Norges miljø- og
biovitenskapelige
universitet

Masteroppgave 2017 30 stp
Fakultet for miljøvitenskap og naturforvaltning

# TAKSERING AV UNGSKOG MED FLYBÅREN LASERSKANNING

## INVENTORY OF YOUNG FOREST USING AIRBORNE LASER SCANNING

Lennart Noordermeer
Skogfag

# FORORD

Gjennom mine studier har jeg utviklet stor interesse for skogbruksplanlegging, ressurskartlegging, og teknologiske utviklinger innen skogbruket. I mitt siste studieår kontaktet jeg derfor Mjøsen Skog sin planavdeling i håp om å finne et aktuelt og spennende tema for min masteroppgave. Det viste seg at det var behov for en grundig evaluering av metoder for laserbasert taksering av ungskog. Planavdelingen, sammen med Blom Geomatics, hadde nettopp tatt i bruk resultatene fra samarbeidsprosjektet mellom MINA og skogplanselskapene 'Laser2'. Mjøsen Skog hadde utført den første kommersielle taksten der ulike egenskaper av ungskogbestand ble kartlagt ved bruk av laserdata i Grue Vest, og resultatene ga rom for forbedring. Jeg var veldig interessert i å undersøke dette temaet nærmere, spesielt på grunn av forbedringspotensialet knyttet til metodene og de praktiske implikasjonene av prosjektet.

Jeg vil takke min hovedveileder, Dr. Ole Martin Bollandsås, for god hjelp og veiledning, samt Dr. Hans Ole Ørka, professor Erik Næsset og professor Terje Gobakken for gode råd og konstruktive innspill. I tillegg vil jeg takke Roar Økseter og Asunción Semper-Pascual for teknisk assistanse. En stor takk rettes til Mjøsen Skog sin planavdeling og Blom Geomatics for gode idéer og de nødvendige dataene, og til Fylkesmannen i Hedmark for finansiering av feltarbeidet.

Norges miljø- og biovitenskapelige universitet

Ås, 15. mai 2017

Lennart Noordermeer

## SAMMENDRAG

Gode data om egenskaper til ungskog er avgjørende for å kunne ta riktige beslutninger knyttet til skogskjøtsel, men fremskaffing av nøyaktige prediksjoner av disse egenskapene ut fra flybåren laser skanning (FLS) data er utfordrende. Hovedmålet med denne oppgaven var å videreutvikle, sammenligne og validere ulike metoder for laserbasert taksering av ungskog, og demonstrere en praktisk tilnærming for kartlegging av behov for ungskogpleie.

Regresjonsmodeller ble kalibrert for FLS-basert prediksjon av dominerende høyde (*Hd*), total gjennomsnittlig trehøyde (*Ht*), totalt treantall (*Nt*) og regulert treantall (*Nd*). To uavhengige datasett ble benyttet til modellkalibrering og validering, bestående av 45 og 64 prøveflater med areal på 256 $m^2$, som ble lagt ut i grandominert ungskog med høyde < 8 m. Høyde- og tetthetsvariabler som representerer den romlige strukturen av punktskyen ble beregnet for prøveflatene, samt en rekke laservariabler som representerer variasjonen av høyde- og tetthetsvariabler innen prøveflatene (*sd*-variabler). Prediksjonsevnen av modellene ble evaluert, der det ble oppnådd gjennomsnittlige prediksjonsfeil (RMSE) på 0.57-0.92, 0.59-1.05, 3412-7059, 327-486 for henholdsvis *Hd*, *Ht*, *Nt* og *Nd*. Inkludering av s*d*-variablene i modellseleksjonen førte til en lavere RMSE for *Hd*, *Ht* og *Nd*, men differansen mellom residualene for de alternative modellene var ikke statistisk signifikant.

Binær logistisk regresjon ble benyttet for prediksjon av behov for ungskogpleie. To ulike sett med prediktorer ble testet: (1) totalt treantall, dominerende høyde og bonitet, og (2) et utvalg av laservariabler og bonitet. Nøyaktigheten av klassifikasjonene ble evaluert ved bruk av stratifisert "2-fold" kryssvalidering, der metodene ga en samlet nøyaktighet på 76 % og 83 % med kappa-indeks på 0.44 og 0.60 for henholdsvis den første og andre modellen. Resultatet tyder på at direkte klassifisering av behov for ungskogpleie ut fra laservariabler er den foretrukne metoden.

Denne oppgaven presenterer en praktisk tilnærming for taksering av ungskog og kartlegging av behov for ungskogpleie. Det ble anbefalt å (1) skaffe et variert datasett til modellkalibrering, (2) aggregere prediksjoner for individuelle beregningsceller, og (3) klassifisere behov for ungskogpleie direkte ut fra bestands- og FLS-data. I tillegg tyder resultatene på at feltarbeidskostnader i operasjonelle takster kan reduseres ved å identifisere ungskogbestand der feltbefaring ikke er nødvendig.

## ABSTRACT

Predicting attributes of young forest is crucial for making sylvicultural treatment decisions in forest management planning. However, producing accurate predictions of such attributes is challenging. In this study, airborne laser scanning (ALS) data were used to develop, compare, and validate various methods for inventory of young forest, and to demonstrate a practical approach for mapping the need for pre-commercial thinning.

Two separate datasets of 45 and 64 sample plots of 256 $m^2$, located in Norway spruce-dominated regeneration stands (< 8 m) in the south-east of Norway, were used to predict dominant height (*Hd*), mean tree height (*Ht*), total stem density (*Nt*) and dominant stem density (*Nd*). Various ALS-derived canopy height and density metrics were calculated along with explanatory variables representing the variation of height and density metrics within sample plots (*sd*-variables). Validation of the predictions revealed root mean squared errors (RMSE) of 0.57-0.92, 0.59-1.05, 3412-7059, 327-486 for *Hd*, *Ht*, *Nt* og *Nd* respectively. Including *sd*-variables in the model selection led to a lower RMSE for *Hd*, *Ht* and *Nd*, but the differences among residuals of the alternative models were not statistically significant.

In addition, the need for pre-commercial thinning was predicted using binary logistic regression. Two different sets of predictors were tested: (1) stem density, dominant height, and site index, and (2) ALS-derived canopy height and density metrics, and site index. Stratified 2-fold cross validation revealed overall accuracies of 76 % and 83 %, and kappa coefficients of 0.44 and 0.60 for the first and second methods respectively, suggesting that classification into tending need classes based on ALS-metrics and site index is the preferred approach.

Based on the obtained results, this thesis provides recommendations for ALS-based inventory of young forest. Methods for predicting attributes of young forest can be improved by (1) obtaining an ample variety within datasets used for model calibration, (2) aggregating predictions for individual grid cells, and (3) implementing direct classification of the need for tending using ALS-data and site index. The obtained results also suggest that costs can be reduced in operational forest management inventories by identifying stands for which a field check is not required.

# Innhold

1 INNLEDNING ..... 3
1.1 Bakgrunn ..... 3
1.2 Ungskog ..... 4
1.3 Problemstilling ..... 6
2 MATERIALE OG METODE ..... 7
2.1 Prosjektområdet ..... 7
2.2 Laserdata ..... 8
2.3 Feltdata ..... 9
2.4 Feltregistrering ..... 12
2.5 Registrering på småflater ..... 13
2.6 Alternative forklaringsvariabler ..... 14
2.7 Modellutvalg ..... 15
2.8 Transformasjoner ..... 18
2.9 Prediksjon av biofysiske egenskaper til ungskog ..... 19
2.10 Prediksjon av behov for ungskogpleie ..... 20
2.10.1 Alternative metoder ..... 20
2.10.2 Logistisk regresjon ..... 20
2.10.3 Stratifisert "2-fold" kryssvalidering ..... 21
3 RESULTATER ..... 23
3.1 Modeller for prediksjon av egenskaper til ungskog ..... 23
3.1.1 Lineære regresjonsmodeller ..... 23
3.1.2 Mixed effects modeller ..... 24
3.2 Validering ..... 25
3.2.1 Kryssvalidering av lineære regresjonsmodeller ..... 25
3.2.2 Validering av lineære regresjonsmodeller med et uavhengig datasett ..... 26
3.2.3 Validering av lineære regresjonsmodeller med aggregerte prøveflater ..... 27
3.2.4 Kryssvalidering av mixed effects modeller ..... 28
3.2.5 Validering av mixed effects modeller med et uavhengig datasett ..... 29
3.3 Signifikans av sd-variablene ..... 30
3.4 Prediksjon av behov for ungskogpleie ..... 30
3.4.1 Klassifisering ut fra treantall, høyde og bonitet ..... 30

3.4.2 Direkte klassifisering ut fra laservariabler og bonitet ..... 31
3.4.3 Validering av modell 1 ..... 32
3.4.4 Validering av modell 2 ..... 33
4 DISKUSJON ..... 34
4.1 Prediktorvariabler ..... 34
4.2 Modeller for prediksjon av treantall og høyde ..... 36
4.3 Prediksjonsevne av modellene ..... 38
4.4 Metoder for prediksjon av behov for ungskogpleie ..... 41
5 IMPLIKASJONER OG ANBEFALINGER ..... 46
5.1 Variasjon innen datasettet ..... 46
5.2 Aggregering ..... 46
5.3 Direkte klassifisering ..... 47
5.4 Bestandsdata om bonitet ..... 47
5.5 Praktisk bruk av dataene ..... 48
5.6 Konklusjon og fremtidig forskning ..... 48
LITERATUR ..... 50
Vedlegg 1 - Flybåren laserskanning ..... 54
Vedlegg 2 - Arealmetoden ..... 55
Vedlegg 3 - Transformasjoner ..... 56
Vedlegg 4 - Boksdiagram for 2015- og 2016-datasettene ..... 60
Vedlegg 5 - Boksdiagram for clustrene ..... 61
Vedlegg 6 - Bonitetsklasser ..... 62
Vedlegg 7 - Resultatoversikt validering ..... 62
Vedlegg 8 - Prediksjon med observerte verdier ..... 63
Vedlegg 9 - Eksempler på kartlegging av behov for ungskogpleie ..... 64

# 1 INNLEDNING

## 1.1 Bakgrunn

For å kunne ta riktige beslutninger knyttet til skogskjøtsel, må skogeieren formulere målsettinger, slik at utviklingen av skogarealet og skogsdriften kan styres i ønsket retning for å nå målene. Fornuftige mål kan for eksempel være å opprettholde et balansekvantum, å maksimere lønnsomheten, eller å ta hensyn til rekreasjon og naturmangfold på skogeiendommen. For å utarbeide meningsfulle mål kreves det informasjon om dagens situasjon, dagens trender, og en nøyaktig oversikt over ressursene som et skogareal inneholder (Eid, 2008). En skogtakst danner dermed grunnlaget for fremtidig skogforvaltning, og er derfor en elementær del av planleggingsprosessen innen skogbruket.

Skogtakster tar sikte på å kartlegge skogressurser ved å formidle data for hver behandlingsenhet (bestand) for taktisk og strategisk planlegging. Gjennom de siste to tiår har flybåren laserskanning (FLS) etter arealmetoden blitt den ledende takstmetoden for skogbruksplanlegging i Norge (Naesset, 1997a, 1997b; Næsset, 2002). FLS systemer var i utgangspunktet utviklet for topografiske formål, hovedsakelig for å generere digitale terrengmodeller (DTM) (Ussyshkin et al., 2010). FLS er en fjernmålingsteknikk, der en aktiv "Light Detection and Ranging" (LiDAR) laser sender ut pulser av nær infrarødt lys (1.064 μm), med en typisk frekvens på ca. 150 kHz (Evans et al., 2009).

Et FLS system sender ut pulser av lysenergi mot bakken med en laser. Under laserskanningen blir punktene distribuert over flykorridoren av en skanningsmekanisme som mest ofte er et oscillerende eller roterende speil (Nordkvist et al., 2013). Sensoren registrerer de eksakte tidspunktene der pulsene blir sendt ut. Laserpulsene treffer bakken, trær, og ulike objekter i flykorridoren, og reflekteres til sensoren som igjen logger tidspunktet der tilbakespredningssignalet blir mottatt av sensoren. Videre blir avstanden mellom sensoren og hvert ekko beregnet med utgangspunkt i lysets hastighet som er konstant. Lasersystemet stedfester ekkoene ved hjelp av et "Global Positioning System" (GPS) og et "Inertial Navigation System" (INS) som posisjonerer ekkoene etter orientering av flyet ("pitch", "yaw", "roll"), og skannevinkel (illustrasjon i Vedlegg 1). Laserekko får tildelt planimetriske koordinater og ellipsoidiske høyder (x,y,z), og resulterende data gir stedfestede, tredimensjonal informasjon i form av en punktsky (illustrasjon i Vedlegg 1) som tilbyr muligheten til å estimere ulike romlige biofysiske egenskaper til trær og skogstrukturer over et stort område

innen kort tid (White et al., 2013). Høyde over bakken til hvert ekko innen punktskyen, eller 'z verdi', kan så beregnes ved å finne differansen mellom terrengnivået og høyden av punktet.

FLS instrumenter som benyttes innen skoginventering er typisk diskrete retursystemer med små strålediameter, eller fotavtrykk, som registrerer et flertall, mest ofte fire eller fem, ekko per puls (Maltamo et al., 2014). Diameteren av laserstrålen er avhengig av både flyhøyde og strålens divergens, men et fotavtrykk av ca. en meter er typisk (Maltamo et al., 2014). Utsendte laserpulser genererer ulike typer data da de blir reflektert fra trekroner, vegetasjon, og skogbunnen. Den romlige fordelingen av registrerte ekko bestemmes av algoritmen som regulerer ved hvilken terskelverdi for signalstyrke et ekko blir registrert av sensoren. En vanlig metode er å registrere en retur når en topp på retursignalet har nådd en viss prosentandel, for eksempel 50 %, av sin maksimumsverdi (illustrasjon i Vedlegg 1) (Nordkvist et al., 2013). Topper som ikke når deteksjonsgrensen blir ikke registrert.

Kun et par år etter at de første kommersielle LiDAR-instrumenter ble introdusert på markedet ble det foretatt en rekke studier der laserteknologien ble anvendt på småskala skogområder med formål å estimere ulike skogvariabler som trehøyde, stående volum, treantall, og diameterfordeling (Nilsson, 1996; Naesset, 1997a, 1997b; Magnussen et al., 1998). Et stort potensial ble påvist for bruk av teknologien innen skogtaksering, og Næsset (2002) utviklet arealmetoden som nå er den mest brukte takstmetoden for skogbruksplanlegging i Norge.

Næsset (2002) presenterte arealmetoden som en to stegs prosedyre. I det første steget blir regresjonsmodeller kalibrert med feltobserverte målinger som beskriver de biofysiske egenskapene til skogen som responsvariabler, og høyde- og tetthets laservariabler som blir ekstrahert fra laserpunktskyen som prediktorer. I det andre steget benyttes regresjonsmodellene til å predikere egenskapene over hele takstområdet (en beskrivelse av arealmetoden er gitt i Vedlegg 2).

### 1.2 Ungskog

Ungskogbestand er problematisk i dagens tilnærming for skoginventering siden arealmetoden ikke egner seg til nøyaktig estimering av tetthet og treslagssammensetning i yngre skog (Næsset et al., 2001; Korhonen et al., 2013; Ørka et al., 2016). Siden lasertakst har blitt tatt i bruk er dagens informasjon om ungskogareal dårligere enn før, da kostnadskrevende feltbefaringer ble gjennomført (Ørka et al., 2015).

Ungskogsfasen er et avgjørende stadium i et bestands omløp der det kan utføres tiltak for å påvirke bestandets utvikling med hensyn til treslagssammensetning, fremtidig vekst, kvalitet og tetthet. Formålet med ungskogpleie er å redusere tetthet i et bestand og begunstige trær med ønskelige egenskaper ved å fjerne konkurrerende individer av dårligere kvalitet og lavere vekstpotensial (Anon., 1969; Fahlvik, 2005). Valg av riktig tidspunkt der ungskogspleien utføres kan ha stor innvirkning på vekst- og kvalitetsegenskapene til et bestand, og avgjører i tillegg lønnsomheten av den første tynningen (Huuskonen et al., 2006). Tidspunkt for ungskogpleie beskrives best etter høyden av dominerende trær, og en høyde på fem meter er ofte typisk for utføring av tiltaket (Varmola et al., 2004). Selv om utføring av ungskogpleie medfører ulike fordeler, blir tiltaket ofte neglisjert av skogeiere, noe som kan resultere i betydelige nåverditap på lang sikt. Mangel på nøyaktige data om ungskogarealene kan være en grunn for at ungskogpleie ofte blir neglisjert, og videreutvikling av metoder for taksering av ungskog kan derfor potensielt bidra til å øke innsats for skogkultur.

I dagens laserbaserte skogtakser blir ulike karakteristikker av ungskogbestand predikert ved hjelp av fjernmålingsdata på lignende måte som for hogstklassene (hkl) 3 – 5. Behov for ungskogpleie blir da utledet enten ut fra estimert tetthet, eller ved bruk av en kalibrert tetthetsindeks (Ørka et al., 2015). Disse metodene har imidlertid ikke produsert tilfredsstillende resultater. Hovedutfordringen er at feilen knyttet til predikert tetthet innen ungskogbestand er for stor ved bruk av kun fjernmålingsdata (Næsset et al., 2001; Ørka et al., 2016). Estimatene på bestandsnivå er gjennomsnitt av aggregerte gridceller innen et bestand og representerer ikke nødvendigvis et reelt behov for ungskogpleie, som i praksis vil være den mest viktige informasjonen med hensyn til fremtidig forvaltning av bestandet. Ut fra dette perspektivet hadde det vært rimelig å vurdere en metode der behov for ungskogpleie blir predikert direkte ut fra laserdata. For tynning av eldre skog har dette gitt lovende resultater (Vastaranta et al., 2011; Pippuri et al., 2012; Watt et al., 2013), men Korhonen et al. (2013) konkluderte med at klassifisering av behov for ungskogpleie direkte ut fra laserdata enten ved bruk av logistisk regresjon eller en "support vector machine" (SVM) ikke kan erstatte feltbefaringer helt. For å skaffe pålitelige data om tetthet og behov for tiltak må ungskogbestand i dag fremdeles kontrolleres i felt, noe som medfører betydelige feltarbeidskostnader. En økning i nøyaktighet av prediksjoner av egenskaper til ungskog og behov for ungskogpleie kan derfor potensielt redusere takstkostnadene vesentlig.

