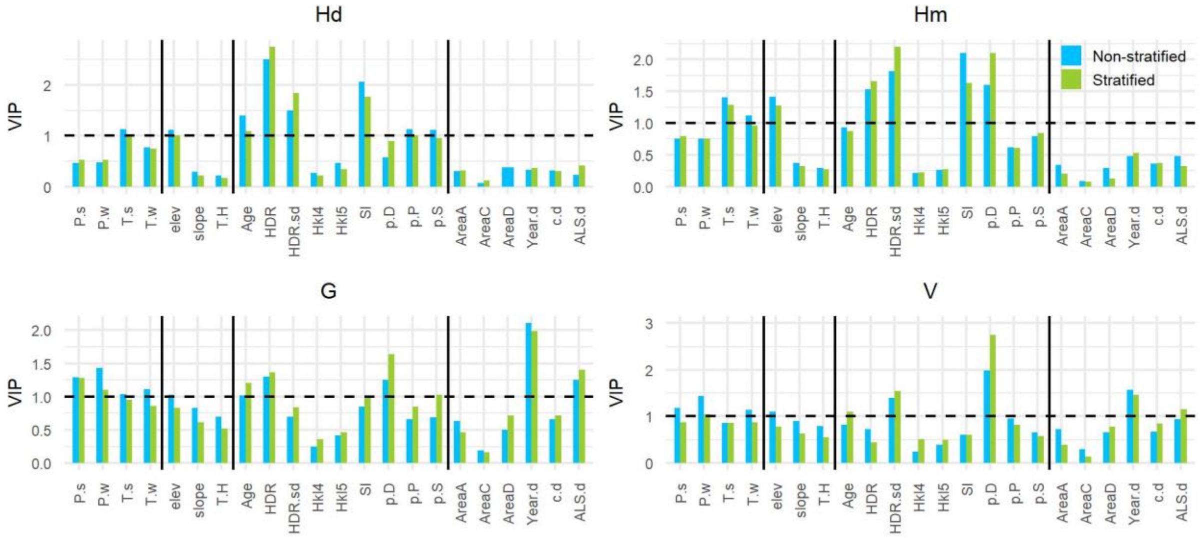

Figure 5. VIP values summarizing the global contribution for each factor to the PLS regression for the non-stratified (blue) and stratified (green) SR16 predictions. Vertical lines separate the different categories of factors analyzed in the PLS regression, see Table 2. The horizontal line indicates the threshold to consider a factor influential.

4. Discussion

This study assessed the accuracy at the level of prediction units (16 m × 16 m cells) of the SR16 forest resource map by comparison with ground reference values of more than 5000 field plots from 33 FMI projects distributed across southern Norway. The results showed the importance of assessing the accuracy at FMI project level because for some FMI projects, the MD% values of the SR16 predictions were substantial (>30% for G and V), which was not evident when the values were calculated across all plots and FMIs in a single calculation.

It was expected that the SR16 predictions were not equally accurate for all FMIs since they were distributed over a large spatial domain, and hence covering substantial ranges of the growth factors that determine stand structure. Regional models such as those used in SR16, calibrated on data from a larger region, will not be equally suitable for each individual and small AOI with its peculiarities, leading to the FMI-specific differences observed in the current study. In general, the challenge is that models developed from empirical data by means of regression, represent average forest conditions for the area from which the data are collected. Thus, a single FMI can comprise forest conditions that on average are different from the average conditions represented by the field plots used to calibrate the models, and systematic prediction errors might therefore occur locally for smaller spatial domains. On the level of a single forest property, this effect of dissimilarity in forest conditions might be even more pronounced, although we did not have property-specific data at hand in this study to illustrate empirically the potential consequences for individual properties. It is also a challenge that the magnitude of the systematic errors is unknown when external models are being applied unless actual field observations from the AOI are available from which the average prediction error could be observed by subtracting the observed plot values from the corresponding predicted values. In cases where such an estimate of the average systematic error is known, local predictions may be corrected. To use such plot observations for calibration of local effects together with predictions from a regional model such as the SR16, may be a cost-effective way of supporting local forest management planning since the number of local field observations could be substantially smaller than what is common practice in operational FMIs.

Among all factors studied in the PLSR analysis, the highest VIP scores were associated with p.D, HDR, HDR.sd, and SI from the category "forest condition”. p.D represents the proportion of deciduous trees in a plot, which has been well documented to affect the derived ALS metrics when keeping all other factors equal (Næsset, 2005, Liang et al., 2007, Villikka et al., 2012). HDR represents tree allometry where low HDR indicates large diameter relative to height, and conversely, high HDR indicates slim trees, typically found in dense stands where the trees compete for light (Hess et al., 2021). The standard deviation of HDR (HDR.sd) represents plot homogeneity. Forest attributes for uneven-aged forests with complex structures and different allometries are challenging to model, and consequently predictions for such forest types also tend to be associated with greater uncertainty compared to those of homogenous forest. SI represents the forest productivity, which is associated to different shapes of crowns and stems. Normally, poorer sites have shorter trees with more rounded crowns and open forest structures with scattered trees. It is not surprising that all these factors related with tree crowns and the structure of the forest are important since they will determine the properties of a prediction model, and thus the appropriateness of the application of a model across study areas (Yates et al., 2018, Tompalski et al., 2019).