Til dags dato har det kun blitt gjort et begrenset antall forsøk på å bruke FLS data for prediksjon av egenskaper til ungskog og behov for ungskogpleie. Næsset et al. (2001) undersøkte hvor

godt laserdata er egnet for prediksjon av høyde og tetthet innen ungskogbestand med høyde opp til 6 meter ved bruk av arealmetoden. Resultatene var lovende med hensyn til trehøyde ($R^2$ = 0.83) men modellen for totalt treantall ga dårligere tilpasning ($R^2$ = 0.42). Närhi et al. (2008) oppnådde tilsvarende resultater for grandominerte bestand med høyde opp til 8 meter. I tillegg til høyde og tetthet innen bestand, vurderte de behov for tiltak ut fra feltmålte data om trehøyde og tetthet. Behov for ungskogpleie ble predikert på to måter: 1) ved bruk av FLS data og 2) basert på predikerte verdier for dominerende høyde og totalt treantall. Metodene ga en samlet nøyaktighet av henholdsvis 72 og 69 %. Korhonen et al. (2013) kombinerte FLS data med tekstur og spektrale karakteristikker fra flybilder for å estimere behov for ungskogpleie ved hjelp av logistisk regresjon og oppnådde 71 % nøyaktighet på bestandsnivå og en kappa-indeks på 0.38. Etter prosjektet 'Laser2' konkluderte Ørka et al. (2015) at dominerende og total gjennomsnittshøyde, samt regulert treantall, kan estimeres med tilfredsstillende nøyaktighet på bestandsnivå. Feilen knyttet til estimert totalt treantall var derimot vesentlig, nemlig 41- 46 % på bestandsnivå. I tillegg presenterte prosjektet en tetthetsindeks og direkte klassifisering av reguleringsbehov som en alternativ løsning for å kartlegge behov for ungskogpleie som ga en samlet nøyaktighet på prøveflatene mellom 55 og 100 %.

Denne rapporten presenterer en praktisk tilnærming for taksering av ungskog og kartlegging av behov for ungskogpleie. Ved å bygge på etablerte modelleringsteknikker som gjennom tidligere forskning har blitt påvist til å være robuste, og ved å teste ulike nye prediktorvariabler blir egenskaper av ungskog og behov for ungskogpleie predikert, og videre blir prediksjonene validert.

### 1.3 Problemstilling

Hovedmålet med denne oppgaven var å videreutvikle, sammenligne og validere ulike metoder for laserbasert taksering av ungskog, og demonstrere en praktisk tilnærming for kartlegging av behov for ungskogpleie.

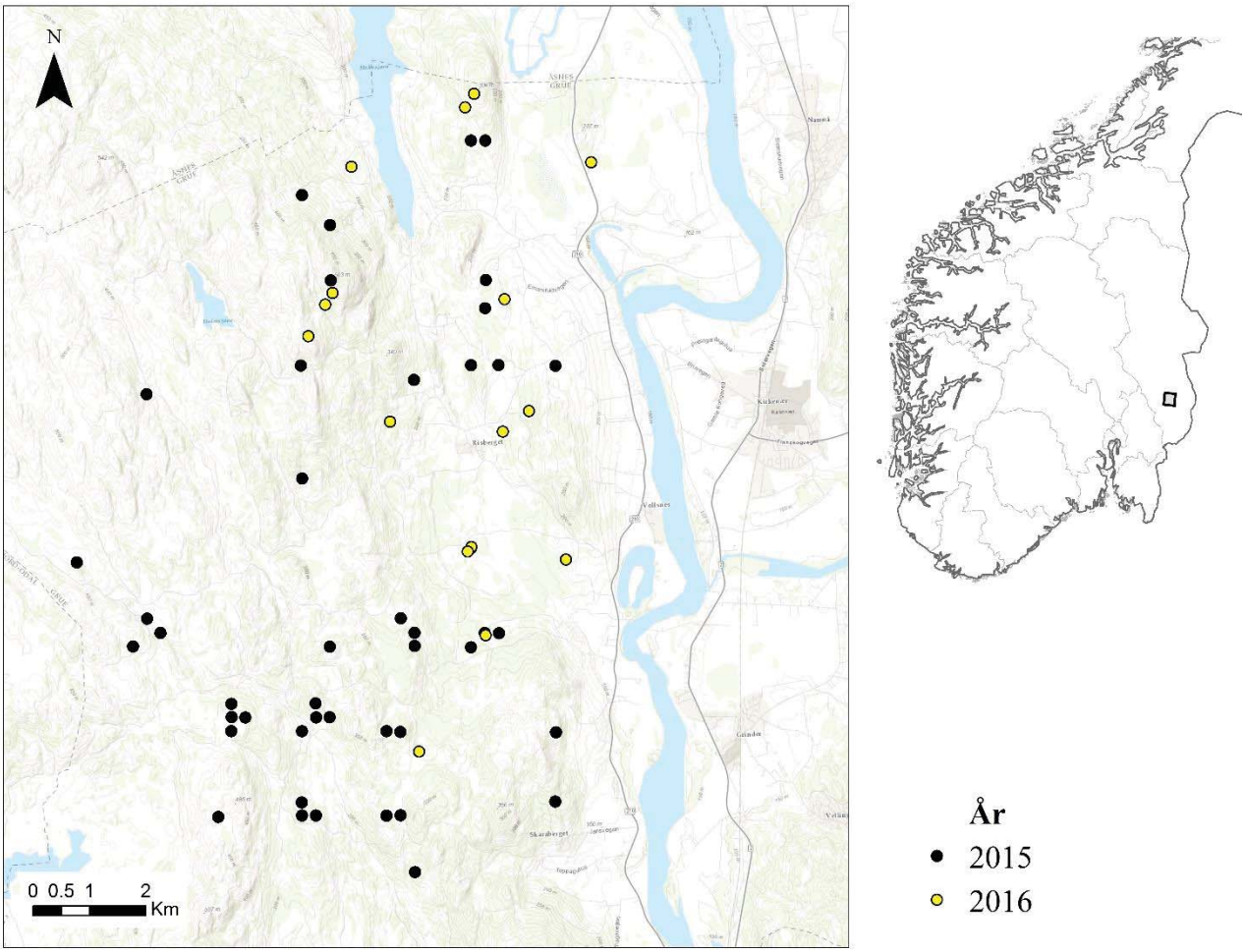
Følgende delmål ble fastsatt for å nå hovedmålet:

1. Utvikle og teste variabler som potensielt kan forbedre modeller for prediksjon av romlige biofysiske egenskaper av ungskog.
2. Utvikle modeller for prediksjon av treantall og høyde.
3. Evaluere prediksjonsevnen av modellene.
4. Teste metoder for prediksjon og kartlegging av behov for ungskogpleie.
5. Formulere anbefalinger for laserbasert taksering av ungskog.

# 2 MATERIALE OG METODE

## 2.1 Prosjektområdet

To ulike datasett ble benyttet i denne studien. Det første datasettet, heretter referert til som 2015-datasettet, ble innsamlet i forbindelse med en områdetakst utført av Mjøsen Skog SA i 2015 i Grue Vest (60°27′N, 11°55′Ø), sør i Hedmark, med formål å kartlegge skogressursene i området. Det ble lagt ut prøveflater i alle hogstklasser gjennom prosjektområdet, og totalt 63 prøveflater ble lagt ut i ungskogbestand. I tillegg ble det målt 64 prøveflater i ungskogbestand av gran i samme området i forbindelse med en kontrolltakst i 2016 (2016-datasettet). Figur 1 viser et oversiktskart over det aktuelle området der beliggenhet av samplingsenhetene er angitt som svarte (2015) og gule (2016) punkter. Beregnet areal i prosjektområdet var på 86 358 dekar produktiv skogsmark.




Figur 1. Oversiktskart over prosjektområdet.

## 2.2 Laserdata

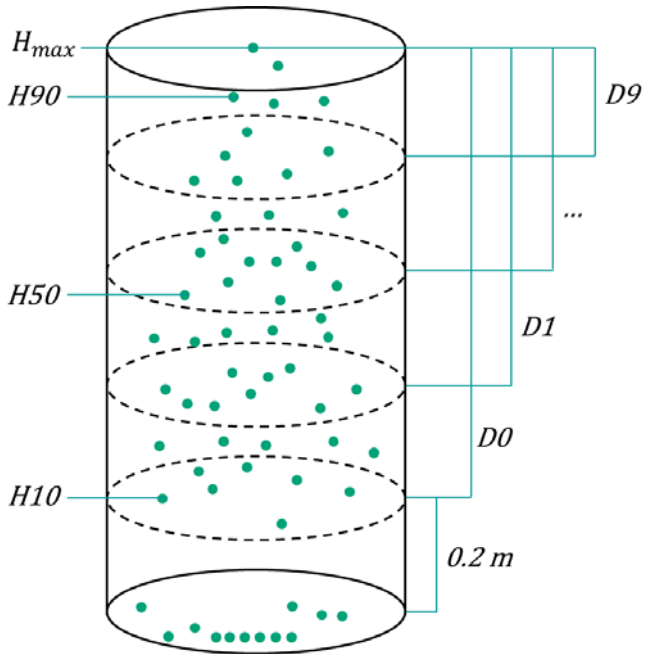
Laserskanningen av takstområdet ble gjennomført av Terratec AS den 8. juni 2014 i forbindelse med prosjektet Geovekst- et samarbeidsprosjekt mellom bl.a. Statens vegvesen, Kartverket, Energi Norge, Telenor, Landbruksdepartementet, og Norges vassdrags- og energidirektoratet. Målet for samarbeidet var å gjennomføre felles kartleggingsprosjekter og etablere et felles sett av geografiske data som tilfredsstiller et bredt brukerbehov. Under laserskanningen av prosjektområdet ble en Leica ALS70 flybåren laserskanner og en MicroIRS IMU ble benyttet. Gjennomsnittlig flyhastighet var 150 kn (278 km/t), repetisjonsrate var 177800 Hz, maksimal skannevinkel var 20 grader, og flyhøyden var ca. 2000 meter, som resulterte i en gjennomsnittlig punkttetthet på ca. 2.5 punkter/ $m^2$. Målemetoden CPOS ble brukt til posisjonering. Laserpunktene ble klassifisert automatisk, og editert til ikke-bakke (klasse = 1) og bakke (klasse = 2).

Laserdataene ble normalisert med LP360 i ArcMap, der planimetriske koordinater og ellipsoidiske høyder av laserpunktene ble beregnet. En "triangulated irregular network" (TIN) modell ble generert av siste returer innen punktskyen ved hjelp av en automatisk algoritme. Terrengnivået ble subtrahert fra den ellipsoidiske høyden til hvert laserpunkt for å beregne en høyde over bakken (HOB) eller 'z verdi' for hvert ekko.

Høyde- og tetthetsrelaterte egenskaper ble ekstrahert fra laserpunktskyen innen hver prøveflate. Egenskapene ble beregnet kun for førstereturer, siden (Ørka et al., 2015) fant at bruk av kun førstereturer ikke førte til signifikant dårligere tilpasning av modeller for prediksjon av egenskaper til ungskog sammenlignet med bruk av alle returer, eller første og siste returer. Det ble antatt at laserekko med HOB > 40 m ikke representerte vegetasjonstreff og punkter med en høyde over denne terskelen ble derfor utelukket fra analysen.

En høydeterskel på 20 cm ble benyttet for å redusere effekten av feilklassifiserte bakketreff på nøyaktigheten av terrengmodellen og for å fjerne effekten av ekko som ikke var relevante for prediksjon av egenskaper av ungskog (Ørka et al., 2015). Høydefordelingen innen punktskyen ble beregnet ut fra hvert punkt sitt HOB, og persentiler av høydefordelingen til laserpunktene ble så beregnet (H10, H20, …, H90, figur 2) der for eksempel H50 er den høyden der 50 % av pulsene traff ved en lavere høyde. Den romlige fordelingen av punkttetthet innen den vertikale dimensjonen i punktskyen av hver prøveflate ble så beregnet ved å dividere høyden mellom

laveste vegetasjonstreff (> 0.2 m) og maksimalt observert laserhøyde i ti identiske fraksjoner. Tetthetsvariablene ble deretter beregnet (D0, D1, …, D9, figur 2) som andel av førstereturer mellom den nedre høydeterskelen for hver fraksjon og maksimal laserhøyde. Eksempelvis ble tetthetsvariabelen *D0* beregnet som antall laserekko som traff med en høyde over 0.2 m, proporsjonal med det totale antallet ekko. Dessuten ble maksimal observert laserhøyde av første returer beregnet, samt gjennomsnittsverdi av første laserreturhøyder og standardavvik og variasjonskoeffisient til første laserpulshøyder.




Figur 2. Illustrasjon av de ulike høyde- og tetthetsvariablene som representerer den romlige strukturen av punktskyen innen en prøveflate (Bollandsås et al., 2008). Hvert punkt representerer et laserekko. Variablene til venstre viser høydevariablene, der $H_{max}$ er maksimum laserhøyder, og *H(x)* er høyden i vegetasjonen der *x* prosent av pulsene traff vegetasjonen ved en lavere høyde. Variablene til høyre illustrerer tetthetsvariablene, der *D(x)* er den relative andelen laserpulser som traff over den nedre terskelen for fraksjon *x*.

### 2.3 Feltdata

Feltarbeidet ble utført av Mjøsen Skog SA i 2015 i forbindelse med områdetaksten i Grue Vest med måling av 63 sirkulære prøveflater i hkl 2. 2015-datasettet besto av 45 prøveflater som var lagt ut i gran bestand, og 18 prøveflater i furu eller lauv bestand. I etterkant av områdetaksten ble det lagt ut 64 kontrollflater i ungskogbestand av gran i høsten 2016, der samme feltinstruks ble fulgt. 2016-datasettet ble oppnådd ved å dele alle hkl 2 granbestand inn i tre klasser for dominerende høyde, der prediksjoner for dominerende høyde fra den utførte taksten ble benyttet, og utvalget ble fordelt over de ulike høydeklassene for å oppnå en viss variasjon innen høydespekteret (Tabell 1).

Tabell 1. Klasser for dominerende høyde

| Høydeklasse | Variasjonsbredde (m) | Antall clustre |
|---|---|---|
| 1 | 2-3 | 5 |
| 2 | 3-4 | 5 |
| 3 | 4-5 | 6 |

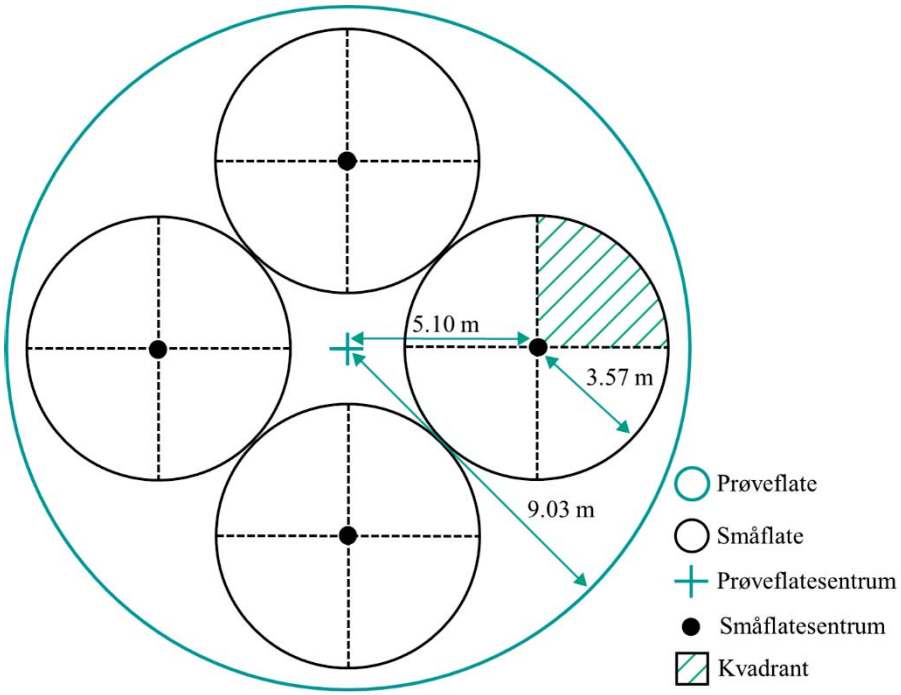
Siden prosjektområdet besto nesten kun av granskog, og det derfor var problematisk å skaffe tilstrekkelige data om furu- og lauvskog innen hver høydeklasse, ble denne studien fokusert på grandominert ungskog, og kun prøveflater som i 2015 var lagt ut i granbestand ble beholdt. Dette resulterte i totalt 45 prøveflater for 2015-datasettet. Målet med kontrolltaksten var å skaffe et uavhengig datasett for validering av modeller for prediksjon av egenskaper til ungskog, og derfor ble datasettene i denne studien ikke lagt sammen, og modellanalysene ble utført på hvert datasett for seg. En oversikt over hovedegenskaper av begge datasettene er gitt i Tabell 2.

Tabell 2. Oppsummering av datasettene. Nt = totalt treantall, Hd = dominerende høyde, Ht = gjennomsnittshøyde, Nd = regulert treantall.

| Egenskap | Variasjonsbredde | Gjennomsnitt | Standardavvik |
|---|---|---|---|
| *2015-datasettet (n = 45)* | | | |
| Andel gran (%) | 3 - 95 | 41 | 26 |
| Andel furu (%) | 0 - 58 | 8 | 13 |
| Andel lauv (%) | 3 - 96 | 51 | 27 |
| Hd | 0.5 - 7.8 | 2.7 | 1.9 |
| Nt | 438 - 27375 | 9588 | 6639 |
| Ht | 0.5 - 4.3 | 1.6 | 1.0 |
| Nd | 63 - 2000 | 1415 | 551 |
| | | | |
| *2016-datasettet (n = 64)* | | | |
| Andel gran (%) | 21 – 100 | 54 | 23 |
| Andel furu (%) | 0 – 61 | 5 | 13 |
| Andel lauv (%) | 0 -79 | 41 | 25 |
| Hd | 1.3 – 7.4 | 3.4 | 1.35 |
| Nt | 456 - 18126 | 5498 | 4253 |
| Ht | 1.0 - 6.7 | 2.8 | 1.3 |
| Nd | 456 - 2000 | 1724 | 361 |

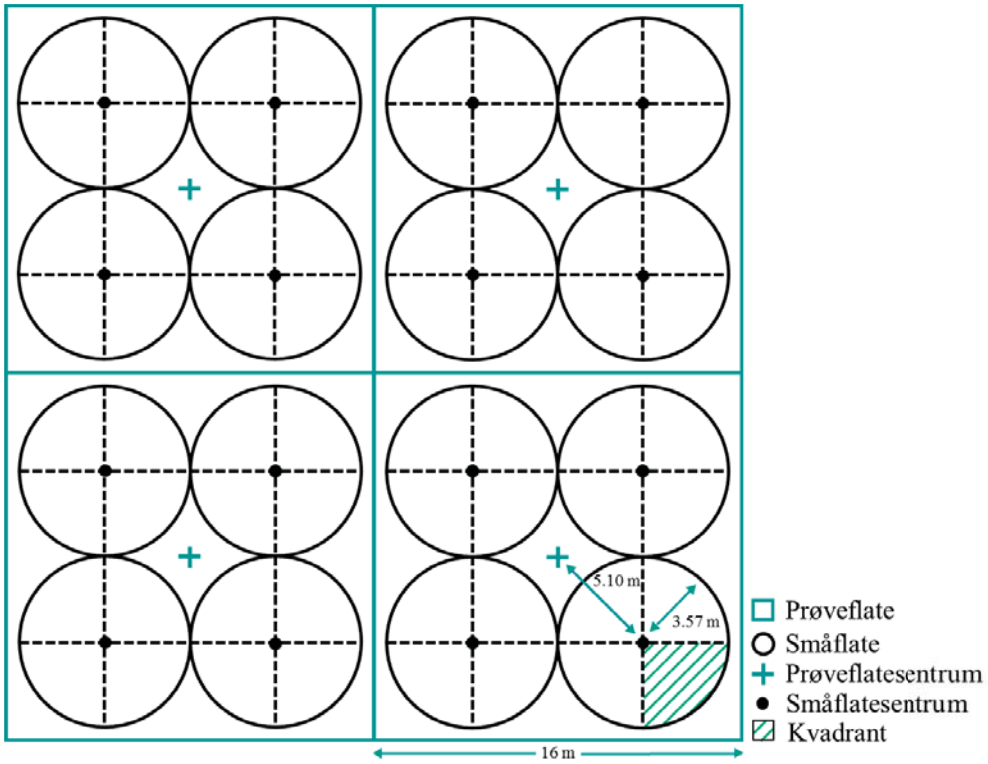
I prøveflatetaksten utført i 2015 ble det for hver prøveflate på 256 m² lagt ut et knippe (cluster) av fire småflater på 40 $m^2$ som representerte den større prøveflata med en radius på 9.03 m fra prøveflatesentrum, dvs. 256 $m^2$ (Figur 3). Hver av de fire småflatene i knippet hadde en radius

på 3,57 m (40 $m^2$), og senterpunktene ble lagt ut 5,1 meter i retningene N, S, Ø, V fra prøveflatesentrum. Summert areal av fire småflater utgjorde 63 % av det totale arealet av prøveflata. Hver småflate ble delt i fire kvadranter av 10 $m^2$ hver. Målebånd og kompass ble benyttet for å lokalisere småflatesentrum.




Figur 3. Illustrasjon av en prøveflate som er representert av fire småflater.

Kontrolltaksten i 2016 var i utgangspunktet delvis designet for å validere prediksjoner innen beregningsceller som var generert i et GIS i forbindelse med områdetaksten. Derfor ble småflatene lagt ut med formål å representere beregningscellene, og ble lagt ut i NØ, SØ, SV og NV retning fra midtpunktet i den samsvarende cellen slik at småflatene i sin helhet skulle befinne seg innenfor den kvadratiske avgrensingen. Dette utleggsmønster resulterte i firkantformede prøveflater siden sirkelformede prøveflater hadde medført en viss overlapp blant prøveflater. Det ble generert koordinater til småflatesentrum som ble lagt ut gjennom studieområdet etter prinsippet for clustersampling der et cluster besto av fire prøveflater, som hver var representert av fire småflater, d.v.s. totalt 16 småflater per cluster (Figur 4). Det ble bestemt hvilke beregningsceller skulle oppsøkes i felt ved tilfeldig utvalg med "Random Extract" funksjonen i QGIS for hver høydeklasse (Tabell 1). Kun celler med originalstørrelse ($256 \leq$ 'Areal' $\leq 259$) ble inkludert i samplet for å unngå overstandere og grenseceller, og koordinatene ble generert utenfor en buffersone av 10 meter fra hkl 2- bestandsgrensene.




Figur 4. Illustrasjon av et cluster av fire prøveflater, representert av totalt 16 småflater.

### 2.4 Feltregistrering

I forbindelse med områdetaksten ble det i 2015 lagt ut et systematisk nett av prøveflater i hogstklassene 2-5 gjennom hele takstområdet. Prøveflatene ble oppsøkt ved hjelp av håndholdt GPS, og selve prøveflatesentrene ble så lagt ut fem meter nord for den oppgitte posisjonen. Kompass og målebånd ble benyttet til å stedfeste sentrene. Hvis beliggenheten av en gitt prøveflate medførte registrering av overstandere, livsløpstrær eller annen eldre vegetasjon, ble senterpunktet flyttet videre i nordlig retning så langt som nødvendig for å unngå slik vegetasjon. Prøveflatene ble deretter innmålt med RTK (Real Time Kinematic), med minst to uavhengige målinger der koordinatene ble midlet. Ved et avvik større enn 20 cm ble posisjonen målt flere ganger, og gjennomsnitt av koordinatene ble registrert.

Ved kontrolltaksten i 2016 ble småflatene oppsøkt i felt direkte etter koordinater som var lagt inn i en GPS på forhånd. Først ble et håndholdt Garmin GPS brukt for å lokalisere clusteret, og til nøyaktig stedfesting ble en Topcon Hiper SR GPS benyttet, som stedfester posisjonen med en nøyaktighet på 10 mm. En manuell kontroll ble utført i ArcMap på forhånd ved hjelp av et ortofoto for å sikre at overstandere eller uforventete objekter skulle unngås.

I tillegg til registrering av skoglige biofysiske egenskaper, ble det i 2015 registrert behandlingsforslag for noen av prøveflatene (n = 35 beliggende i granbestand). Under kontrolltaksten i 2016 ble det tatt bilder av hver prøveflate i felt for å dokumentere skogtilstanden.

### 2.5 Registrering på småflater

Hver småflate ble delt i fire kvadranter på 10 $m^2$ hver i henhold til retningene N, S, Ø, V. En stang med lengde 3.57 m ble brukt for avgrensing av flatene. På hver kvadrant ble følgende egenskaper registrert:

Tabell 3. Registrerte egenskaper

| Egenskap | Beskrivelse |
|---|---|
| Totalt treantall- Gran | Det totale treantallet for gran (trær over 50 cm høyde) |
| Totalt treantall- Furu | Det totale treantallet for furu (trær over 50 cm høyde) |
| Totalt treantall- Lauv | Det totale treantallet for lauv (trær over 50 cm høyde) |
| Regulert treantall – Gran | Det regulerte treantallet for gran (trær over 50 cm høyde)<br>• Maksimalt to utviklingsdyktige trær totalt per kvadrant, siden 2000 trær/ha ble ansett å være fulltett<br>• Den minste avstanden mellom to regulerte trær er 1 m |
| Regulert treantall – Furu | Det regulerte treantallet for furu (trær over 50 cm høyde)<br>• Se gran for registreringsinstruksjoner |
| Regulert treantall – Lauv | Det regulerte treantallet for lauv (trær over 50 cm høyde)<br>• Se gran for registreringsinstruksjoner |
| Trehøyde totalt treantall | Høyde (dm) på det første treet som inngår i totalt treantall i hver kvadrant |
| Trehøyde regulert treantall | Høyde (dm) på det først treet som inngår i regulert treantall i hver kvadrant |

Ved høydemåling ble en høydestang brukt for trær < 4 m, og en Vertex høydemåler for trær > 4 m. Innen hver kvadrant ble rekkefølgen av de registrerte trærne bestemt ved å sample med klokka.

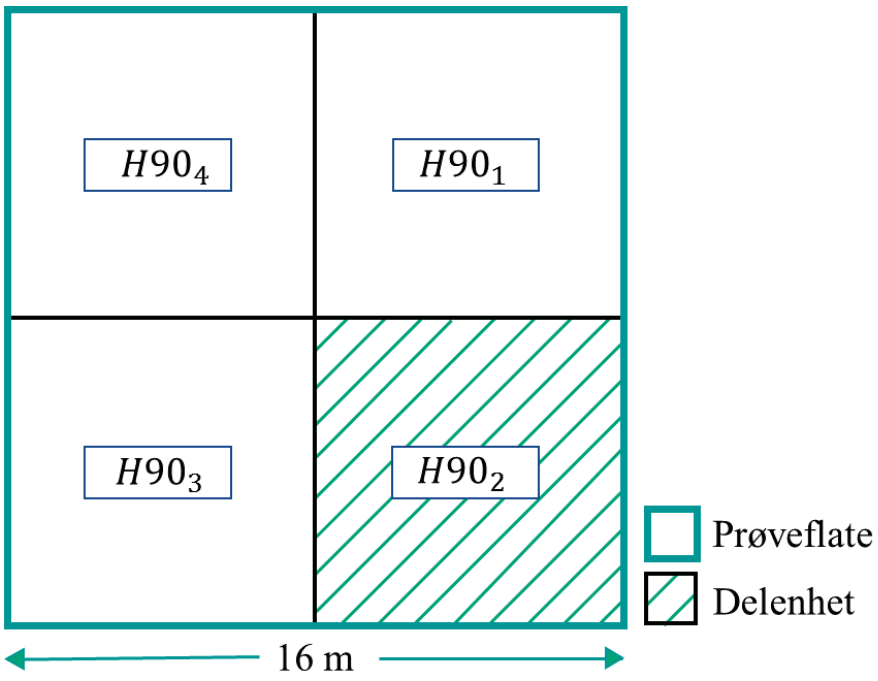
Siden laserskanningen ble utført i 2014 ble det trukket ett toppskudd for høydemålinger i taksten i 2015, og to toppskudd i kontrolltaksten i 2016 for å korrigere for høydetilveksten som trærne hadde oppnådd i mellomtiden. Trær som hadde en høyde > 5 desimeter under feltregistreringen men som ikke hadde nådd terskelverdien da skanningen ble utført i 2014 ble

dermed ikke registrert, og inngikk ikke i totalt treantall. For totalt treantall-lauv ble tette grupper av oppslag fra samme stubbe telt som ett individ.

### 2.6 Alternative forklaringsvariabler

Under dataprosesseringen ble prøveflatene fra begge datasettene delt inn i fire delenheter med identisk størrelse på 64 $m^2$, og høyde- og tetthetsvariabler ble beregnet for hver delenhet og i tillegg for hele prøveflata. Standardavvik for høyde- og tetthetsvariabler blant delenhetene i hver prøveflate ble beregnet med det formål å representere variasjonen av høyde og tetthetsvariabler innen prøveflata (*sd*-variabler). Eksempelvis ble variasjonen av laservariabelen *H90* innen en prøveflate beregnet ut fra standardavvik til H90-persentilene innen prøveflata:

*sdH90* = sd($H90_1$ + $H90_2$ + $H90_3$ + $H90_4$) (Figur 5).




Figur 5. Illustrasjon av beregning av laservariabelen sdH90.

Det ble undersøkt om *sd*-variablene ga en signifikant forbedring av prediksjonsevnen ved å sammenligne prediksjonsevnen av modeller der de alternative variablene var inkludert i modellseleksjonen med prediksjonsevnen av modeller der *sd*-variablene ikke inngikk i modellseleksjonen. Dette resulterte i to forskjellig modeller for hver av de fire skoglige variabler, som ble testet hver for seg. Kryssvalidering ble benyttet for å evaluere prediksjonsevnen av de alternative modellene, og differansen mellom residualene av modeller

der *sd*-variablene var inkludert og residualene av modeller der de alternative variablene ikke var inkludert ble beregnet. En tosidig t-test ($\alpha = 0.05$) ble benyttet for å teste om denne differansen var statistisk signifikant, der p-verdien indikerte sannsynligheten for at differansen mellom prediksjonsfeilene er forskjellig fra null.

I tillegg til laservariabler som representerer egenskaper av punktskyen innen prøveflata ble boniteten ekstrahert fra bestandskartet for hver prøveflate. Kartet fra den siste taksten, som ble utført i prosjektområdet i 2016, ble benyttet.

I det første steget blir regresjonsmodeller kalibrert med feltobserverte målinger som beskriver de biofysiske egenskapene til skogen som responsvariabler, og høyde- og tetthets laservariabler som blir ekstrahert fra laserpunktskyen som prediktorer. I det andre steget benyttes regresjonsmodellene til å predikere egenskapene over hele takstområdet (en beskrivelse av arealmetoden er gitt i Vedlegg 2).

### 2.7 Modellutvalg

Regresjonsmodeller ble kalibrert med feltobserverte målinger som beskriver de biofysiske egenskapene til skogen som responsvariabel, og laservariabler som prediktorvariabler. I regresjonsanalysen for 2015-datasettet ble multippel lineær regresjon benyttet. Den relative tilpasningen av alternative modeller ble evaluert etter "Akaike Information Criterion" (AIC) og "Bayesian Information Criterion" (BIC). Prediktorvariablene ble valgt ved hjelp av en automatisk stegvis algoritme- 'regsubsets()' funksjonen i R, der både "forward selection"- og "backward elimination" etter AIC og BIC ble anvendt. For å unngå overtilpasning ble det satt en øvre grense på tre forklaringsvariabler per modell. Modellen med lavest AIC og BIC ble foretrukket. Begge kriteriene gir et relativt estimat av informasjonen som går tapt når en gitt modell brukes til å representere dataene, favoriserer minimering av residualfeil og straffer modeller for flere prediktorvariabler med et straffeledd for antall parametere i modellen, som bidrar til å unngå overtilpasning (Burnham et al., 2004). Modellutvalg etter AIC og BIC resulterer dermed i en avveining mellom graden modellen som helhet er tilpasset kalibreringsdataene og kompleksiteten av modellen (Kuha, 2004).

Utgangsmodellen ble formulert som:

$$
\begin{aligned}
Y = {} & \beta_0 + \beta_1 bonitet + \beta_2 Hmax + \beta_3 Hmean + \beta_4 Hsd + \beta_5 Hcv + \beta_6 H10 + \beta_7 H20 \\
& + \beta_8 H30 + \beta_9 H40 + \beta_{10} H50 + \beta_{11} H60 + \beta_{12} H70 + \beta_{13} H80 + \beta_{14} H90 \\
& + \beta_{15} D0 + \beta_{16} D1 + \beta_{17} D2 + \beta_{18} D3 + \beta_{19} D4 + \beta_{20} D5 + \beta_{21} D6 + \beta_{22} D7 \\
& + \beta_{23} D8 + \beta_{24} D9 + \beta_{25} sdHmax + \beta_{26} sdHmean + \beta_{27} sdHsd \\
& + \beta_{28} sdHcv + \beta_{29} sdH10 + \beta_{30} sdH20 + \beta_{31} sdH30 + \beta_{32} sdH40 \\
& + \beta_{33} sdH50 + \beta_{34} sdH60 + \beta_{35} sdH70 + \beta_{36} sdH80 + \beta_{37} sdH90 \\
& + \beta_{38} sdD0 + \beta_{39} sdD1 + \beta_{40} sdD2 + \beta_{41} sdD3 + \beta_{42} sdD4 + \beta_{43} sdD5 \\
& + \beta_{44} sdD6 + \beta_{45} sdD7 + \beta_{46} sdD8 + \beta_{47} sdD9 + \varepsilon
\end{aligned}
$$

Der: $Y$ = regulert gjennomsnittshøyde, total gjennomsnittshøyde, regulert treantall eller totalt treantall, *bonitet* = H40-bonitet, *Hmax* = maksimalt observert laserhøyde av første returer, *Hmean* = gjennomsnittsverdi av første laserreturhøyder, *Hsd/Hcv* = standardavvik og variasjonskoeffisient til første laserpulshøyder, *H10 – H90* = høydepersentilene av første laserreturer, *D0 – D9* = lasertetthet i punktskyen tilsvarende andelen av første laserreturer ovenfor fraksjonen av det totale antallet første returer, *sdHmean – sdD0* = standardavvik til de ulike laservariablene av delenhetene innen prøveflata, $\varepsilon$ = feilleddet, antatt uavhengig og $N(0,\sigma^2)$.

Siden prøveflatene som ble målt i 2016 ble lagt ut etter prinsippet for cluster sampling, var observasjonene innen clustrene romlig autokorrelert, det vil si at samsvaret mellom observasjoner innen et cluster er relatert til avstanden mellom observasjonene. Clustrene resulterte også i en hierarkisk struktur på dataene. For å ta hensyn til disse aspektene ble en "mixed effects" (ME) modellanalyse benyttet. En ME modell inneholder både faste og tilfeldige effekter, som kan være nyttig i en situasjon hvor gjentatte målinger er gjort på de samme statistiske enhetene (Zuur et al., 2009). ME modeller beskriver forholdet mellom en responsvariabel og flere prediktorvariabler, med koeffisienter som kan variere med hensyn til en eller flere grupperingsvariabler. Faste effekter omfatter lineære regresjonsparametere som er av primær interesse, og tilfeldige effekter er uobserverte tilfeldige variabler som er assosiert med et sampel av grupperte enheter fra en større populasjon. ME modeller representerer dermed kovariansen relatert til grupperingen av dataene ved å knytte tilfeldige effekter til observasjoner som har det samme nivået av en grupperingsvariabel. ME modellering egner seg til cluster sampling, for å ta hensyn til romlige autokorrelasjoner, og i analyser der datasettet har en hierarkisk struktur (Zuur et al., 2009). *Clusternummer* ble brukt som tilfeldig effekt i

modelleringen, og funksjonen 'lme4' ble benyttet i R. Mauya et al. (2015) og Breidenbach et al. (2007) viste at denne modelleringsteknikken egner seg til beskriving av sammenhenger mellom egenskaper av grupperte samplingsenheter og laservariabler. Stegvis utvelgelse etter AIC og BIC ble benyttet for å bestemme hvilke prediktorvariabler skulle inngå i modellene. Modellene ble vurdert etter både AIC og BIC.

For modellene kalibrert med 2015-datasettet ble determinasjonskoeffisientene $R^2$ og $R^2_{adj}$ dokumentert. For ME-modellene kalibrert med 2016-datasettet ble både den marginale $R^2$ ($R^2_m$), som kun tar hensyn til de faste effektene ved å ekskludere den tilfeldige effekten innen regresjonen, og den betingete $R^2$ ($R^2_c$) som også tar hensyn til tilfeldige effekter innen modellen, dokumentert. $R^2$ er et statistisk mål som angir hvor stor andel av variansen i responsvariabelen er forklart av forklaringsvariablene:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

Der $y_i$ = observert verdi av observasjon *i*, $\hat{y}_i$ = predikert verdi av observasjon *i*, $\bar{y}$ = feltmålt gjennomsnittsverdi, og *n* = antall observasjoner i samplet.

$R^2_{adj}$ er en modifisert versjon av $R^2$ som er justert for antallet prediktorvariabler i modellen. $R^2_{adj}$ øker bare hvis en gitt prediktorvariabel som legges til forbedrer modellen i større grad enn det som ville være forventet ved tilfeldighet. Formelen for $R^2_{adj}$ er:

$$R^2_{adj} = 1 - \frac{(\hat{y}_i - \bar{y})^2/(n - (k+1))}{(y_i - \bar{y})^2/(n-1)}$$

Der *n* = antall observasjoner i samplet og *k* = antall forklaringsvariabler.