Year.d from the category “other factors” was another factor with high VIP score, especially for G. The immediate interpretation is that our correction to adjust the SR16 predictions to the date of the field acquisition was not optimal. We assumed that the proportion of change for G was the same as for V, but a specific growth model for G could have been better because the tree's lateral growth responsible for diameter and basal area increment is sensitive to changes in density. Therefore, the growth of G and V could be slightly different at different stand densities. Another possible reason is undetected thinnings in the period between the data acquisitions. Depending on the thinning strategy and intensity, the thinnings may not substantially influence the stand heights (Skovsgaard and Vanclay, 2008) but will reduce the basal area. Therefore, the fact that year.d was important in the VIP analysis, emphasizes the importance of using proper growth models and procedures to detect disturbances when acquisitions are from different points in time.

From the VIP results, it is still challenging to provide general guidance on where SR16 will perform well, especially because we don't have access to the data used for modeling. However, we observed that the most important factors explaining the variability of the differences between predictions and ground reference values represented forest and canopy structure. This is illustrated by the prediction accuracies for the FMIs B1 and B19. Field data for both FMIs were acquired the same year, the FMIs were from the same region, and they had similar mean values of the forest attributes. The forest structures (represented by p.D, HDR, and SI) were nevertheless different (see Appendix A). Consequently, MD% values for Hd, Hm, G, and V differed substantially between the two FMIs (respectively -4%, -10%, 37%, and 34% for B1 and 0%, -3%, -2%, and 0% for B19).

To the extent of our knowledge, just studies from other ecoregions have examined factors influencing the results of regional models constructed with NFI data (Guerra-Hernández et al., 2022). However, we can compare our result against other Nordic nationwide forest attribute maps constructed with ALS and NFI plots (Nilsson et al., 2017), and other studies using SR16 predictions (Hauglin et al., 2021, Rahlf et al., 2021). Our main objective focused on the systematic differences, but there are not that many studies with results for MD%, and therefore, we will also discuss the RMSE, which is a commonly reported uncertainty statistic in forest inventory (Persson and Ståhl, 2020).

Nilsson et al. (2017) report plot-level results for three independent areas in northern, mid, and southern Sweden using leave-one-out cross-validation. The absolute MD ranged from 8 to 9% for HL, 16 to 22% for G, and 15 to 20% for V, i.e., larger than in the current study. However, the RMSE values from the Swedish study were similar in magnitude to ours, and even smaller for V, ranging from 10% to 11% for HL, 20% to 27% for G, and 19% to 25% for V.

Using non-stratified and species-specific SR16 models, Hauglin et al. (2021) reported overall plot level RMSE values ranging from 12 to 15% for HL, 31 to 33% for G, and 35 to 42% for V,. They concluded that the use of separate models for each main tree species improved the prediction accuracy. In the present study, the overall plot level RMSE% values were for all forest attributes smaller than the ones reported by Hauglin et al. (2021), but the stratified models did not always improve the results. In fact, in the current study, more predictions were considered outliers after using the stratified SR16 models compared to those made by the non-stratified models, meaning that the stratified predictions were further away from the true value in those plots. However, it should be mentioned that the stratified SR16 models were different between the two studies. Hauglin et al. (2021) only used three models based on dominant tree species, instead of our 12. A more detailed stratification has the advantage that the models potentially could be more accurate and precise if the stratification criteria are relevant. However, the downsides of constructing more detailed models over more general ones are that each stratum-wise model would have fewer ground reference plots available for model construction and that there would be a greater risk of applying a stratified model to a different stratum due to errors in the forest classification. In the studies compared here, stratum information was obtained from different sources that might be associated with different levels of uncertainty. Hauglin predicted the dominant tree species using a model dependent on metrics derived from Sentinel-2 images (Breidenbach et al., 2021), while we based the stratification entirely on manual stand-wise photointerpretation.

Rahlf et al. (2021) estimated the volume of mature spruce forest with the SR16 and an "adjusted" SR16 that added local sample plots in the modeling phase. The results for validation on independent FMI plots were -2% and 18% for MD and RMSE, respectively using the SR16, and -11% MD and 21% RMSE using the adjusted SR16, so no improvement was observed. The results from Rahlf et al. (2021) are within the ranges reported at the FMI project level in the present study. However, their validation data were limited to 60 plots distributed across six forest stands. The limited data material might have been a reason for the small effect of using local sample plot information.