For å vurdere multikollinearitet ble den såkalte "variance inflation factor" (VIF) regnet ut for hver forklaringsvariabel som ble valgt til å inngå i kandidatmodellene. Ved en VIF > 10 for én eller flere forklaringsvariabler ble den respektive modellen tilpasset på nytt, der den forklaringsvariabelen med VIF > 10, med lavest signifikansnivå, ble fjernt fra modellseleksjonen. VIF for prediktorvariabel *j* ble beregnet etter formel:

$VIF_j = \frac{1}{1-R_j^2}$ for j = 1, …, k

Der $R_j^2$ er determinasjonskoeffisienten til alle forklaringsvariabelene *j* - *k* som blir testet.

### 2.8 Transformasjoner

Effekten av ulike transformasjoner på modelltilpasningen ble undersøkt ved å kalibrere modeller med følgende transformasjoner:

- Logaritmisk (log) transformasjon av responsvariablene
- Log-transformasjon av alle prediktorvariabler
- Log-transformasjon av respons- og prediktorvariablene
- Kvadratrottransformasjon (sqrt) av responsvariablene
- Sqrt-transformasjon av alle prediktorvariabler
- Sqrt-transformasjon av respons- og prediktorvariablene
- Ingen transformasjon

Modellene ble kalibrert for hver transformasjonstype, og det ble evaluert om de nødvendige modellforutsetninger var tilfredsstilt siden brudd på antagelsene kan resultere i ugyldige prosedyrer, resultater og konklusjoner. Det ble derfor undersøkt om residualplottene til de tilpassede modellene for hver transformasjonstype viste tegn på en ikke-lineær sammenheng, om variansen var konstant, om residualene var normalfordelt, om residualene var spredt likt langs variasjonsbredden av prediktorvariablene, og om det var noen observasjoner med stor innflytelse ("Cooks distance" > 1).

På grunnlag av ovennevnte analyse, ble en log-log transformasjon anvendt på variablene fra begge datasettene siden den i alle tilfeller tilfredsstilte de angitte antagelsene om feilleddet i høyeste grad. Et eksempel på en residualanalyse for modellene for dominerende høyde ved bruk av de ulike transformasjonene er gitt i Vedlegg 3.

Siden den logaritmiske transformasjonen som ble anvendt på responsvariablene førte til en negativ skjevhet ved prediksjon (Wiant, 1979), ble alle tilbaketransformerte prediksjoner korrigert for skjevhet ved hjelp av en "ratio of means" estimator foreslått av Snowdon (1992). Skjevheten ble beregnet ut fra forholdet mellom middeltallet av observerte verdier og

middeltallet av tilbaketransformerte prediksjoner. De endelige aritmetiske prediksjonene ble så multiplisert med forholdstallet for å korrigere for skjevhet.

### 2.9 Prediksjon av biofysiske egenskaper til ungskog

Prediksjonsevnen av modellene ble vurdert ved hjelp av både "leave one out" kryssvalidering (Arlot et al., 2010)**,** og validering ved bruk av de respektive uavhengige datasettene. Ved kryssvalidering ble én observasjon (prøveflate) tatt ut av datasettet om gangen og modellen ble kalibrert med de resterende observasjonene. En "out of sample" prediksjon ble beregnet for hver observasjon, og differansen mellom observert og predikert verdi ble beregnet for hver prediksjon. I tillegg ble prediksjonsevnen evaluert ved å benytte modellene som var kalibrert med 2015-datasettet til å predikere for prøveflatene fra 2016-datasettet, og omvendt. I tillegg ble prediksjonsevnen av modellene som var kalibrert med 2015-datasettet testet på clusternivå i 2016-datasettet for å vurdere effekten av aggregering av grupperte enheter. Gjennomsnittet av alle predikerte og observerte skoglige variabler ble beregnet innen hvert cluster, og differansen mellom observert og predikert verdi ble beregnet for hvert cluster. Prediksjoner for regulert treantall > 2000 trær/ha ble gjort om til 2000 siden feltobserverte verdier også var trunkert på 2000.

Ut fra differansen mellom observerte og predikerte verdier for hver skoglig variabel ble PRESS-statistikken beregnet som summen av kvadratene av prediksjonsfeil. PRESS-statistikken er et mål på prediksjonsevnen av en modell til et sampel av observasjoner som ikke selv ble brukt til kalibrering av modellen:

$$PRESS = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} d_i^2$$

Der $y_i$ = observert-, og $\hat{y}_i$ = predikert verdi for observasjon *i*, og *d* = differansen mellom observert og predikert verdi for observasjon i.

"Root mean squared error" (RMSE) og RMSE relativ til feltmålt gjennomsnittsverdi (RMSE%) ble brukt som endelige estimatorer for gjennomsnittlig prediksjonsfeil. En lavere verdi indikerer bedre prediksjonsevne av en modell:

$$RMSE = \sqrt{\frac{PRESS}{n}} \qquad \rightarrow \qquad RMSE\% = 100 * \frac{RMSE}{\bar{y}}$$

## 2.10 Prediksjon av behov for ungskogpleie

Det ble vurdert om det var behov for ungskogpleie på prøveflatene innen femårsperioden påfølgende etter laserskanningen. For 2015-datasettet ble det brukt subjektive behandlingsforslag registrert gjennom feltarbeidet, og for 2016-datasettet ble behov for ungskogpleie subjektivt tolket fra bildene som ble tatt i felt. Det ble registrert behandlingsforslag på kun en del av prøveflatene i 2015 (n = 36), og fra 2016-datasettet ble kun observasjoner der det på etterkant var synlig på bildene om det var behov for ungskogpleie eller ikke (n = 46) benyttet. Dette resulterte i et datasett med 82 feltobservasjoner, der 58 observasjoner ble klassifisert som 'behov' og 24 observasjoner ble klassifisert som 'ikke behov'.

### 2.10.1 Alternative metoder

To alternative metoder ble testet og evaluert for klassifikasjon av behov for ungskogpleie. I den første metoden ble behov for ungskogpleie predikert på grunnlag av totalt treantall, dominerende høyde og bonitet, og i den andre metoden ble behov for ungskogpleie predikert med et utvalg av laservariabler og bonitet.

### 2.10.2 Logistisk regresjon

Siden observasjonene ble delt inn i to klasser for behov for ungskogpleie, nemlig 'behov' og 'ikke behov', ble binær logistisk regresjon benyttet som klassifikator. Funksjonen for generaliserte lineære modeller (GLM) i R ble benyttet for å tilpasse modellene. Modellen ble formulert som:

$$\pi(x) = \frac{e^{\beta_0+\beta_1 x_1+\beta_2 x_2+\beta_3 x_3}}{1 + e^{\beta_0+\beta_1 x_1+\beta_2 x_2+\beta_3 x_3}}$$

Der $\pi(x)$ er sannsynlighet for behov for ungskogpleie, som antas å være en binær respons som resultat av et Bernoulli-forsøk (0/1, der 0 = 'ikke behov', 1 = 'behov'), $\beta_0$ = konstantledd, $x_{1\text{-}3}$ = kvantitative prediktorvariabler.

For den andre metoden ble ulike kandidatmodeller konstruert ved bruk av både stegvise algoritmer: 'stepAIC' og 'bestglm' funksjonene i R, og subjektiv utvelgelse. For begge metodene ble VIF-indeksen beregnet for de valgte prediktorvariablene, der det ble antatt multikollinearitet ved en VIF > 10. De endelige modellene ble sammensatt ved manuell eksperimentering og bestemt etter høyest samlet nøyaktighet og Cohen's Kappa-indeks (κ) (Landis et al., 1977) . For å unngå overtilpasning ble det satt en øvre grense på tre forklaringsvariabler for de logistiske modellene, siden minst ti observasjoner per forklaringsvariabel ble ansett som tilstrekkelig (Peduzzi et al., 1996).

Sannsynligheten som den logistiske regresjonsmodellen fremskriver som output er i utgangspunktet en kontinuerlig responsvariabel med en verdi mellom '0' og '1'. Denne verdien blir konvertert til enten '1' eller '0' ved å dele skala i to ved en viss terskel, som mest ofte er satt akkurat i midten på 0.5, der predikerte sannsynligheter med en verdi > 0.5 blir konvertert til '1' og prediksjoner < 0.5 til '0'. Ulike terskelverdier ble testet i denne studien for å undersøke om nøyaktigheten av klassifikasjonen kunne forbedres, og i tillegg ble det testet om en gitt terskel kunne føre til tilnærmet optimal nøyaktighet for en av to klassene, slik at enten 'behov' eller 'ikke behov' klassen kunne predikeres med nær optimalt resultat. Også ble det eksperimentert med et flertall klasser for sannsynlighet, der sannsynligheten som den logistiske modellen ga som output ble benyttet for å identifisere observasjoner som ga pålitelige klassifikasjoner.

### 2.10.3 Stratifisert "2-fold" kryssvalidering

Nøyaktigheten av klassifikasjonene ble testet ved bruk av stratifisert "2-fold" kryssvalidering (Kohavi, 1995), der samplet ble delt i to separate subsett av 41 observasjoner, der forholdet mellom '0' og '1' observasjoner innen subsettene var lik, for å sørge for at modellen ble kalibrert med et subsett som er representativ for datasettet som helhet. Observasjonene som ble valgt for begge klassene innen subsettene ble valgt tilfeldig for å unngå bias. Modellen ble kalibrert på halvparten av samplet og testet på den andre halvparten, og vice versa. For validering av den første modellen som inneholdte totalt treantall og dominerende høyde som prediktorer, ble modellen først kalibrert med *observerte* verdier for tetthet og høyde, og prediksjonene ble så utført ved bruk av *predikerte* verdier som input for modellen, siden observerte verdier for totalt treantall og dominerende høyde i praksis ikke er tilgjengelige for hele prosjektområdet. Feilmatriser ble generert for hver kandidatmodell og ut fra feilmatrisene ble samlet nøyaktighet, κ, og "producer's accuracy" og "user's accuracy" beregnet. Samlet nøyaktighet gjenspeiler hvor mange prosent av observasjonene er klassifisert riktig for begge klassene. "Producer's accuracy" er prosenten riktig klassifiserte observasjoner innen klassene observert i felt, og "user's accuracy" er prosenten riktige klassifikasjoner innen klassene predikert av modellen. κ antas å være et mer robust mål for nøyaktighet enn samlet nøyaktighet, ettersom κ tar hensyn til muligheten for at en riktig klassifikasjon oppstår tilfeldig. κ ble beregnet ut fra feilmatrisene:

$$\kappa = 1 - \frac{1 - \rho_0}{1 - \rho_e}$$

Der $\rho_0$ er den relative observerte overenstemmelse blant feltobserverte og modellpredikerte klassifikasjoner, og $\rho_e$ er den hypotetiske sannsynligheten for at klassifikasjonen ble bestemt

tilfeldig. Hvis utfallet blant modellpredikert og feltobservert klassifikasjon stemmer 100 % overens er $\kappa = 1$. $\kappa \leq 0$ hvis utfallet ikke er forskjellig fra det som kan forventes når observasjonene får tildelt en klasse helt tilfeldig (Cohen, 1960).

# 3 RESULTATER

## 3.1 Modeller for prediksjon av egenskaper til ungskog

### 3.1.1 Lineære regresjonsmodeller

En oversikt over de tilpassede log-log modellene for dominerende høyde (*Hd*), totalt treantall (*Nt*), gjennomsnittlig høyde (*Ht*) og regulert treantall (*Nd*) er gitt i Tabell 4. Følgende lineære regresjonsmodeller ble kalibrert med 2015-datasettet (n = 45):

Tabell 4. Parameterestimat med korresponderende standardavvik og signifikansnivå for de valgte prediktorvariablene, og determinasjonskoeffisientene ($R^2$ og $R^2_{adj}$) for de fire tilpassede regresjonsmodellene.

| **Laservariabler** | **Responsvariabler** | | | |
|---|---|---|---|---|
| | *log(Hd)* | *log(Nt)* | *log(Ht)* | *log(Nd)* |
| *Konstantledd* | 0.87 ± 0.25 ** | 8.9459 ± 0.37 *** | 1.07 ± 0.27 *** | 8.34 ± 0.22 *** |
| *logD0* | 0.48 ± 0.10 *** | | | 0.55 ± 0.13 *** |
| *logH50* | 0.41 ± 0.08 *** | | | 0.84 ± 0.26 ** |
| *logsdD7* | -0.16 ± 0.08 * | | | |
| *logH90* | | 0.3101 ± 0.16 . | | |
| *logsdH30* | | -0.3398 ± 0.12 ** | | |
| *logD2* | | 0.4875 ± 0.12 *** | 0.45 ± 0.05 *** | |
| *logsdD5* | | | -0.29 ± 0.08 *** | |
| *logsdD8* | | | 0.20 ± 0.09 * | |
| *logH70* | | | | -1.06 ± 0.23 *** |
| $R^2$ | 0.87 | 0.36 | 0.77 | 0.59 |
| $R^2_{adj}$ | 0.86 | 0.32 | 0.75 | 0.56 |

Signif. koder: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.'

Modellene for *Hd*, *Nt*, *Ht* og *Nd* forklarte henholdsvis 87, 36, 77 og 59 prosent av variasjonen i responsvariablene. *Bonitet* ble ikke valgt som prediktorvariabel for noen av modellene. For *Ht* ble det valgt en tetthetsvariabel og to laservariabler som representerer den romlige variasjonen av laserpunkter innen den horisontale dimensjonen i punktskyen (*sd*-variabler).

### 3.1.2 Mixed effects modeller

En oversikt over de tilpassede ME modellene for de fire skogvariablene, kalibrert med 2016-datasettet (n = 64), er gitt i Tabell 5.

Tabell 5. Parameterestimat med korresponderende standardavvik og signifikansnivå for de valgte prediktorvariabler, og determinasjonskoeffisientene ($R^2_m$ og $R^2_c$) for de fire tilpassede regresjonsmodellene.

| **Laservariabler** | **Responsvariabler** | | | |
|---|---|---|---|---|
| | ***log(Hd)*** | ***log(Nt)*** | ***log(Ht)*** | ***log(Nd)*** |
| *Konstantledd* | 1.30 ± 0.11 *** | 10.11 ± 0.51 *** | 0.81 ± 0.11 *** | 8.25 ± 0.22 *** |
| *logHmean* | 0.29 ± 0.08 *** | | | |
| *logH20* | 0.11 ± 0.06 . | | 0.21 ± 0.06 ** | -0.17 ± 0.05 ** |
| *logD0* | 0.30 ± 0.07 *** | | | |
| *logH40* | | -0.81 ± 0.20 *** | | |
| *logD3* | | 0.64 ± 0.20 ** | | |
| *logsdD7* | | 0.21 ± 0.11 . | | |
| *logH70* | | | 0.32 ± 0.08 *** | |
| *logD9* | | | | 0.30 ± 0.07 *** |
| *logsdH40* | | | | -0.06 ± 0.03 * |
| $R2_m$ | 0.74 | 0.27 | 0.61 | 0.35 |
| $R2_c$ | 0.86 | 0.80 | 0.89 | 0.64 |

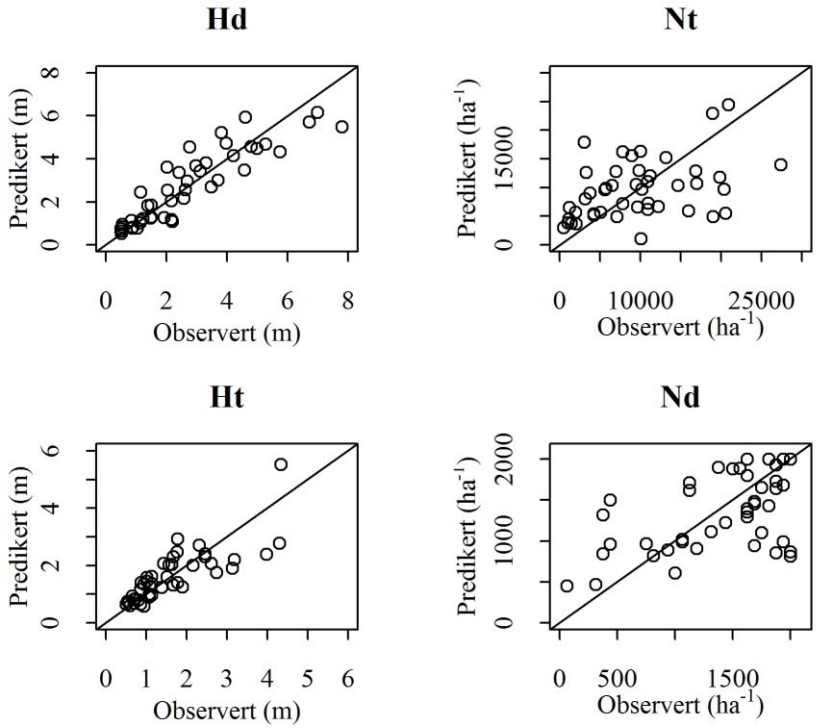
Signif. koder: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.'

Modellen for *Hd* inkluderte to høyde- og én tetthetsvariabel. For både *Nt* og *Nd* ble det valgt en høydevariabel, en tetthetsvariabel, og en *sd*-variabel. Prediktorvariablene *logH20* og *logsdD7* som ble valgt i modellene for henholdsvis *Hd* og *Nt*, hadde p-verdier i overkant av 0.05. For *Ht* ble det i modellseleksjonen kun valgt to prediktorer, som begge var høydevariabler. Den marginale $R^2$ av modellen for *Hd* var høyest.

## 3.2 Validering

### 3.2.1 Kryssvalidering av lineære regresjonsmodeller

I Figur 6 er predikerte verdier for dominerende høyde, totalt treantall, total gjennomsnittlig høyde og regulert treantall plottet mot de respektive feltobserverte verdiene.




Figur 6. Predikert vs. feltobservert høyde og tetthet for modellene som ble kalibrert med 2015-datasettet, oppnådd ved "leave one out" kryssvalidering.

Modellen for totalt treantall hadde, med 66.6 %, den høyeste gjennomsnittlige feilen relativ til feltmålt gjennomsnittverdi (Tabell 6). Modellen for dominerende høyde hadde lavest RMSE%. Gjennomsnittlig differanse mellom feltmålt og predikert regulert treantall var -48.07. For både observert og predikert regulert treantall er det en metning på 2000 trær/ha.

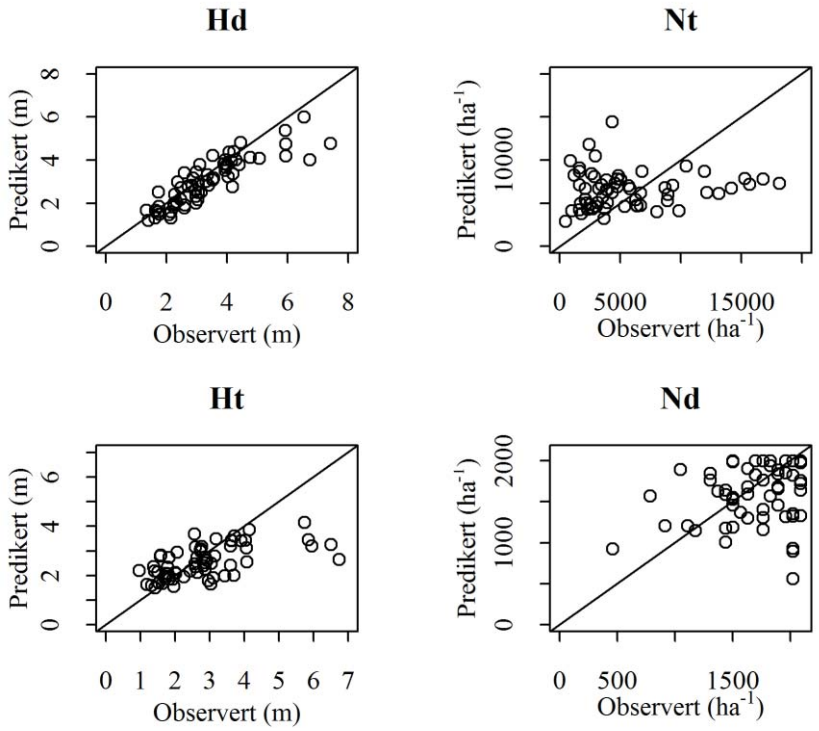
Tabell 6. Nøyaktigheten av modellene ved kryssvalidering for dominerende høyde (*Hd*), totalt treantall (*Nt*), total gjennomsnittshøyde (*Ht*) og regulert treantall (*Nd*).

| | | | | Differanse mellom predikert og feltmålt verdi | | | |
|---|---|---|---|---|---|---|---|
| Respons | RMSE | RMSE% | $\bar{y}$ | min | max | gjennomsnitt | standardavvik |
| *Hd* (m) | 0.81 | 29.8 | 2.72 | -2.29 | 1.79 | 0.00 | 0.82 |
| *Nt* (ha$^{-1}$) | 6390 | 66.6 | 9588 | -14943 | 14890 | 0.02 | 6462 |
| *Ht* (m) | 0.59 | 37.1 | 1.60 | -1.58 | 1.20 | 0.01 | 0.60 |
| *Nd* (ha$^{-1}$) | 486 | 34.4 | 1415 | -1186 | 1063 | -48.07 | 490 |

Der $\bar{y}$ = gjennomsnittlig feltmålt verdi

### 3.2.2 Validering av lineære regresjonsmodeller med et uavhengig datasett

Validering av de fire modellene på et uavhengig datasett ga, sammenlignet med kryssvalideringen, lavere gjennomsnittlige feil (RMSE) for dominerende høyde, totalt treantall og regulert treantall, men en høyere RMSE for total gjennomsnitthøyde (Tabell 7). RMSE relativ til feltmålt gjennomsnittverdi (RMSE%) var lavere for *Hd* og *Nd*. Gjennomsnittlig differanse mellom observert og predikert verdi var størst for totalt treantall (1021).




Figur 7: Predikert vs. feltobservert høyde og tetthet for modellene kalibrert med 2015-datasettet, validert med 2016-datasettet.

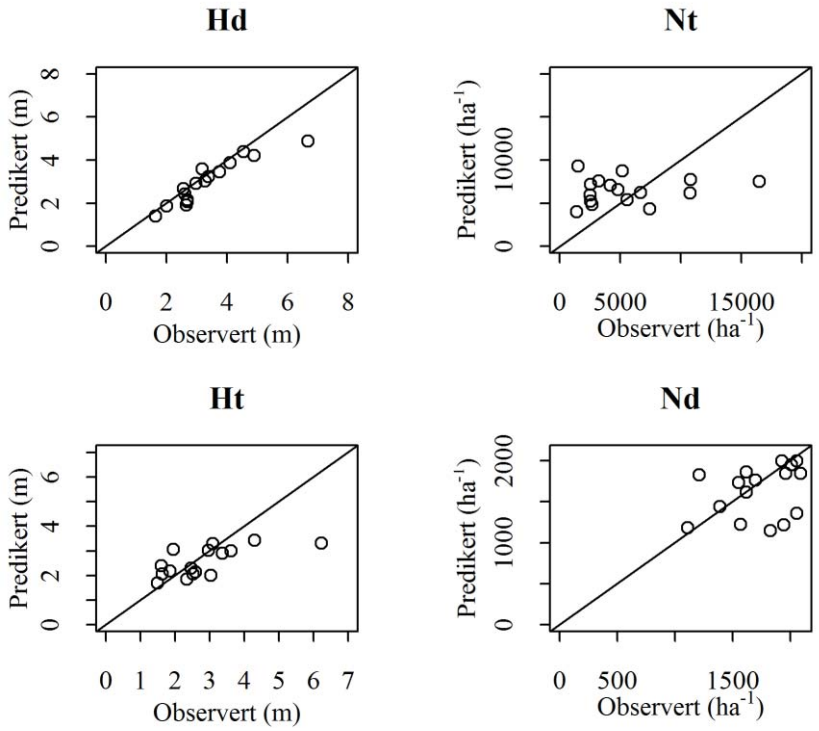
Tabell 7. Nøyaktigheten av modellene ved validering med 2016-datasettet for dominerende høyde (*Hd*), totalt treantall (*Nt*), total gjennomsnittshøyde (*Ht*) og regulert treantall (*Nd*).

| | | | | Differanse mellom predikert og feltmålt verdi | | | |
|---|---|---|---|---|---|---|---|
| Respons | RMSE | RMSE% | $\bar{y}$ | min | max | gjennomsnitt | standardavvik |
| *Hd* (m) | 0.74 | 27.2 | 2.72 | -2.71 | 0.83 | -0.33 | 0.67 |
| *Nt* (ha$^{-1}$) | 4606 | 48.0 | 9588 | -10818 | 10157 | 1021 | 4526 |
| *Ht* (m) | 1.05 | 66.1 | 1.60 | -4.08 | 1.25 | -0.26 | 1.03 |
| *Nd* (ha$^{-1}$) | 428 | 30.2 | 1415 | -1456 | 852 | -96 | 420 |

Der $\bar{y}$ = gjennomsnittlig feltmålt verdi

### 3.2.3 Validering av lineære regresjonsmodeller med aggregerte prøveflater

Prediksjonene ble aggregert til clusternivå (Figur 8). Ved aggregering av observasjoner innen clustre ble en lavere RMSE oppnådd for alle de fire responsvariablene. Det samme gjelder for RMSE%. De gjennomsnittlige differansene mellom observert og predikert verdi forble konstant for alle de fire skoglige variablene.




Figur 8: Predikert vs. feltobservert høyde og tetthet for modellene kalibrert med 2015-datasettet validert med aggregerte prøveflater fra 2016-datasettet.

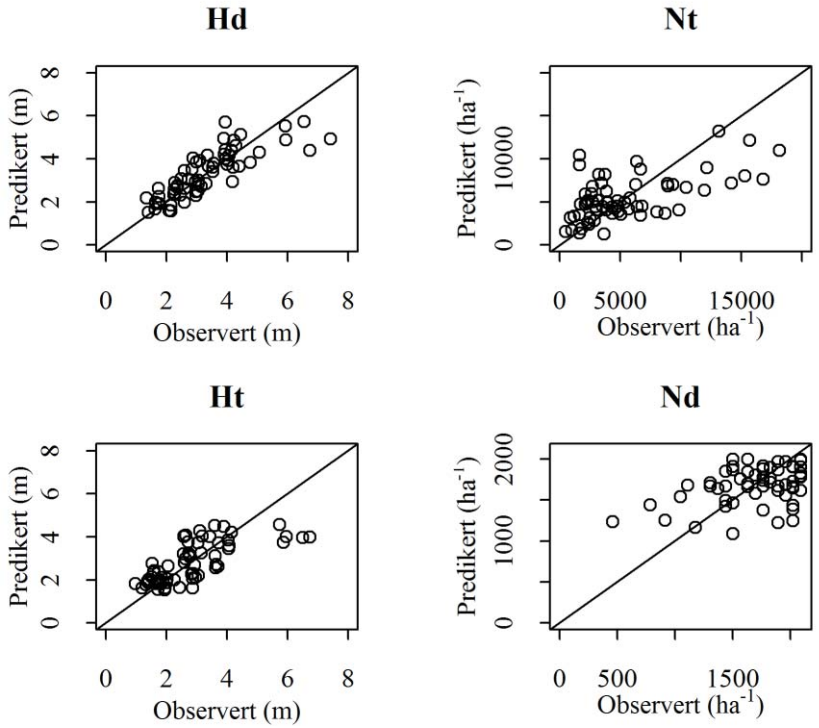
Tabell 8. Nøyaktigheten av modellene ved validering med 2016-datasettet for dominerende høyde (*Hd*), totalt treantall (*Nt*), total gjennomsnittshøyde (*Ht*) og regulert treantall (*Nd*).

| | | | | Differanse mellom predikert og feltmålt verdi | | | |
|---|---|---|---|---|---|---|---|
| Respons | RMSE | RMSE% | $\bar{y}$ | min | max | gjennomsnitt | standardavvik |
| *Hd* (m) | 0.57 | 21.0 | 2.72 | -1.77 | 0.41 | -0.33 | 0.48 |
| *Nt* (ha$^{-1}$) | 4154 | 43.3 | 9588 | -8893 | 7788 | 1021 | 4159 |
| *Ht* (m) | 0.93 | 58.2 | 1.60 | -2.91 | 1.13 | -0.26 | 0.92 |
| *Nd* (ha$^{-1}$) | 367 | 25.9 | 1415 | -719 | 622 | -96 | 365 |

Der $\bar{y}$ = gjennomsnittlig feltmålt verdi

### 3.2.4 Kryssvalidering av mixed effects modeller

Ved kryssvalidering av mixed effects modellene ble en forholdvis lav RMSE for *Nt* oppnådd (Tabell 9). Gjennomsnittlig differanse mellom observert og predikert *Nt* var -12.6. RMSE relativ til feltmålt gjennomsnittverdi (RMSE%) var med 62.1 % høyest for totalt treantall.




Figur 9: Predikert vs. feltobservert høyde og tetthet for mixed effects modellene kalibrert med 2016-datasettet, oppnådd ved kryssvalidering.

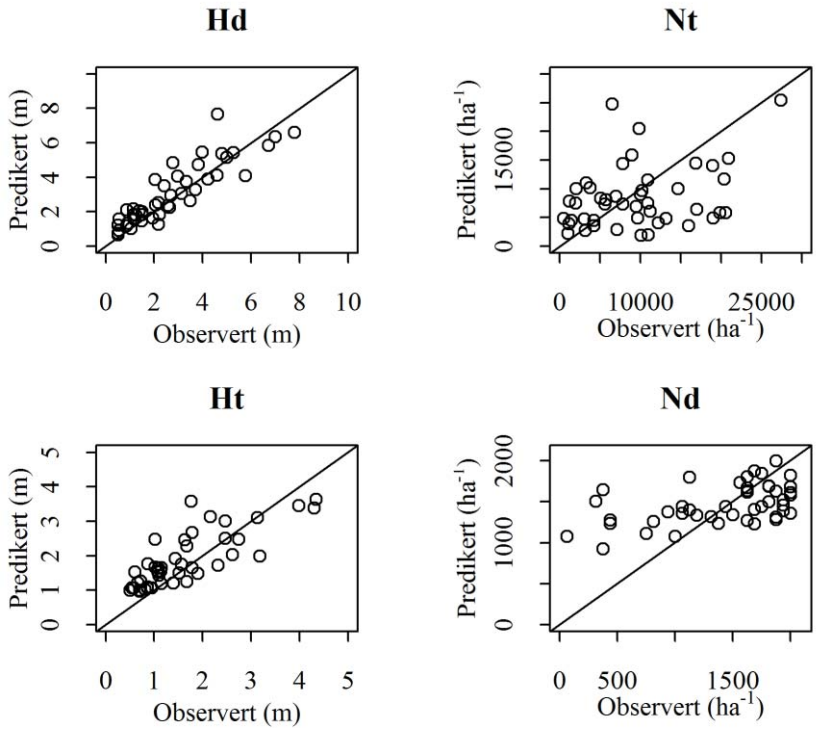
Tabell 9. Nøyaktigheten av modellene ved kryssvalidering for dominerende høyde (*Hd*), totalt treantall (*Nt*), total gjennomsnittshøyde (*Ht*) og regulert treantall (*Nd*).

| | | | | Differanse mellom predikert og feltmålt verdi | | | |
|---|---|---|---|---|---|---|---|
| Responsvariabel | RMSE | RMSE% | $\bar{y}$ | min | max | gjennomsnitt | standardavvik |
| *Hd* (m) | 0.74 | 22.0 | 3.35 | -2.49 | 1.78 | 0.00 | 0.74 |
| *Nt* (ha$^{-1}$) | 3412 | 62.1 | 5498 | -9100 | 8825 | -12.6 | 3439 |
| *Ht* (m) | 0.90 | 32.0 | 2.81 | -2.74 | 1.48 | 0.00 | 0.91 |
| *Nd* (ha$^{-1}$) | 327 | 19.0 | 1724 | -773 | 779 | -10.20 | 330 |

Der $\bar{y}$ = gjennomsnittlig feltmålt verdi

### 3.2.5 Validering av mixed effects modeller med et uavhengig datasett

Validering av de fire modellene på 2015-datasettet ga, sammenlignet med kryssvalideringen, en høyere RMSE for dominerende høyde, totalt treantall og regulert treantall, og en høyere RMSE% for tre av fire responsvariabler (Tabell 10). Gjennomsnittlig differanse mellom observert og predikert verdi var 1021 for totalt treantall og -752 for regulert treantall.




Figur 10: Predikert vs. feltobservert høyde og tetthet for modellene fra 2015-datasettet, validert med 2016-datasettet.

Tabell 10. Nøyaktigheten av modellene ved validering med 2016-datasettet for dominerende høyde (*Hd*), totalt treantall (*Nt*), total gjennomsnittshøyde (*Ht*) og regulert treantall (*Nd*).

| | | | | Differanse mellom predikert og feltmålt verdi | | | |
|---|---|---|---|---|---|---|---|
| Responsvariabel | RMSE | RMSE% | $\bar{y}$ | min | max | gjennomsnitt | standardavvik |
| *Hd* (m) | 0.92 | 27.5 | 3.35 | -1.65 | 3.06 | 0.42 | 0.87 |
| *Nt* (ha$^{-1}$) | 7059 | 128.4 | 5498 | -14642 | 18271 | -752 | 7057 |
| *Ht* (m) | 0.64 | 22.7 | 2.81 | -1.18 | 1.83 | 0.35 | 0.59 |
| *Nd* (ha$^{-1}$) | 474 | 27.5 | 1724 | -637 | 1273 | 136 | 476 |

Der $\bar{y}$ = gjennomsnittlig feltmålt verdi

### 3.3 Signifikans av *sd*-variablene

Modeller som ble tilpasset med *sd*-variabler som potensielle prediktorer i modellseleksjonen ga en lavere RMSE for dominerende høyde (*Hd*), total gjennomsnittshøyde (*Ht*) og regulert treantall (*Nd*), og høyere RMSE for totalt treantall (*Nt*) (Tabell 11). Differansene mellom residualene for de alternative modellene var ikke statistisk signifikante.

Tabell 11. Sammenligning av prediksjonsevne for modeller med- og uten *sd*-variabler.

| Respons | Modell | Prediktorer | $R2_{adj}$ | RMSE | RMSE% | D | Sign. |
|---|---|---|---|---|---|---|---|
| log(*Hd*) | 1 | *logD0+logH50+logsdD7* | 0.86 | 0.81 | 29.1 | 0.00 | IS |
| | 2 | *logD1+logH50+logHsd* | 0.85 | 0.84 | 31.1 | | |
| log(*Nt*) | 1 | *logH90+logsdH30+logD2* | 0.32 | 6390 | 66.7 | -12.69 | IS |
| | 2 | *logH80+logH60+logD0* | 0.31 | 5916 | 61.7 | | |
| log(*Ht*) | 1 | *logD2+logsdD5+logsdD8* | 0.75 | 0.59 | 37.1 | -0.01 | IS |
| | 2 | *logH20+logD1+logH10* | 0.73 | 0.59 | 37.2 | | |
| log(*Nd*) | 1 | *logD0+logH70+logH50* | 0.56 | 486 | 34.4 | -26.91 | IS |
| | 2 | *logH60+logD2+logH80* | 0.55 | 439 | 31.0 | | |

Model = alternative modeller (1,2) der *sd* variablene ble inkludert i modellseleksjonen (1), og ikke inkludert (2), D = gjennomsnittlig differanse av residualer, Sign. = statistisk signifikans *(α = 0.05)*, IS = ikke signifikant.

### 3.4 Prediksjon av behov for ungskogpleie

#### 3.4.1 Klassifisering ut fra treantall, høyde og bonitet

Den første logistiske modellen ble kalibrert med behov/ ikke behov observasjoner som respons, og feltobserverte verdier for totalt treantall og dominerende høyde i kombinasjon med bonitet som prediktorer. En oppsummering av modell 1 er gitt i Tabell 12.

Tabell 12. Parameterestimat med korresponderende standardavvik og signifikansnivå for de valgte prediktorvariablene til den første tilpassede logistiske regresjonsmodellen.

| | **Behov** |
|---|---|
| *Konstantledd* | -1440 ± 4.06 *** |
| *Nt* | 3.69e-04 ± 0.00 *** |
| *Hd* | 0.96 ± 0.28 *** |
| *Bonitet* | 0.65 ± 0.20 ** |

Signif. codes: 0 '***' 0.001 '**

### 3.4.2 Direkte klassifisering ut fra laservariabler og bonitet

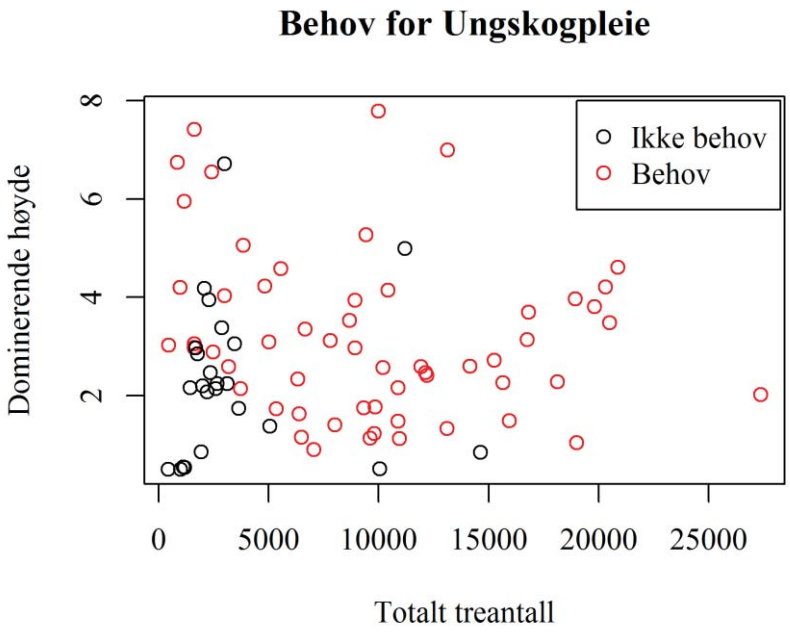
For den andre logistiske modellen ble det valgt tre prediktorvariabler. En oppsummering av modell 2 er gitt i Tabell 13.