Our results might also be compared (both RMSE and MD) to previous Nordic studies to show the differences between predictions made by regional (SR16) and local prediction models (traditional FMI) at the spatial level of an FMI. Næsset (2007) reviewed the results of accuracy assessments of several Nordic local FMIs. The studies were from various geographical regions, comprised different area sizes and used diverse numbers of training plots. Næsset (2007) reported MD values between -5 and 3%, and RMSE values between 3 and 6% for Hd and Hm, respectively. For G and V, the MD values ranged from -3.6 to 8.4%, and the RMSE from 10 to 21%. Using the study of Næsset to portrait what is reported in the literature for locally calibrated predictions, we concluded that local FMIs provide MD and RMSE values of greater quality than the SR16 predictions, so in terms of accuracy, local FMIs seems to be a better option. However, decisions concerning which data to use for a particular purpose should be based not only on desired levels of accuracy but also on the data acquisition costs relative to the benefit in terms of the suitability of the data for decision making (Kangas, 2010). Further studies should focus on this topic and analyze the benefits and costs of local FMIs versus a regional product such as SR16 using, for example, cost-plus-loss analysis as an analytical method. In cost-plus-loss analysis the economic losses caused by decisions based on inaccurate data are added to the forest inventory's total costs (Burkhart et al., 1978, Ruotsalainen et al., 2019). The method with the lowest total cost is considered the best alternative.

5. Conclusion

We examined the potential systematic prediction errors of the Norwegian SR16 forest resource map on data from 33 different local FMI projects across Norway, which represent a great diversity of forest conditions. The results show large MD% and RMSE% values for certain individual FMIs, which were not evident when all plots across all the FMIs were analyzed together. The use of stratification did not improve the predictions, and differences between FMIs were found to be caused by factors representing forest structure, such as the proportion of deciduous trees, the height-diameter ratio and site index. Thus, the use of SR16 for particular AOIs where the forest conditions deviate from the average forest conditions of the region for which the models were constructed, are prone to systematic prediction errors. To assess the magnitude of systematic prediction errors and as a means of correcting systematic prediction errors, a sample of local field plots is expedient. Expected losses due to suboptimal decisions caused by inaccurate data need to be considered before nationwide forest maps are used for forest management.

Acknowledgments

The authors would like to thank the four forest owner associations Allskog SA, AT Skog SA, Glommen Mjøsen Skog SA, and Viken Skog SA for sharing their ground reference data. We would also thank Stephanie Eisner for preparing the climatic data.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research has been funded by the private research fund Skogtiltaksfondet and the Norwegian forest trust fund (Utviklingsfondet for skogbruket).

Data availability statement

The SR16 maps used in this study are publicly available on the web (http://kilden.nibio.no, Skog, SR16). The field data are owned by private inventory companies and are not publicly available.

REFERENCES

ANON. 1987. Handbok for planlegging i skogbruket. Landbruksforlaget Oslo, Norway.

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.

BLINGSMO, K. 1988. Tilvekstfunksjoner. (Growth models) *Research paper from Norwegian Forest Research Institute, Ås*, 8 pp.

BRAASTAD, H. 1966. Volumtabeller for bjoerk [Volume tables for birch]. *(In Norwegian with English summary),* Meddelelser fra Det norske Skogforsøksvesen 21, 265-365.

BRANTSEG, A. 1967. Furu sønnafjells. Kubering av staaende skog. Funksjoner og tabeller [Volume functions and tables for Scots pine. South Norway]. *Meddelelser fra det Norske Skogforsøksvesen*.

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.

BREIDENBACH, J., WASER, L. T., DEBELLA-GILO, M., SCHUMACHER, J., RAHLF, J., HAUGLIN, M., PULITI, S. & ASTRUP, R. 2021. National mapping and estimation of forest area by dominant tree species using Sentinel-2 data. *Canadian Journal of Forest Research,* 51, 365-379.

BURKHART, H. E., STUCK, R. D., LEUSCHNER, W. A. & REYNOLDS, M. R. 1978. Allocating inventory resources for multiple-use planning. *Canadian Journal of Forest Research,* 8, 100-110.

CHIRICI, G., GIANNETTI, F., MCROBERTS, R. E., TRAVAGLINI, D., PECCHI, M., MASELLI, F., CHIESI, M. & CORONA, P. 2020. Wall-to-wall spatial prediction of growing stock volume based on Italian National Forest Inventory plots and remotely sensed data. *International Journal of Applied Earth Observation and Geoinformation,* 84, 101959.

DAYAL, B. S. & MACGREGOR, J. F. 1997. Improved PLS algorithms. *Journal of Chemometrics,* 11, 73-85.

DELBECK, K. 1965. Methods for calculation of increment for open forests. *Tidsskrift for Skogbruk,* 73, 5-45.

ERIKSSON, L., BYRNE, T., JOHANSSON, E., TRYGG, J. & VIKSTRÖM, C. 2013. *Multi-and megavariate data analysis basic principles and applications*, Umetrics Academy.