Tabell 13. Parameterestimat med korresponderende standardavvik og signifikansnivå for de valgte prediktorvariablene til den andre tilpassede logistiske regresjonsmodellen.

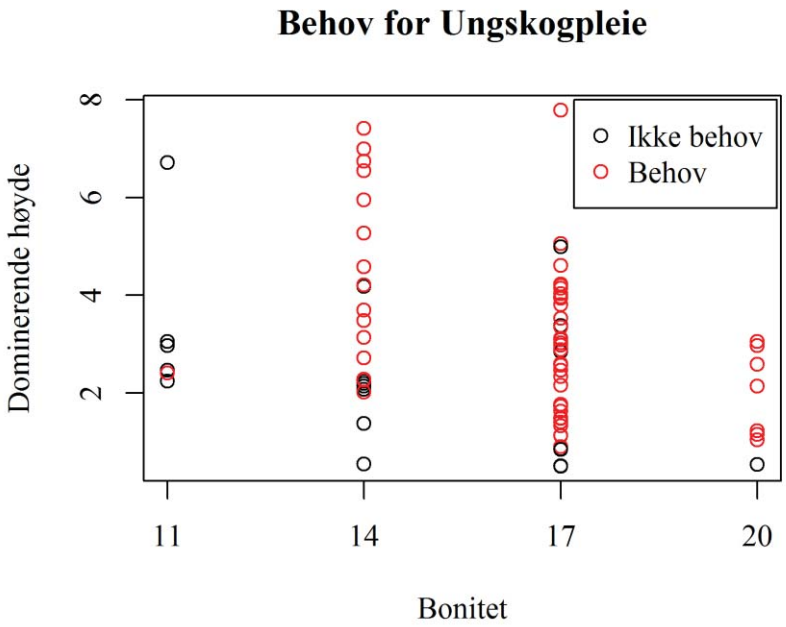
| | **Behov** |
|---|---|
| *Konstantledd* | -10.89 ± 3.23 *** |
| *D1* | 14.77 ± 3.97 *** |
| *H20* | -1.89 ± 0.78 * |
| *Bonitet* | 0.62 ± 0.19 *** |

Signif. codes: 0 '***' 0.01 '*'

I Figur 11 er observasjonene om behov for ungskogpleie (n=82) plottet mot dominerende høyde og totalt treantall, og Figur 12 viser behov for ungskogpleie som funksjon av dominerende høyde og bonitet.




Figur 11: Behov/ ikke behov for ungskogpleie for observasjonene som funksjon av dominerende høyde og totalt treantall, der svarte sirkler representerer observasjoner der det ikke ble registrert behov for ungskogpleie, og røde sirkler representerer observasjoner med behov for ungskogpleie innen 5 år.




Figur 12: Behov/ ikke behov for ungskogpleie for observasjonene som funksjon av dominerende høyde og bonitet, der svarte sirkler representerer observasjoner der det ikke ble registrert behov for ungskogpleie, og røde sirkler representerer observasjoner med behov for ungskogpleie innen 5 år.

### 3.4.3 Validering av modell 1

En terskel på 0.5 ble benyttet i klassifikasjonen for å bestemme om den kontinuerlige sannsynligheten som ble beregnet av modellen for hver observasjon ble konvertert til '1' eller '0', siden denne terskelen ga høyest samlet nøyaktighet og kappa-indeks. Tabell 14 viser nøyaktigheten av den første modellen ved "2-fold" kryssvalidering.

Tabell 14. Feilmatrise for kryssvalidering av modell 1

| Modell | Feltobservasjoner | | | |
|---|---|---|---|---|
| | Ikke behov | Behov | Total | User's accuracy |
| Ikke behov | 16 | 12 | 28 | 0.57 |
| Behov | 8 | 46 | 54 | 0.85 |
| Total | 24 | 58 | 82 | |
| Producer's accuracy: | 0.67 | 0.79 | | |
| Samlet nøyaktighet: | 0.76 | | | |
| Kappa: | 0.44 | | | |

Modellen ga en samlet nøyaktighet på 76 % med $\kappa = 0.44$. Av totalt 24 observasjoner som ble klassifisert som 'ikke behov' i felt ble 67 % klassifisert riktig av modellen, og av 58 observasjoner som ble klassifiser som 'behov' i felt ble 79 % klassifisert riktig av modellen. Totalt ble 28 observasjoner klassifisert som 'ikke behov' av modellen, hvorav 57 % var klassifisert riktig. Totalt 54 observasjoner ble klassifisert som 'behov' som ga en "user's accuracy" på 85 % for denne klassen. Totalt ble 62 observasjoner klassifisert riktig av modellen og 20 observasjoner ble feilklassifisert.

### 3.4.4 Validering av modell 2

For klassifikasjonen ved bruk av den andre modellen ble en terskel på 0.6 benyttet. Samsvar mellom feltobservert og modellpredikert behov for ungskogpleie ved "2-fold" kryssvalidering er gitt i Tabell 15.

Tabell 15. Feilmatrise for kryssvalidering av modell 2

| Modell | Feltobservasjoner | | | |
|---|---|---|---|---|
| | Ikke behov | Behov | Total | User's accuracy |
| Ikke behov | 18 | 8 | 26 | 0.69 |
| Behov | 6 | 50 | 56 | 0.89 |
| Total | 24 | 58 | 82 | |
| Producer's accuracy: | 0.75 | 0.86 | | |
| Samlet nøyaktighet: | 0.83 | | | |
| Kappa-indeks: | 0.60 | | | |

Ved bruk av den andre modellen økte den samlede nøyaktigheten til 0.83 % med $\kappa = 0.60$. For observasjonene som ble klassifisert som 'ikke behov' i felt ble 75 % klassifisert riktig av modellen. Av de 58 observasjonene som ble klassifiser som 'behov' i felt ble 86 % klassifisert riktig av modellen. Det ble oppnådd en "user's accuracy" på 69 % og 89 % for henholdsvis 'ikke behov' og 'behov' klassene. Til sammen ble 68 observasjoner klassifisert riktig av modellen og 14 observasjoner ble feilklassifisert.

# 4 DISKUSJON

Hovedmålet med denne oppgaven var å videreutvikle, sammenligne og validere ulike metoder for laserbasert taksering av ungskog, og demonstrere en praktisk tilnærming for kartlegging av behov for ungskogpleie. Delmålene var å (1) utvikle og teste nye prediktorvariabler som potensielt kan øke nøyaktigheten av prediksjoner av egenskaper til ungskog, (2) utvikle modeller for prediksjon av treantall og høyde, (3) evaluere prediksjonsevnen av modellene, (4) teste metoder for prediksjon av behov for ungskogpleie, og (5) formulere anbefalinger for taksering av ungskog. Resultatene oppnådd i denne studien har gitt ulike nye innsikter om laserbasert taksering av ungskog og prediksjon av behov for ungskogpleie. I dette kapittelet blir delmålene 1 - 4 diskutert, og de viktigste konklusjonene blir presentert for hvert delmål. De praktiske implikasjonene av studien diskuteres i kapittel 5, der ulike anbefalinger til aktører som jobber innenfor skogtaksering blir presentert, samt forslag til fremtidig forskning på laserbasert taksering av ungskog.

## 4.1 Prediktorvariabler

Det ble testet om ulike ukonvensjonelle prediktorvariabler kunne føre til en signifikant forbedring av prediksjonsevnen av modeller for prediksjon av høyde og tetthet av ungskog. Variablene som ble testet var *sd*-variablene, som representerer variasjonen av høyde- og tetthetsvariabler innen prøveflata, og bonitet. Innføring av disse variablene i modellseleksjonen resulterte i en lavere gjennomsnittlig prediksjonsfeil ved kryssvalidering av tre av de fire modellene som ble kalibrert med 2015-datasettet, men det kunne ikke konkluderes med at den gjennomsnittlige differansen av residualene blant de alternative modellene var statistisk signifikant $(\alpha = 0.05)$.

Bollandsås et al. (2008) viste at laservariabler som ble beregnet fra forskjellige romlige skalaer innen prøveflater kan bidra til å forbedre prediksjoner av egenskaper til naturlig foryngelse i flersjiktet granskog. Hypotesen om at en lignende tilnærming potensielt kan forbedre metoder for prediksjon av egenskaper til ung granskog virket derfor i utgangspunktet fornuftig. Siden *sd*-variablene representerer romlige egenskaper av punktskyen som ikke blir representert av høyde- og tetthetsvariabler, var forventningen at *sd*-variablene ville tilby tilleggsinformasjon som kunne forbedre prediksjonsevnen av modellene. Når både en høyde- og en tetthetsvariabel allerede er inkludert i en gitt modell, kan det antas at tilføring av en variabel av samme type ikke forbedrer modellen i like stor grad som en prediktor som representerer en annen romlig

egenskap av punktskyen. Valg av totalt seks *sd*-variabler av den stegvise algoritmen for modellseleksjon i totalt åtte modeller bekrefter dette perspektivet til en viss grad.

Videre ble det for modellene for dominerende høyde og totalt treantall som ble kalibrert med 2015-datasettet valgt en høyde-, en tetthets-, og en *sd*- variabel av den stegvise algoritmen. For ME modellene for regulert og totalt treantall ble det valgt variabler av de samme tre typene. De utvalgte forklaringsvariablene som ble valgt for de ovennevnte modellene, der modellutvalg etter AIC og BIC ble benyttet, viste at et variert sett av prediktorvariabler gir bedre tilpasning for modeller for prediksjon av romlige egenskaper av ungskog.

Det påpekes også at s*d*-variablene ga en lettere arbeidsflyt i modellseleksjonen, siden det ikke ble oppdaget noen sterke korrelasjoner mellom forklaringsvariabler (VIF > 10) når en *sd*-variabel ble valgt som prediktor i en gitt modell. I modeller der *sd*-variablene ikke ble inkludert ble det bestandig oppnådd høye VIF-verdier for noen av de valgte prediktorene, siden det regelmessig ble valgt to høyde- eller tetthetsvariabler for en modell som ga tilnærmet lik informasjon, som for eksempel *H20* og *H30*. Kombinasjonene resulterte i en høy VIF-indeks for begge variablene, og modellseleksjonen måtte derfor gjentas for å tilfredsstille krav om VIF < 10 for de individuelle prediktorene. Innføring av *sd*-variablene gjorde det derfor lettere å unngå multikollinearitet i modellseleksjonen.

For ingen av de kalibrerte modellene ble bonitet valgt som forklaringsvariabel, men det påpekes at det var forholdsvis lite variasjon i bonitetsklasser innen samplet. For 2015- og 2016-datasettetene hadde henholdsvis 60 % og 63 % av observasjonene bonitetsklasse 17 (Vedlegg 6). Det kan derfor ikke utelukkes at resultatene hadde vært annerledes med en større variasjonsbredde i bonitet. I tillegg bør det bemerkes at nøyaktigheten av bestandsdata om bonitet ikke har blitt testet i denne studien, og i operative takster bestemmes bonitet ved manuell og subjektiv tolkning av flybilder, som kan medføre feilaktig klassifisering. I tillegg kan det være betydelige variasjoner av bonitet innen bestand (Amateis et al., 2016). Planen var opprinnelig å teste om stratifisering etter bonitet kunne potensielt forbedre metoder for kartlegging av ungskog. Men siden det var en skjev fordeling av observasjoner blant bonitetsklassene, og antallet observasjoner med lav bonitet ikke var tilstrekkelig for å skaffe et representativt stratum, ble bonitet inkludert som kontinuerlig prediktorvariabel i modellseleksjonen. En effekt av denne innføringen kunne dermed ikke dokumenteres.

En betydelig andel av observasjonene ble lagt ut enten helt eller delvis i overlappsoner mellom flykorridorer, og det ble beregnet verdier for punkttetthet > 4 $m^{-1}$ for 44 av totalt 109 observasjoner. I analysen ble det ikke tatt høyde for en eventuell bias som oppstår ved bruk av prøveflater med høyere punkttetthet, siden det ble antatt at denne effekten var minimalt. I tillegg kan en slik bias også forklares av skannevinkelen, gitt at overlappsoner befinner seg langs kantene av flykorridoren der skannevinkelen er størst.

### 4.2 Modeller for prediksjon av treantall og høyde

Regresjonsmodeller ble kalibrert for prediksjon av regulert og totalt treantall, total gjennomsnittshøyde, og dominerende høyde av ung granskog. Effekten av ulike transformasjoner ble undersøkt i regresjonsanalysen, der modeller med log-log transformasjon ga bedre tilpasning og tilfredsstilte de angitte antagelsene om feilleddet i høyeste grad.

Determinasjonskoeffisientene som ble beregnet for de ulike modellene var sammenlignbare med resultater oppnådd i tidligere studier. Det ble oppnådd en $R^2$ på 0.87 for den lineære modellen for dominerende høyde, tilsvarende resultater ble presentert av Ørka et al. (2016) ($R^2$ på 0.74- 0.88) og Næsset et al. (2001) ($R^2$ på 0.83). For modellen for totalt treantall ble det oppnådd en $R^2$ på 0.37, som også var som forventet ut fra tidligere forskning der Ørka et al. (2016) presenterte en $R^2$ på 0.07 – 0.26, og Næsset et al. (2001) en $R^2$ på 0.42.

Selv om log-log modellene tilfredsstilte antagelsene om feilleddet i en høyere grad enn modeller med andre transformasjoner, ble det ikke konstatert konstant varians for alle tilpassede modeller. Residualene var ikke spredt helt jevnt og tilfeldig langs variasjonsbredden av den predikerte variabelen, som kan medføre bias for standard feil estimater av visse parametere, og resultere i ugyldige konklusjoner.

En automatisk algoritme ble benyttet for valg av forklaringsvariabler i modellseleksjonen. Det bemerkes at det er en viss uenighet om hvorvidt det er ønskelig å bruke automatiserte metoder for modellseleksjon (Flom et al., 2007). Gitt det høye antallet potensielle prediktorvariabler, og siden det ikke var realistisk å velge ut potensiell gode prediktorer ved manuell eksperimentering, ble en automatisk algoritme benyttet. Som et alternativ kan en automatisk algoritme også kombineres med faglige vurderinger knyttet til valg av prediktorer i den

endelige modellen. Automatiske metoder kan for eksempel brukes for å avgjøre om det er høye persentiler som gjennomgående blir valgt for eksempel – og om tetthetsvariable er signifikante.

Den relative tilpasningen av alternative modeller ble evaluert etter AIC- og BIC-statistikkene. Modellene med lavest verdi for AIC og BIC ble valgt som endelig modell, som i alle tilfeller også var modellen med høyest $R^2$. Men som tabell 11 viser, betyr en lavere AIC og BIC, eller en høyere $R^2$, ikke nødvendigvis at en gitt modell gir bedre prediksjoner. En prosedyre i modellseleksjonen der, i tillegg til en vurdering av tilpasningskriteriene, også en vurdering av prediksjonsfeil blir gjort som for eksempel baseres på kryssvalidering av hver kandidatmodell, kan potensielt resultere i modeller med bedre prediksjonsevne. I denne studien ble et slikt trinn i modellseleksjonen ikke benyttet siden det av praktiske årsaker ble bestemt at tilpasningskriteriene AIC og BIC skulle bli brukt. Men gitt at modellene spesifikt ble kalibrert med det formål å predikere egenskaper av ungskog, kunne en RMSE beregnet ut fra kryssvalidering også ha vært hensiktsmessig å benytte som kriterium.

Det ble forventet at modeller for prediksjon av trehøyde skulle omfatte minst en høydevariabel, siden det ble antatt at det var en sterk sammenheng mellom laservariabler som representerer høydefordelingen av ekko i punktskyen og den faktiske høyden av vegetasjonen. Mot forventningen ble det for modellen for gjennomsnittlig totalt trehøyde, kalibrert med 2015-datasettet, valgt to *sd*-variabler, og en tetthetsvariabel. Utvalget av prediktorer antyder at det var forholdsvis lite variasjon i feltmålte verdier for total gjennomsnittlig trehøyde innen 2015-datasettet, og at det var samvariasjon mellom mange av tetthet- og høydevariablene. Dette stemmer overens med karakteristikkene av 2015-datasettet, der gjennomsnittlig trehøyde var på 1.6 m, med et standardavvik på 1 m, og 76 % av observasjonene hadde en total gjennomsnittshøyde < 2 m.

P-verdiene for alle prediktorer i totalt åtte modeller viste at koeffisientene til de ulike laservariabler var signifikant forskjellige fra null *($\alpha = 0.1$).* De fleste prediktorer hadde et signifikansnivå på *$\alpha = 0.01$.* Siden den relative tilpasningen av modellene ble evaluert etter AIC og BIC, ble det ikke spesifisert noen krav på p-verdier av individuelle prediktorer. Som et resultat ble et flertall prediktorer inkludert med en p-verdi > 0.05. Dette fordi modellutvalg etter AIC og BIC er rettet mot å konstruere modeller som generelt er tilpasset dataene, i motsetning til utvalg etter p-verdier av de individuelle prediktorvariablene der man kan spesifisere et signifikansnivå for de individuelle prediktorene.

Koeffisientene i regresjonsmodellen for dominerende høyde som ble kalibrert med 2015-datasettet viste at det var en positiv sammenheng mellom prediktorvariablene *D0* og *H50* og dominerende trehøyde, gitt at de andre prediktorene i modellen forble konstant. Et logisk resultat, siden det var en sterk positiv korrelasjon mellom dominerende høyde og disse to variablene hver for seg. Men i den samme modellen hadde prediktoren *sdD7* en negativ innvirkning på responsvariabelen, som var overraskende siden denne variabelen faktisk var positiv korrelert med responsen, og VIF statistikkene for alle prediktorer i modellen var med < 5 forholdsvis lav og multikollinearitet ble derfor ikke konstatert. Fortegnet som er motsatt det som var forventet kan forklares av samvariasjon mellom denne variabelen og *D0*, der en korrelasjon på 0.79 mellom de respektive prediktorene var forholdsvis høy.