FITJE, A. & VESTJORDET, E. 1977. Stand height curves and new tariff tables for Norway spruce. Medd. Nor. inst. skogforsk. 34 (2): 23–68. *Norwegian with English summary*.

GOODWIN, N. R., COOPS, N. C. & CULVENOR, D. S. 2006. Assessment of forest structure with airborne LiDAR and the effects of platform altitude. *Remote Sensing of Environment,* 103, 140-152.

GSCHWANTNER, T., ALBERDI, I., BAUWENS, S., BENDER, S., BOROTA, D., BOSELA, M., BOURIAUD, O., BREIDENBACH, J., DONIS, J., FISCHER, C., GASPARINI, P., HEFFERNAN, L., HERVÉ, J.-C., KOLOZS, L., KORHONEN, K. T., KOUTSIAS, N., KOVÁCSEVICS, P., KUČERA, M., KULBOKAS, G., KULIEŠIS, A., LANZ, A., LEJEUNE, P., LIND, T., MARIN, G., MORNEAU, F., NORD-LARSEN, T., NUNES, L., PANTIĆ, D., REDMOND, J., REGO, F. C., RIEDEL, T., ŠEBEŇ, V., SIMS, A., SKUDNIK, M. & TOMTER, S. M. 2022. Growing stock monitoring by European National Forest Inventories: Historical origins, current methods and harmonisation. *Forest Ecology and Management,* 505, 119868.

GUERRA-HERNÁNDEZ, J., BOTEQUIM, B., BUJAN, S., JURADO-VARELA, A., MOLINA-VALERO, J. A., MARTÍNEZ-CALVO, A. & PÉREZ-CRUZADO, C. 2022. Interpreting the uncertainty of model-based and design-based estimation in downscaling estimates from NFI data: a case-study in Extremadura (Spain). *GIScience & Remote Sensing,* 59, 686-704.

HANSEN, M. C., POTAPOV, P. V., MOORE, R., HANCHER, M., TURUBANOVA, S. A., TYUKAVINA, A., THAU, D., STEHMAN, S. V., GOETZ, S. J., LOVELAND, T. R., KOMMAREDDY, A., EGOROV, A.,

CHINI, L., JUSTICE, C. O. & TOWNSHEND, J. R. 2013. High-resolution global maps of 21st-century forest cover change. *Science,* 342**,** 850-3.

HAUGLIN, M., RAHLF, J., SCHUMACHER, J., ASTRUP, R. & BREIDENBACH, J. 2021. Large scale mapping of forest attributes using heterogeneous sets of airborne laser scanning and National Forest Inventory data. *Forest Ecosystems,* 8**,** 65.

HESS, A. F., MINATTI, M., COSTA, E. A., SCHORR, L. P. B., DA ROSA, G. T., DE ARRUDA SOUZA, I., BORSOI, G. A., LIESENBERG, V., STEPKA, T. F. & ABATTI, R. 2021. Height-to-diameter ratios with temporal and dendro/morphometric variables for Brazilian pine in south Brazil. *Journal of Forestry Research,* 32**,** 191-202.

HILL, A., BUDDENBAUM, H. & MANDALLAZ, D. 2018. Combining canopy height and tree species map information for large-scale timber volume estimations under strong heterogeneity of auxiliary data and variable sample plot sizes. *European Journal of Forest Research,* 137**,** 489-505.

HUANG, C., GOWARD, S. N., MASEK, J. G., THOMAS, N., ZHU, Z. & VOGELMANN, J. E. 2010. An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks. *Remote Sensing of Environment,* 114**,** 183-198.

JUTRAS-PERREAULT, M.-C., GOBAKKEN, T. & ØRKA, H. O. 2021. Comparison of two algorithms for estimating stand-level changes and change indicators in a boreal forest in Norway. *International Journal of Applied Earth Observation and Geoinformation,* 98**,** 102316.

KANGAS, A., ASTRUP, R., BREIDENBACH, J., FRIDMAN, J., GOBAKKEN, T., KORHONEN, K. T., MALTAMO, M., NILSSON, M., NORD-LARSEN, T., NÆSSET, E. & OLSSON, H. 2018. Remote sensing and forest inventories in Nordic countries - roadmap for the future. *Scandinavian Journal of Forest Research,* 33**,** 397-412.

KANGAS, A. S. J. E. J. O. F. R. 2010. Value of forest information. 129**,** 863-874.

LIANG, X., HYYPPÄ, J. & MATIKAINEN, L. 2007. Deciduous-coniferous tree classification using difference between first and last pulse laser signatures. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences,* 36.

MALTAMO, M., PACKALEN, P. & KANGAS, A. 2021. From comprehensive field inventories to remotely sensed wall-to-wall stand attribute data — a brief history of management inventories in the Nordic countries. *Canadian Journal of Forest Research,* 51**,** 257-266.

MASSEY, A. & MANDALLAZ, D. 2015. Design-based regression estimation of net change for forest inventories. *Canadian Journal of Forest Research,* 45**,** 1775-1784.

MCROBERTS, R. E., TOMPPO, E. O. & NÆSSET, E. 2010. Advances and emerging issues in national forest inventories. *Scandinavian Journal of Forest Research,* 25**,** 368-381.

MOEN, A. 1999. *Vegetation,* Hønefoss, Norwegian Mapping Authority.

MONNET, J. M., GINZLER, C. & CLIVAZ, J. C. 2016. Wide-area mapping of forest with national airborne laser scanning and field inventory datasets. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.,* XLI-B8**,** 727-731.

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.

NÆSSET, E. 2005. Assessing sensor effects and effects of leaf-off and leaf-on canopy conditions on biophysical stand properties derived from small-footprint airborne laser data. *Remote Sensing of Environment,* 98**,** 356-370.

NÆSSET, E. 2007. Airborne laser scanning as a method in operational forest inventory: Status of accuracy assessments accomplished in Scandinavia. *Scandinavian Journal of Forest Research,* 22**,** 433-442.

NÆSSET, E. 2009. Effects of different sensors, flying altitudes, and pulse repetition frequencies on forest canopy metrics and biophysical stand properties derived from small-footprint airborne laser data. *Remote Sensing of Environment,* 113**,** 148-159.

NÆSSET, E. 2014. Area-Based Inventory in Norway – From Innovation to an Operational Reality. *In:* MALTAMO, M., NÆSSET, E. & VAUHKONEN, J. (eds.) *Forestry Applications of Airborne Laser Scanning: Concepts and Case Studies.* Dordrecht: Springer Netherlands.

NÆSSET, E., BOLLANDSÅS, O. M. & GOBAKKEN, T. 2005. Comparing regression methods in estimation of biophysical properties of forest stands from two different inventories using laser scanner data. *Remote Sensing of Environment,* 94**,** 541-553.

NILSSON, M., NORDKVIST, K., JONZÉN, J., LINDGREN, N., AXENSTEN, P., WALLERMAN, J., EGBERTH, M., LARSSON, S., NILSSON, L., ERIKSSON, J. & 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**,** 447-454.

NORD-LARSEN, T. & SCHUMACHER, J. 2012. Estimation of forest resources from a country wide laser scanning survey and national forest inventory data. *Remote Sensing of Environment,* 119**,** 148-157.

PERSSON, H. J. & STÅHL, G. 2020. Characterizing Uncertainty in Forest Remote Sensing Studies. *Remote Sensing,* 12**,** 505.

RAHLF, J., HAUGLIN, M., ASTRUP, R. & BREIDENBACH, J. 2021. Timber volume estimation based on airborne laser scanning — comparing the use of national forest inventory and forest management inventory data. *Annals of Forest Science,* 78**,** 49.

ROSNER, B. 1983. Percentage Points for a Generalized ESD Many-Outlier Procedure. *Technometrics,* 25**,** 165-172.

RUOTSALAINEN, R., PUKKALA, T., KANGAS, A., VAUHKONEN, J., TUOMINEN, S. & PACKALEN, P. 2019. The effects of sample plot selection strategy and the number of sample plots on inoptimality losses in forest management planning based on airborne laser scanning data. *Canadian Journal of Forest Research,* 49**,** 1135-1146.

SHARMA, R. P., BRUNNER, A., EID, T. & OYEN, B. H. 2011. Modelling dominant height growth from national forest inventory individual tree data with short time series and large age errors. *Forest Ecology and Management,* 262**,** 2162-2175.

SKAUGEN, T. E., HANSSEN-BAUER, I. & FØRLAND, E. J. 2003. Adjustment of dynamically downscaled temperature and precipitation data in Norway. (RegClim Report No. 20/02). Norwegian Meteorological Institute

SKOVSGAARD, J. P. & VANCLAY, J. K. 2008. Forest site productivity: a review of the evolution of dendrometric concepts for even-aged stands. *Forestry: An International Journal of Forest Research,* 81**,** 13-31.

TOMPALSKI, P., WHITE, J. C., COOPS, N. C. & WULDER, M. A. 2019. Demonstrating the transferability of forest inventory attribute models derived using airborne laser scanning data. *Remote Sensing of Environment,* 227**,** 110-124.

TOMPPO, E., GSCHWANTNER, T., LAWRENCE, M., MCROBERTS, R. & GODINHO-FERREIRA, P. 2010. *National forest inventories: Pathways for common reporting*.

TVEITO, O. E., BJØRDAL, I., SKJELVÅG, A. O. & AUNE, B. 2005. A GIS-based agro-ecological decision system based on gridded climatology. *Meteorological Applications,* 12**,** 57-68.

VAN DER VOET, H. 1994. Comparing the predictive accuracy of models using a simple randomization test. *Chemometrics and Intelligent Laboratory Systems,* 25**,** 313-323.