### 4.3 Prediksjonsevne av modellene

Prediksjonsevnen av de ulike modellene ble evaluert ved hjelp av kryssvalidering, alternerende ved bruk av de to forskjellige datasettene, og effekten av aggregering ble testet innen clustrene. Nøyaktigheten av prediksjonene var som forventet, der Næsset et al. (2001) og Ørka et al. (2015) også konkluderte med en forholdsvis lav gjennomsnittlig feil for dominerende og total gjennomsnittshøyde og regulert treantall, og større prediksjonsfeil for modeller for totalt treantall.

Den gjennomsnittlige prediksjonsfeilen for dominerende høyde var, med 0.57 – 0.92 m, forholdsvis lav. Resultatet er som forventet ut fra tidligere forskning, der Ørka et al. (2016) oppnådde en RMSE på 0.74 – 1.15 m. Fra et forvaltningsperspektiv er dette et viktig resultat siden tidspunktet for ungskogpleie som regel bestemmes etter dominerende høyde (Varmola et al., 2004). Denne studien bekrefter også at regulert treantall kan predikeres med forholdsvis lav gjennomsnittlig feil. Denne egenskapen er viktig både for strategisk skogbruksplanlegging, for å kunne beregne prognoser for et bestands fremtidig utvikling, og for taktisk planlegging, for å kunne vurdere om supplering er et ønsket tiltak på visse plantefelt. Total gjennomsnittlig trehøyde er muligens minst viktig fra et forvaltningssynspunkt, men modellene for denne variabelen ga også robuste prediksjoner.

Som forventet var nøyaktigheten av prediksjonene for totalt treantall forholdsvis lav. Forsøket på å forbedre prediksjonsevnen av modellene ved å introdusere nye prediktorer resulterte ikke i en statistisk signifikant reduksjon av den gjennomsnittlige prediksjonsfeilen. Men selv om nøyaktigheten av prediksjonene for totalt treantall ikke var som ønsket, bemerkes det at mulighetene for prediksjon av denne egenskapen innen ungskog per i dag er begrenset.

Alternativet, som er kostnadsintensive feltmålinger i hvert bestand ved bruk av prøveflater, fører ikke nødvendigvis til økt nøyaktighet (Ørka et al., 2015).

Aggregering av prøveflater innen clustre førte til en betydelig lavere RMSE for alle fire responsvariablene. Dette resultatet var som forventet, siden tilfeldige feil, det vil si feil som slår ut til begge sider av den 'sanne' verdien av en observasjon som kan oppstå på grunn av ulike ukontrollerte variabler, opphever hverandre når grupperte observasjoner er sammenslått ved å midle de observerte og predikerte verdiene av de individuelle samplingsenhetene.

Forventningen var at valideringsresultatene ville bli bedre enn funnene fra tidligere studier, siden denne studien var fokusert kun på granbestand. Det ble forutsatt at variasjonen mellom prøveflater som ble lagt ut kun i granbestand var mindre enn variasjonen mellom prøveflater som hadde blitt lagt ut i bestand av alle treslag, og at prediksjonsevnen av modeller for prediksjon av de fire skoglige variablene derfor skulle bli bedre (Næsset et al., 2001; Ørka et al., 2016), men dette var ikke tilfellet. Denne konstateringen antyder at stratifisering etter treslag, som ikke eksplisitt har blitt testet i denne studien siden dataene ikke var tilstrekkelige, ikke nødvendigvis skal gi en betydelig økning i nøyaktighet av prediksjoner innen ungskog.

En mulig forklaring for at valideringsresultatene oppnådd i denne studien ikke ga noen bedre resultater enn resultatene som ble oppnådd i tidligere studier er forskjellen mellom metoden for utlegging av prøveflater, som medførte ulikheter blant datasettene og en begrenset grad av variasjon innen 2016-datasettet. Prøveflatene målt i 2015 ble lagt ut systematisk, mens 2016-datasettet ble samplet i clustre fordelt over klasser i henhold til predikert dominerende høyde. Clustrene ble fordelt utover østsiden av prosjektområdet. Systematisk sampling er en samplingsmetode hvor samplingsenhetene velges ut etter et bestemt mønster, i dette tilfellet et rutenett spredt utover hele prosjektområdet, som fører til en mest mulig regelmessig fordeling av enhetene utover populasjonen, og at samplingsenhetene blir spredt utover arealet på best mulig måte. Siden clustrene ble fordelt over de spesifiserte høydeklassene, ble observasjonene ikke spredt regelmessig utover prosjektområdet, og 2016-datasettet var derfor ulik 2015-datasettet. Gjennomsnittlig dominerende høyde i 2015-datasettet var med 2.7 m betydelig lavere enn 3.4 m i 2016-datasettet (Tabell 2). Gjennomsnittlig tetthet, som antas å avta ved økende høyde, var med 9588 trær/ha i 2015-datasettet betydelig høyere enn 5498 trær/ha i 2016-datasettet.

Forskjellene i egenskaper mellom datasettene tyder på at 2016-datasettet ikke var representativ for ungskog i hele prosjektområdet. En utvalgsmetode som ikke bare baseres på predikert

dominerende høyde, men også tetthet som kriterium kunne ha blitt bruk for utlegging av prøveflater. Men siden det på grunn av tidligere forskning ble antatt at feilen knyttet til prediksjonene for totalt treantall var betydelig, og at det er en viss sammenheng mellom dominerende høyde og totalt treantall, ble utvalget basert kun på de spesifiserte klassene for dominerende høyde. Som et alternativ kunne en utvalgsmetode ha blitt brukt der samplet ble valgt med det formål å dekke variasjonen innen høyde- og tetthetsvariabler beregnet fra punktskyen, siden laserdataene var tilgjengelige før igangsetting av feltarbeidet.

I tillegg til ulikheter innen egenskaper av treantall og høyde, var 2015-datasettet mer variert enn 2016-datasettet (Tabell 2, Vedlegg 4), siden observasjonene i 2016-datasettet var clustret sammen i grupper av fire observasjoner. Ved å legge ut prøveflater i clustre ble det gjennom feltarbeidet i 2016 ikke oppnådd like mye variasjon blant prøveflatene, siden grupperte samplingsenheter som regel ikke varierer like mye som individuelle observasjoner som ligger spredt utover et større areal.

Videre ble det gjennom feltarbeidet i 2016 kun samlet inn feltobservasjoner fra østsiden av studieområdet, siden det i feltarbeidsperioden (oktober - november) allerede hadde akkumulert betydelige snømengder i høyereliggende områder på vestsiden av prosjektområdet, som antageligvis ville ha ført til en økning i frekvens og intensitet av målefeil. Den geografiske variasjonen innen prosjektområdet ble dermed i mindre grad dekket gjennom feltarbeidet i 2016, som også kan ha bidratt til ulikheter blant egenskaper av de respektive datasettene.

Modellene fra begge datasettene ble validert ved hjelp av "leave one out" kryssvalidering og ved å predikere vekselvis for prøveflater fra de alternative datasettene. Det bemerkes at sistnevnte metode for validering er foretrukket, siden det angående kryssvalidering eksisterer en viss uenighet om begrunnelse av antallet observasjoner som tas ut om gangen (Kohavi, 1995). I denne studien ble kryssvalidering benyttet i tillegg til validering med et uavhengig datasett for å oppnå en mer grundig evaluering av prediksjonsevnen av modellene ved å benytte dataene i mest mulig grad. "Leave one out" metoden ble anvendt på grunn av dens enkelhet og fordi den blir brukt mest i praksis. Men det påpekes at en økning i antallet observasjoner som blir tatt ut for hvert steg medfører en større bias og høyere prediksjonsfeil (Kohavi, 1995). Kryssvalidering anses derfor som en pessimistisk estimator for prediksjonsfeil, siden kun en del av dataene blir brukt til kalibrering av modellen for hvert steg.

Resultatene viste at den gjennomsnittlige differansen mellom predikert og feltmålt verdi, eller skjevheten av prediksjonene, var betydelig større ved valideringer da et uavhengig datasett ble

brukt. Dette kan forklares igjen av de forskjellene mellom de respektive datasettene. En gitt modell er best egnet til å representere sammenhengen mellom feltobservasjoner og laservariabler innen akkurat det datasettet modellen ble kalibrert med. Når modellene deretter blir brukt for å predikere egenskaper av prøveflater innen et annet datasett som er ulik datasettet som modellene ble tilpasset med, oppnår man en viss bias som resultatene i denne studien viste.

Da modellene ble validert på de respektive uavhengige valideringsdatasettene viste det seg at modellene som ble kalibrert med 2016-datasettet generelt sett ga dårligere prediksjoner enn modellene som ble kalibrert med 2015-datasettet. Det ble oppnådd en høyere RMSE for modellene fra 2016-datasettet for *Hd*, *Nt*, og *Nd*, og en høyere RMSE% *Hd*, *Ht* og *Nd*. Dette kan muligens forklares av den begrensede variasjonen innen 2016-datasettet (Vedlegg 4), som førte til ekstrapolering, der det ble predikert utenfor intervallet av x-verdiene i 2016-datasettet.

### 4.4 Metoder for prediksjon av behov for ungskogpleie

En logistisk modell ble testet som klassifikator for prediksjon av behov for ungskogpleie. Observasjonene ble klassifisert ut fra predikert treantall og dominerende høyde i kombinasjon med observert bonitet og direkte ved hjelp av laservariabler og bonitet. Stratifisert "2-fold" kryssvalidering viste at modellene var i stand til å predikere behov for ungskogpleie med en samlet nøyaktighet på henholdsvis 76 og 83 % og kappa-indeks på henholdsvis 0.44 og 0.60. Fra et forvaltningssynspunkt er prediksjoner for behov for ungskogpleie viktig i forbindelse med planlegging av skogskjøtsel, og påvisning av en akseptabel nøyaktighet av et slikt verktøy er derfor relevant for skogbruksplanlegging.

Den første modellen ble testet fordi det ble antatt at det er en sammenheng mellom totalt treantall og behov for ungskogpleie, der behov for ungskogpleie øker med økt tetthet. Også dominerende høyde ble antatt å være en viktig variabel, der behov for ungskogpleie øker med økende dominerende høyde (Varmola et al., 2004). I tillegg ble det antatt at ungskog på høyere bonitet med gitte høyde- og tetthetsegenskaper er mer sannsynlig for å ha behov for ungskogpleie enn ungskog med samme høyde- og tetthetsegenskaper på lavere bonitet, på grunn av hurtigere vekst. Den tilpassede modellen viste at disse hypotesene var riktige, der p-verdiene indikerte at alle de tre variablene var signifikante og positive fortegn indikerte en positiv korrelasjon mellom behov for ungskogpleie og de tre prediktorene. Ulempen med denne metoden er at man innfører en viss modellfeil ved å bruke predikerte verdier for totalt treantall og høyde. Prediksjonsevnen av modellen er dermed avhengig av nøyaktigheten av

prediksjonene for disse variablene, og som valideringsresultatene for modeller for totalt treantall viser, er feilen knyttet til prediksjoner for tetthet betydelig. En lavere nøyaktighet ved validering av den første modellen i forhold til den andre modellen, og en høyere nøyaktighet ved bruk av observerte verdier for dominerende høyde og totalt treantall både som "training" og "testing data" ved validering av den første modellen (Vedlegg 8), støtter denne påstanden.

Den andre metoden der behov for ungskogpleie ble predikert ut fra ulike laservariabler og bonitet ble testet i denne studien både for å unngå innføring av den ovennevnte modellfeilen, og fordi bestemmelse av behov for ungskogpleie er en subjektiv vurdering, som antageligvis ikke kun er avhengig av tetthet, høyde og bonitet. Selv da observerte verdier for dominerende høyde og totalt treantall ble brukt både som "training-" og "testing data" viste det seg at behov for ungskogpleie ikke kunne predikeres med betydelig høyere nøyaktighet (Vedlegg 8), der en samlet nøyaktighet på 87 % ble oppnådd. Dette understreker subjektiviteten knyttet til bestemmelsen av behov for ungskogpleie, og antyder at behov for ungskogpleie ikke er en enkel funksjon av ovennevnte variablene. Direkte klassifisering ved bruk av laservariabler har derfor gitt bedre resultater både i denne studien og i tidligere forskning (Ørka et al., 2015), og bør derfor foretrekkes.

Den samlede nøyaktigheten som ble oppnådd ved bruk av direkte klassifisering tilsvarte resultatene som ble oppnådd av Korhonen et al. (2013) og Ørka et al. (2015), som også brukte logistisk regresjon og oppnådde en samlet nøyaktighet på henholdsvis 77 og 80 % med kappa-indeks på henholdsvis 0.55 og 0.45. En metodisk forskjell var innføring av bestandsdata om bonitet, som i denne studien bidro til å øke nøyaktigheten av klassifikasjonene. Ulike logistiske modeller ble testet, også modeller der bonitet ikke inngikk som prediktor. Hovedtrenden var at modeller som ikke inkluderte *bonitet* som prediktor ga lavere nøyaktighet enn modeller der *bonitet* inngikk som prediktor. Korhonen et al. (2013) brukte ulike spektrale og teksturale karakteristikker fra flybilder i kombinasjon med laservariabler som prediktorvariabler i regresjonen, og Ørka et al. (2015) brukte kun laservariabler.

En aspekt som påpekes angående datasettet som ble benyttet i denne studien er autokorrelasjonen mellom prøveflater innen clustre. En antagelse som ligger til grunn for logistisk regresjon er at utfallene i et bernoulli-forsøk er uavhengige av hverandre (Wang, 1993). En betydelig andel av observasjonene som inngikk i analysen var clustret sammen, der 49 av totalt 82 prøveflater grenset direkte til minst en annen prøveflate, som resulterte i brudd på antagelsen om uavhengighet for en betydelig andel av observasjonene. ME modellanalysen

viste at den tilfeldige effekten *clusternummer* forklarte forholdsvis mye av variasjonen innen romlige egenskaper blant prøveflatene (Tabell 5). Variasjonen blant prøveflater innen clustre var mindre enn variasjonen mellom prøveflater som ble lagt ut individuelt (Tabell 2, vedlegg 4), og clustrene var forholdsvis homogene (vedlegg 5). Ifølge Wang (1993) øker variansen innen forholdet mellom antallet suksesser og det totale antallet observasjoner i et bernoulli-forsøk ettersom datasettet blir mer homogent. Effekten av å bruke et forholdsvis homogent datasett forventes derfor å produsere for optimistiske resultater.

Det påpekes også at det kan være en del usikkerhet knyttet til feltobservasjonene om behov for ungskogpleie. Observasjonene som ble benyttet i denne studien ble samlet inn av to forskjellige studenter gjennom feltarbeidet utført i 2015, og en student som ut fra bildene som ble tatt i felt gjennom feltarbeidet i 2016 har tolket behandlingsforslag subjektivt. Ingen av observasjonene har blitt bekreftet, og denne usikkerheten kan medføre en viss bias.

Det ble eksperimentert med ulike terskelverdier for bestemmelse av '0' og '1' klassifikasjoner. Ved en terskelverdi på 0.5 og 0.6 for henholdsvis modell 1 og modell 2 ble høyest samlet nøyaktighet oppnådd. I tillegg ble det testet om de beregnete sannsynligheter som den logistiske modellen ga som output kunne bli benyttet for å identifisere prøveflater der behov for ungskogpleie kunne predikeres med nær optimal nøyaktighet. Alle sannsynligheter > 0.8 ble identifisert ved validering av modell 2 (45 av 82 observasjoner) med forventningen om at det faktisk var behov for ungskogpleie for disse observasjonene. Men overraskende nok var det i dette subsettet fem observasjoner, som alle hadde en beregnet sannsynlighet for behov for ungskogpleie > 0.9, som faktisk ble klassifisert som 'ikke behov' i felt. Som regel hadde disse observasjonene en bonitet på 17 eller 20, og enten en observert dominerende høyde på > 3 meter eller observert totalt treantall på > 2000 $ha^{-1}$, som førte til en høy predikert sannsynlighet for behov for ungskogpleie. Det ble også gjort et forsøk på å identifisere prøveflater der prediksjoner for 'ikke behov' klassen var pålitelig. Alle observasjoner med en sannsynlighet for behov for ungskogpleie < 0.2 ble identifisert, men det var to observasjoner med en beregnet sannsynlighet < 0.2 som faktisk ble klassifisert som 'behov' i felt.

Det ble også testet om flere klasser for sannsynlighet for behov for ungskogpleie kunne forbedre resultatet. Tre klasser ble lagt: (1) 'ikke behov', (2) 'usikker' og (3) 'behov' med sannsynlighet (P) på henholdsvis $P < 0.33$, $0.33 < P < 0.66$ og $P > 0.66$. Men resultatet var at flere observasjoner som ellers hadde blitt klassifisert riktig, ble tildelte klassen 'usikker', og at

like mange av observasjonene i klassene (1) og (3) ble feilklassifisert. Ulike forsøk der observasjonene ble delt inn i forskjellige sannsynlighetsklasser førte ikke til en forbedring av resultatene, og metoden som ga høyest samlet nøyaktighet var dermed fremdeles den opprinnelige tilnærmingen der observasjonene ble delt inn i kun to klasser for 'behov' og 'ikke behov'.

Ved prediksjon av behov for ungskogpleie er det spesielt viktig at det unngås at et gitt ungskogareal blir klassifisert som 'ikke behov' mens det faktisk er behov på arealet. En slik feilklassifikasjon vil føre til at arealet blir glemt og at ungskogpleien ikke blir utført når det faktisk er behov for tiltak. Derfor er en user's accuracy på 69 % oppnådd ved validering av modell 2 et lovende resultat for 'ikke behov' klassen predikert av den logistiske modellen (tabell 15). Det er i mindre grad viktig at et gitt ungskogareal blir klassifisert som 'behov' mens det faktisk ikke er behov. En slik feilklassifikasjon vil i det verste fall føre til at et mannskap blir sendt til et gitt bestand uten at det er behov for ungskogpleie. En eventuell justering av terskelen, som i dette tilfellet ble satt på 0.6, kan derfor vurderes for å fremme nøyaktigheten innen 'ikke behov' klassen.

Kartlegging av behov for ungskogpleie over hele prosjektområdet medfører som fordel at ikke nødvendigvis alle bestand må bli oppsøkt i felt for å registrere behandlingsforslag. Når det er predikert enten behov eller ikke behov for ungskogpleie jevnt over stort sett hele arealet innen et gitt bestand, kan man gå ut fra at det ikke er nødvendig å oppsøke det bestandet i felt. Ved å identifisere slike bestand kan man potensielt redusere feltarbeidskostnader vesentlig.