VEGA, C., RENAUD, J.-P., SAGAR, A. & BOURIAUD, O. 2021. A new small area estimation algorithm to balance between statistical precision and scale. *International Journal of Applied Earth Observation and Geoinformation,* 97**,** 102303.

VERBESSELT, J., ZEILEIS, A. & HEROLD, M. 2012. Near real-time disturbance detection using satellite image time series. *Remote Sensing of Environment,* 123**,** 98-108.

VESTJORDET, E. 1967. Functions and tables for volume of standing trees. Norway spruce. *Meddelelser fra det Norske Skogforsoksvesen***,** 539–574.

VILLIKKA, M., PACKALÉN, P. & MALTAMO, M. 2012. *The suitability of leaf-off airborne laser scanning data in an area-based forest inventory of coniferous and deciduous trees*.

WASER, L. T., GINZLER, C. & REHUSH, N. 2017. Wall-to-Wall Tree Type Mapping from Countrywide Airborne Remote Sensing Surveys. *Remote Sensing,* 9**,** 766.

WHITE, J. C., WULDER, M. A., VARHOLA, A., VASTARANTA, M., COOPS, N. C., COOK, B. D., PITT, D. & WOODS, M. 2013. A best practices guide for generating forest inventory attributes from airborne laser scanning data using an area-based approach. *The Forestry Chronicle,* 89**,** 722-723.

WOLD, S., SJÖSTRÖM, M. & ERIKSSON, L. 2001. PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems,* 58**,** 109-130.

YATES, K. L., BOUCHET, P. J., CALEY, M. J., MENGERSEN, K., RANDIN, C. F., PARNELL, S., FIELDING, A. H., BAMFORD, A. J., BAN, S., BARBOSA, A., DORMANN, C. F., ELITH, J., EMBLING, C. B., ERVIN, G. N., FISHER, R., GOULD, S., GRAF, R. F., GREGR, E. J., HALPIN, P. N., HEIKKINEN, R. K., HEINÄNEN, S., JONES, A. R., KRISHNAKUMAR, P. K., LAURIA, V., LOZANO-MONTES, H., MANNOCCI, L., MELLIN, C., MESGARAN, M. B., MORENO-AMAT, E., MORMEDE, S., NOVACZEK, E., OPPEL, S., CRESPO, G. O., PETERSON, A. T., RAPACCIUOLO, G., ROBERTS, J. J., ROSS, R. E., SCALES, K. L., SCHOEMAN, D., SNELGROVE, P., SUNDBLAD, G., THUILLER, W., TORRES, L. G., VERBRUGGEN, H., WANG, L., WENGER, S., WHITTINGHAM, M. J., ZHARIKOV, Y., ZURELL, D. & SEQUEIRA, A. M. M. 2018. Outstanding Challenges in the Transferability of Ecological Models. *Trends in Ecology & Evolution,* 33**,** 790-802.

Appendix A.

Table A1. Summary of the factors that represent the differences between forest inventory projects (FMI). The factors are defined in Table 2.

	P.s		P.w		T.s		T.w		elev		slope		T.H	
FMI	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
A1	98.28	11.49	193.57	26.39	12.71	0.59	-2.29	0.92	250	84	16.60	8.83	1.55	1.25
A2	93.85	4.03	110.48	14.15	12.33	0.67	-3.02	0.64	310	87	12.99	7.08	1.17	0.89
A3	90.12	8.13	99.17	17.87	12.38	0.66	-3.19	0.69	314	91	14.21	7.58	1.24	1.02
A4	100.67	7.78	106.56	8.22	11.88	0.68	-4.37	0.69	412	92	9.29	4.10	0.84	0.65
A5	83.76	9.35	122.96	23.57	12.64	0.59	-2.44	0.76	241	83	15.60	7.81	1.57	1.00
A6	98.01	3.97	168.62	12.41	13.03	0.53	-2.85	0.85	100	70	12.67	8.04	1.17	1.08
A7	80.07	4.22	96.54	7.18	12.63	0.65	-2.65	0.78	260	100	14.90	7.39	1.38	1.03
A8	105.03	12.94	111.97	13.56	13.05	0.67	-2.35	0.80	208	84	14.39	7.56	1.26	1.01
B1	75.45	3.05	36.42	3.56	11.31	0.68	-7.90	0.31	634	101	10.13	6.47	0.95	0.83
B2	87.99	1.60	89.08	8.54	15.17	0.35	-1.71	0.54	154	37	6.93	5.21	0.65	0.59
B3	50.84	10.27	59.28	18.68	11.13	1.39	-6.18	1.03	667	158	15.83	7.51	1.45	1.04
B4	88.05	3.38	68.17	2.80	14.49	0.44	-4.44	0.37	223	58	7.94	5.31	0.72	0.63
B5	93.52	3.52	72.96	7.24	14.04	0.73	-4.88	0.48	346	101	10.58	6.81	0.96	0.89
B6	85.58	3.89	60.60	4.57	14.08	0.49	-5.34	0.28	340	84	9.25	6.25	0.79	0.67
B7	88.22	7.66	64.24	12.18	13.52	0.94	-5.02	0.62	427	123	10.50	7.01	0.94	0.87
B8	83.31	5.63	50.79	7.24	13.99	0.82	-5.95	0.48	331	113	4.67	3.67	0.41	0.45
B9	86.75	2.54	83.02	1.79	15.47	0.44	-2.35	0.41	120	50	6.57	4.63	0.62	0.54
B10	93.43	5.15	59.99	5.33	13.45	0.55	-4.00	0.30	432	65	8.41	4.64	0.82	0.61
B11	109.06	2.41	62.83	3.36	13.78	1.20	-5.08	0.61	356	132	13.27	7.46	1.18	1.02
B12	104.57	9.87	71.08	7.58	12.29	1.11	-6.77	0.49	583	163	9.62	6.00	0.88	0.75
B13	92.39	5.33	77.96	13.18	14.83	0.98	-3.21	0.75	230	123	11.48	7.56	1.13	1.02
B14	82.43	1.69	81.96	1.03	15.93	0.20	-1.69	0.31	57	23	6.22	4.57	0.64	0.54
B15	93.10	2.77	65.51	5.95	14.24	0.54	-5.00	0.32	311	80	9.01	5.48	0.82	0.64
B16	106.70	5.60	62.83	7.58	12.68	1.25	-6.47	0.65	534	169	10.37	7.34	0.92	0.84
B17	87.30	9.60	47.35	8.74	11.92	0.91	-8.01	0.38	567	142	9.99	6.50	0.93	0.86
B18	103.63	5.86	66.46	8.44	13.61	1.11	-4.76	0.67	372	146	9.36	5.64	0.87	0.68
B19	92.51	12.13	49.08	12.14	11.59	1.14	-7.74	0.35	637	167	11.19	6.16	1.01	0.81
B20	84.48	3.20	45.66	7.18	14.60	0.71	-4.27	0.32	267	84	9.60	5.64	0.81	0.64
C1	178.09	25.01	321.91	76.68	13.67	1.02	1.11	1.50	129	118	17.51	8.59	1.58	1.25
D1	101.00	4.72	134.85	9.15	15.67	0.29	-0.54	0.53	100	45	9.76	5.81	0.90	0.67
D2	95.27	2.49	107.62	5.69	15.68	0.28	-1.20	0.47	112	44	12.93	7.94	1.10	0.96
D3	104.97	7.88	86.79	13.77	14.24	1.25	-3.29	1.07	293	150	12.21	7.21	1.09	0.90
D4	113.56	9.05	175.97	21.66	15.20	0.39	0.15	0.64	119	61	12.38	6.40	1.13	0.81

	Age		HDR		HDR.sd		p.D		p.P		p.S		SI	
FMI	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
A1	115	46	68.59	13.78	9.23	5.33	0.79	0.26	0.79	0.26	0.06	0.20	11.85	3.03
A2	84	32	79.88	14.83	12.57	5.21	0.60	0.33	0.60	0.33	0.27	0.33	13.79	3.75
A3	106	43	77.61	14.58	17.16	4.16	0.59	0.36	0.59	0.36	0.32	0.37	11.44	3.58
A4	78	43	82.12	10.85	10.12	3.68	0.73	0.23	0.73	0.23	0.09	0.19	11.98	2.00
A5	86	38	79.38	14.45	13.54	5.99	0.74	0.26	0.74	0.26	0.13	0.22	12.85	2.93
A6	85	48	84.97	16.89	13.74	7.24	0.80	0.27	0.80	0.27	0.10	0.26	12.54	2.99
A7	91	39	75.44	12.90	11.18	4.09	0.79	0.25	0.79	0.25	0.14	0.25	12.62	3.49
A8	86	39	85.06	14.77	13.34	4.97	0.73	0.29	0.73	0.29	0.17	0.29	12.83	3.20
B1	93	36	75.84	13.90	9.04	3.20	0.11	0.24	0.11	0.24	0.81	0.30	9.44	2.26
B2	66	33	85.01	15.45	10.85	4.92	0.45	0.35	0.45	0.35	0.46	0.37	14.93	4.36
B3	106	44	69.95	13.80	10.47	4.93	0.07	0.21	0.07	0.21	0.80	0.28	9.57	1.91
B4	64	29	96.83	14.89	13.30	6.07	0.55	0.37	0.55	0.37	0.37	0.38	15.53	3.28
B5	65	30	87.57	12.68	11.83	5.78	0.77	0.29	0.77	0.29	0.16	0.28	15.91	3.41
B6	71	32	89.30	14.82	10.89	4.42	0.59	0.35	0.59	0.35	0.33	0.35	14.22	3.36
B7	70	35	89.34	14.81	12.52	5.99	0.70	0.32	0.70	0.32	0.20	0.31	14.69	3.63
B8	57	26	94.88	17.14	13.30	6.54	0.72	0.32	0.72	0.32	0.18	0.33	17.54	2.78
B9	74	45	90.47	17.73	13.36	5.22	0.54	0.34	0.54	0.34	0.34	0.36	16.07	5.33
B10	90	53	76.41	10.80	9.96	3.73	0.88	0.18	0.88	0.18	0.04	0.13	14.08	3.24
B11	78	34	81.46	16.81	10.67	4.18	0.41	0.37	0.41	0.37	0.43	0.42	14.35	4.28
B12	77	41	79.05	12.74	9.89	4.16	0.89	0.20	0.89	0.20	0.05	0.16	13.58	4.35
B13	86	39	78.19	15.24	11.89	4.26	0.53	0.35	0.53	0.35	0.35	0.39	14.79	4.83
B14	76	36	87.96	17.95	13.92	6.20	0.37	0.31	0.37	0.31	0.49	0.38	14.60	3.84
B15	71	32	89.30	14.63	12.03	4.80	0.52	0.35	0.52	0.35	0.41	0.37	14.78	3.85
B16	92	28	80.94	14.46	9.93	3.86	0.77	0.32	0.77	0.32	0.18	0.32	12.76	2.97
B17	81	33	78.97	14.50	9.15	4.35	0.41	0.39	0.41	0.39	0.46	0.42	11.60	3.21
B18	90	46	80.72	15.07	11.27	4.23	0.39	0.34	0.39	0.34	0.50	0.39	13.82	4.47
B19	83	39	78.58	14.79	10.56	4.31	0.53	0.39	0.53	0.39	0.35	0.40	11.63	3.85
B20	90	55	77.28	12.90	9.08	3.34	0.15	0.24	0.15	0.24	0.81	0.27	12.53	2.56
C1	87	41	77.77	18.03	13.45	6.50	0.33	0.41	0.33	0.41	0.52	0.40	15.03	6.68
D1	82	34	77.56	14.67	13.33	5.14	0.37	0.35	0.37	0.35	0.49	0.34	14.35	3.46
D2	88	34	80.18	17.86	14.60	5.57	0.38	0.32	0.38	0.32	0.47	0.35	13.29	3.17
D3	91	40	78.77	16.96	12.59	5.30	0.30	0.33	0.30	0.33	0.61	0.38	12.60	3.59
D4	91	40	72.78	16.52	12.22	5.12	0.24	0.33	0.24	0.33	0.53	0.33	13.54	3.74

	ALS.d		Year.d	c.d*	HKl3*	Hkl4*	Hkl5*
FMI	mean	sd	mean				
A1	4	2	3	0 %	19 %	14 %	67 %
A2	3	0	1	0 %	22 %	32 %	46 %
A3	4	0	6	0 %	22 %	14 %	64 %
A4	1	0	0	0 %	59 %	17 %	24 %
A5	3	0	1	0 %	28 %	27 %	45 %
A6	2	0	0	0 %	46 %	3 %	51 %
A7	5	0	4	0 %	28 %	19 %	53 %
A8	4	1	0	0 %	35 %	27 %	38 %
B1	3	0	2	0 %	34 %	23 %	43 %
B2	4	0	1	100 %	44 %	29 %	27 %
B3	6	1	5	100 %	29 %	17 %	54 %
B4	3	0	1	100 %	45 %	31 %	24 %
B5	2	0	1	100 %	43 %	30 %	27 %
B6	5	0	3	100 %	41 %	35 %	24 %
B7	4	0	3	100 %	43 %	32 %	25 %
B8	3	0	0	100 %	39 %	51 %	10 %
B9	4	0	0	100 %	41 %	20 %	40 %
B10	3	0	2	14 %	14 %	47 %	38 %
B11	3	0	3	25 %	25 %	34 %	42 %
B12	5	0	4	100 %	38 %	31 %	31 %
B13	2	0	0	12 %	12 %	38 %	51 %
B14	4	0	0	100 %	41 %	23 %	36 %
B15	4	1	3	100 %	28 %	49 %	23 %
B16	3	0	2	14 %	14 %	37 %	49 %
B17	2	1	0	100 %	34 %	37 %	29 %
B18	2	0	0	16 %	16 %	39 %	45 %
B19	3	0	2	100 %	41 %	25 %	34 %
B20	3	0	2	6 %	6 %	45 %	48 %
C1	3	1	7	0 %	17 %	33 %	50 %
D1	3	0	2	0 %	23 %	26 %	51 %
D2	4	2	3	0 %	26 %	16 %	58 %
D3	4	1	1	0 %	28 %	21 %	50 %
D4	0	0	3	0 %	32 %	18 %	50 %

* Percentage of plots in the FMI