En ulempe som oppstår ved bruk av logistisk regresjon for klassifisering av behov for ungskogpleie er at modellen antar at det er en lineær sammenheng mellom respons- og prediktorvariablene. Beregningsceller der vegetasjonen har nådd en dominerende høyde som overstiger den øvre grensa for økonomisk forsvarlig ungskogpleie har derfor forholdsvis stor sannsynlighet for å bli klassifisert som 'behov', siden modellen går ut fra en positiv sammenheng mellom vegetasjonshøyde og sannsynlighet for behov for ungskogpleie. For å unngå feilklassifikasjon av ungskogareal der overhøyden er så høy at det ikke lenger lønner seg å utføre ungskogpleie, hadde det vært en mulighet å klassifisere gridceller med en høy verdi for dominerende høyde, for eksempel > 7 meter, som 'ikke behov'. Siden det kun var et par observasjoner med en slik høy verdi for dominerende høyde i datasettet som ble benyttet i denne studien, har dette ikke blitt testet, men en slik seleksjon kan utføres med tilfredsstillende

nøyaktighet siden prediksjoner for dominerende høyde som regel er forholdsvis pålitelig. Det kan også være hensiktsmessig å kartlegge alle bestand der det er høy bonitet, for å oppsøke de bestandene uansett klassifikasjon, siden det spesielt ved hurtig voksende bestand er viktig at et eventuelt behov for ungskogpleie blir registrert.

Selv om de kontinuerlige verdiene som den logistiske modellen ga som output ikke kunne benyttes for å identifisere pålitelige prediksjoner i denne studien, er en samlet nøyaktighet på 83 % et lovende resultat for praktisk bruk av metoden innen operasjonelle takster, som antyder at FLS data, i kombinasjon med bestandsdata om bonitet, kan brukes i fremtiden for å redusere mengden av feltarbeid i operasjonelle takster.

# 5 IMPLIKASJONER OG ANBEFALINGER

Fremskaffing av nøyaktige prediksjoner av egenskapene til ungskog ut fra FLS data er utfordrende, men resultatene fra denne studien gir innsikt i metoder for prediksjon av romlige egenskaper til ungskog og behov for ungskogpleie. En rekke anbefalinger ble formulert for laserbasert taksering av ungskog, som blir diskutert i dette sub-kapittelet.

## 5.1 Variasjon innen datasettet

Det anbefales å samle inn et variert datasett gjennom prøveflatetaksten, der variasjonsbredden innen høyde og tetthetsegenskaper av ungskog er dekket i best mulig grad. Valideringsresultatene viste at modeller som ble kalibrert med et forholdsvis variert datasett (2015-datasettet) førte til mer nøyaktige prediksjoner enn modeller som ble kalibrert med et mindre variert datasett (2016-datasettet, Vedlegg 7). Dette understreker betydningen av å oppnå mest mulig variasjon i materialet brukt til modellkalibrering.

Også for å kunne predikere behov for ungskogpleie ved hjelp av logistisk regresjon bør datasettet være representativ for alle typer ungskog innen prosjektområdet. I operasjonelle takster bør en avveining mellom kostnader og gevinster vurderes. Et økt antall prøveflater medfører en økning i kostnader, og det er derfor viktig å finne metoder for hensiktsmessig allokering og utlegging av feltobservasjoner. Systematisk sampling kan resultere i et datasett som representerer populasjonen på best mulig måte, men som alternativ kan et utvalg som baseres på høyde og tetthetsvariabler beregnet på forhånd vurderes for å oppnå et variert datasett.

## 5.2 Aggregering

Valideringsresultatene viste at gjennomsnittlige prediksjonsfeil ble redusert ved aggregering av prøveflatene innen clustre, og denne effekten forventes å øke når flere enheter innen bestand blir aggregert. Det anbefales derfor å aggregere prediksjoner for individuelle gridceller, for å produsere nøyaktige estimater på bestandsnivå. Det antas dermed at skogeiere og skogforvaltere er spesielt interessert i nøyaktige data for hele bestand som behandlingsenheter, i motsetning til prediksjoner for individuelle gridceller. Ørka et al. (2015) viste at prediksjoner for de samme fire skoglige variablene som også har blitt testet i denne studien ga en lavere prediksjonsfeil på bestandsnivå enn feil som ble oppnådd ved prediksjon for individuelle observasjoner, enten ved kryssvalidering eller prediksjoner for beregningsflater som ble brukt

for validering. Dette bekrefter gevinsten av å aggregere prediksjoner for beregningsceller til bestandsnivå.

### 5.3 Direkte klassifisering

Det anbefales å benytte direkte klassifisering av behov for ungskogpleie, der laservariabler og bonitet inngår som prediktorvariabler i logistisk regresjon. Med en samlet nøyaktighet på 83 % oppnådd med stratifisert "2-fold" kryssvalidering viste denne studien at det er mulig å oppnå akseptabel nøyaktighet ved klassifisering av behov for ungskogpleie ved bruk av denne metoden.

Bruk av to klasser for behov for ungskogpleie ga best resultat i denne studien, og det anbefales derfor ikke å benytte noen flere sannsynlighetsklasser siden flere observasjoner som ellers hadde blitt klassifisert riktig som enten '0' eller '1' i denne studien ble klassifisert som 'usikker', uten at en reduksjon i feilklassifikasjoner innen de to ytterste sannsynlighetsklassene ble oppnådd. Flere klasser for behov for ungskogpleie forventes derfor å medføre støy i dataprosesseringen og gjøre tolkningen av prediksjonene mer komplisert. I tillegg anbefales det å teste noen forskjellige terskelverdier for klassifikasjonen, for å kunne oppnå høyest mulig samlet nøyaktighet.

Den samlede nøyaktigheten som ble oppnådd ved kryssvalidering viste at prediksjonsevnen av den logistiske modellen ikke ble redusert da modellen ble kalibrert med kun halvparten av observasjonene (n = 41). Dette antyder at, for prediksjon av behov for ungskogpleie, det ikke er nødvendig å investere ytterlige i innsamling av flere prøveflater enn det som allerede blir innsamlet for ungskog-stratumet ved prøveflatetaksten. Ørka et al. (2015) viste at et antall på mellom 40 og 50 prøveflater er tilstrekkelig for taksering av ungskog i en områdetakst, som er det samme antallet som benyttes i et stratum for eldre skog (Næsset, 2004).

### 5.4 Bestandsdata om bonitet

Resultatene oppnådd i denne studien antyder at *bonitet* som forklaringsvariabel ikke bidrar til å forklare variasjonen i høyde og tetthet av ungskog i like stor grad som laservariabler. Men for prediksjon av behov for ungskogpleie kan *bonitet* potensielt forbedre nøyaktigheten av klassifikasjonene oppnådd ved logistisk regresjon. Det anbefales derfor å inkludere bestandsdata om bonitet i klassifiseringen.

Som et alternativ kan det vurderes en metode der behov for ungskogpleie blir klassifisert direkte ut fra laservariabler, der bestand med høy bonitet blir identifisert og oppsøkt i felt uansett klassifikasjon. Dette for å unngå at tidspunktet for ungskogpleie passeres innen bestand der vegetasjonen vokser forholdvis hurtig.

### 5.5 Praktisk bruk av dataene

Som nevnt ovenfor anbefales det for operasjonelle områdetakster et sluttprodukt der ungskogegenskaper blir presentert på bestandsnivå i skogbruksplanen. Det anbefales videre at prediksjonene oppnådd med klassifiseringen av behov for ungskogpleie blir benyttet for å identifisere bestand der feltbefaring ikke er nødvendig, ved å subjektiv velge ut bestand der det er predikert enten behov eller ikke behov for ungskogpleie nærmere jevnt over hele bestandet. Spesielt for bestand hvorav nesten hele arealet blir klassifisert som å ha behov for ungskogpleie innen fem år kan det være hensiktsmessig å unnlate å oppsøke bestandet i felt for å redusere feltarbeidskostnader. I vedlegg 9 gis det noen eksempler på bestand i prosjektområdet der markbefaring ikke var nødvendig for registrering av behandlingsforslag siden nesten hele arealet ble klassifisert enten som 'behov' eller 'ikke behov' av den logistiske modellen.

Ved kartlegging av behov for ungskogpleie anbefales det å klassifisere gridceller med en høy verdi for predikert dominerende høyde, for eksempel > 7 meter, som 'ikke behov' for å redusere feilklassifikasjoner, siden det er stor sannsynlighet for at økonomisk forsvarlig tidspunkt for ungskogpleie har utløpt når skogbruksplanen presenteres til skogeieren. Videre anbefales det å oppsøke ungskogbestand med høy bonitet uansett klassifikasjon, siden det spesielt for hurtig voksende bestand er viktig at et eventuelt behov for ungskogpleie blir registrert.

### 5.6 Konklusjon og fremtidig forskning

Bruk av FLS data medfører store fordeler for taksering av ungskog, og resultatene fra denne studien har vist at laserdata kan benyttes for å predikere egenskaper av ungskog og kartlegge behov for ungskogpleie. Ved en områdetakst blir laserskanningen utført utover hele prosjektområdet, og laserdataene tilhørende ungskogarealet er uansett tilgjengelige. Økningen i takstkostnader ved bruk av disse dataene er derfor begrenset, og overstiger antageligvis ikke kostnadene av alternative takstmetoder eller kostnader som påløper fordi det tas beslutninger basert på feilaktig eller manglende informasjon.

Videre forskningsfokus anbefales på stratifisering etter bonitet, muligheter for bestemmelse av treslagssammensetning, og betydningen av antall feltmålinger. Stratifisering etter bonitet har gitt lovende resultater for eldre skog (Næsset, 2002), men datasettet som ble benyttet i denne studien var ikke tilstrekkelig for denne analysen. Bruk av bildematching eller en kombinasjon av FLS data og karakteristikker beregnet fra flybilder kan eventuelt gi en forbedring av bestemmelse av treslag i ungskogbestand. Ved å gjøre færre målinger kan feltarbeidskostnadene potensielt reduseres.

## LITERATUR


Amateis, R. L., & Burkhart, H. E. (2016). Simulating the effects of site index variation within loblolly pine plantations using an individual tree growth and yield model.

Anon. (1969). Beståndsvård och produktionsekonomi. Kungliga Skogsstyrelsen, Stockholm.

Arlot, S., & Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Statistics surveys, 4*, 40-79.

Bollandsås, O. M., Hanssen, K. H., Marthiniussen, S., & Næsset, E. (2008). Measures of spatial forest structure derived from airborne laser data are associated with natural regeneration patterns in an uneven-aged spruce forest. *Forest Ecology and Management, 255*(3), 953-961.

Breidenbach, J., McGaughey, R. J., Andersen, H.-E., Kändler, G., & Reutebach, S. (2007). *A mixed effects model to estimate stand volume by means of small footprint airborne lidar data for an American and a German study site.* Paper presented at the Proceedings of ISPRS workshop laser scanning.

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference understanding AIC and BIC in model selection. *Sociological methods & research, 33*(2), 261-304.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and psychological measurement, 20*(1), 37-46.

Eid, T. (2008). Planlegging for beslutninger I skog. Introduksjon til skogregistrering i SKOG205.

Evans, J. S., Hudak, A. T., Faux, R., & Smith, A. (2009). Discrete return lidar in natural resources: Recommendations for project planning, data processing, and deliverables. *Remote Sensing, 1*(4), 776-794.

Fahlvik, N. (2005). *Aspects of precommercial thinning in heterogeneous forests in southern Sweden* (Vol. 2005).

Flom, P. L., & Cassell, D. L. (2007). Stopping stepwise: Why stepwise and similar selection methods are bad, and what you should use. *NorthEast SAS Users Group (NESUG): Statistics and Data Analysis*.

Huuskonen, S., & Hynynen, J. (2006). Timing and intensity of precommercial thinning and their effects on the first commercial thinning in Scots pine stands. *Silva Fennica, 40*(4), 645.

Kohavi, R. (1995). *A study of cross-validation and bootstrap for accuracy estimation and model selection.* Paper presented at the Ijcai.

Korhonen, L., Pippuri, I., Packalén, P., Heikkinen, V., Maltamo, M., & Heikkilä, J. (2013). Detection of the need for seedling stand tending using high-resolution remote sensing data. *Silva Fennica, 47*(2), 1-20.

Kuha, J. (2004). AIC and BIC: Comparisons of Assumptions and Performance. *Sociological methods & research, 33*(2), 188-229.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *biometrics*, 159-174.

Magnussen, S., & Boudewyn, P. (1998). Derivations of stand heights from airborne laser scanner data with canopy-based quantile estimators. *Canadian journal of forest research, 28*(7), 1016-1031.

Maltamo, M., Næsset, E., & Vauhkonen, J. (2014). Forestry applications of airborne laser scanning. *Concepts and case studies. Manag For Ecosys, 27*, 2014.

Mauya, E. W., Ene, L. T., Bollandsås, O. M., Gobakken, T., Næsset, E., Malimbwi, R. E., & Zahabu, E. (2015). Modelling aboveground forest biomass using airborne laser scanner data in the miombo woodlands of Tanzania. *Carbon balance and management, 10*(1), 28.

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*(1), 88-99.

Næsset, E. (2004). Practical large-scale forest stand inventory using a small-footprint airborne scanning laser. *Scandinavian Journal of Forest Research, 19*(2), 164-179.

Naesset, E. (1997a). Determination of mean tree height of forest stands using airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing, 52*(2), 49-56.

Naesset, E. (1997b). Estimating timber volume of forest stands using airborne laser scanner data. *Remote sensing of environment, 61*(2), 246-253.

Næsset, E., & Bjerknes, K.-O. (2001). Estimating tree heights and number of stems in young forest stands using airborne laser scanner data. *Remote sensing of environment, 78*(3), 328-340.

Närhi, M., Maltamo, M., Packalén, P., Peltola, H., & Soimasuo, J. (2008). Kuusen taimikoiden inventointi ja taimikonhoidon kiireellisyyden määrittäminen laserkeilauksen ja metsäsuunnitelmatietojen avulla.

Nilsson, M. (1996). Estimation of tree heights and stand volume using an airborne lidar system. *Remote sensing of environment, 56*(1), 1-7.

Nordkvist, K., & Olsson, H. (2013). *Laserskanning och digital fotogrammetri i skogsbruket* (1401-1204). Retrieved from

Ørka, H. O., Gobakken, T., & Næsset, E. (2015). Taksering av ungskog med flybåren laserscanning etter arealmetoden. Prosjektrapport.

Ørka, H. O., Gobakken, T., & Næsset, E. (2016). Predicting Attributes of Regeneration Forests Using Airborne Laser Scanning. *Canadian Journal of Remote Sensing, 42*(5), 541-553.

Peduzzi, P., Concato, J., Kemper, E., Holford, T. R., & Feinstein, A. R. (1996). A simulation study of the number of events per variable in logistic regression analysis. *Journal of clinical epidemiology, 49*(12), 1373-1379.

Pippuri, I., Kallio, E., Maltamo, M., Peltola, H., & Packalén, P. (2012). Exploring horizontal area-based metrics to discriminate the spatial pattern of trees and need for first thinning using airborne laser scanning. *Forestry*, cps005.

Snowdon, P. (1992). Ratio methods for estimating forest biomass. *NZJ For Sci, 22*, 54-62.

Ussyshkin, V., & Theriault, L. (2010). *ALTM ORION: bridging conventional lidar and full waveform digitizer technology*: na.

Varmola, M., & Salminen, H. (2004). Timing and intensity of precommercial thinning in Pinus sylvestris stands. *Scandinavian Journal of Forest Research, 19*(2), 142-151.

Vastaranta, M., Holopainen, M., Yu, X., Hyyppä, J., Hyyppä, H., & Viitala, R. (2011). Predicting stand-thinning maturity from airborne laser scanning data. *Scandinavian Journal of Forest Research, 26*(2), 187-196.

Wang, Y. H. (1993). On the number of successes in independent trials. *Statistica Sinica*, 295-312.

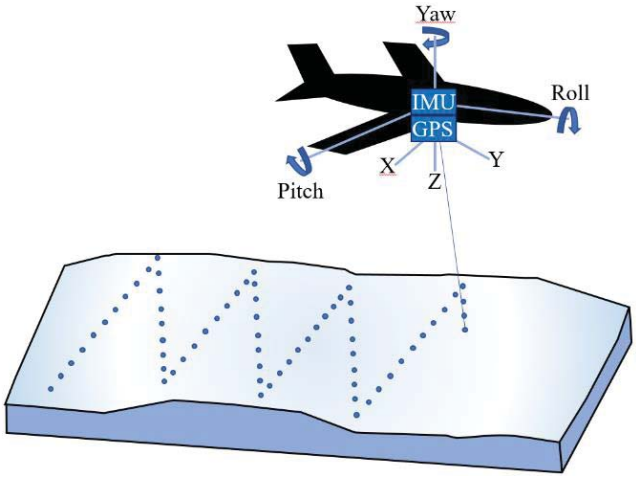
Watt, M. S., Meredith, A., Watt, P., & Gunn, A. (2013). Use of LiDAR to estimate stand characteristics for thinning operations in young Douglas-fir plantations. *New Zealand Journal of Forestry Science, 43*(1), 18.

White, J. C., Wulder, M. A., Varhola, A., Vastaranta, M., Coops, N. C., Cook, B. D., . . . Woods, M. (2013). A best practices guide for generating forest inventory attributes from airborne laser scanning data using an area-based approach. *For. Chron, 89*(722723), 5.

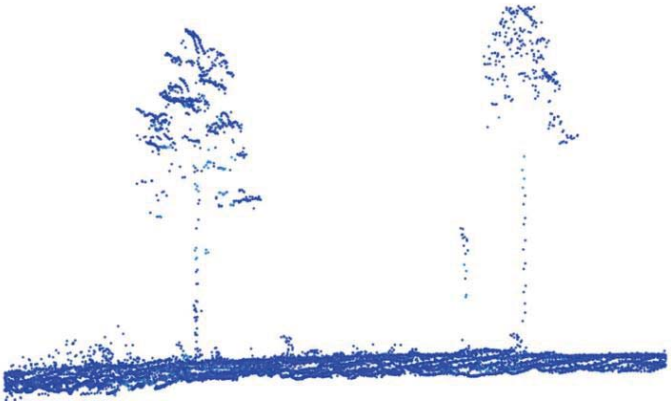
Wiant, H. V., Harner, E.J. (1979). Percent bias and standard error in logarithmic regression. Forest Science 20: 167-8.

Zuur, A. F., Ieno, E. N., Walker, N. J., Saveliev, A. A., & Smith, G. M. (2009). Mixed effects modelling for nested data *Mixed effects models and extensions in ecology with R* (pp. 101-142): Springer.

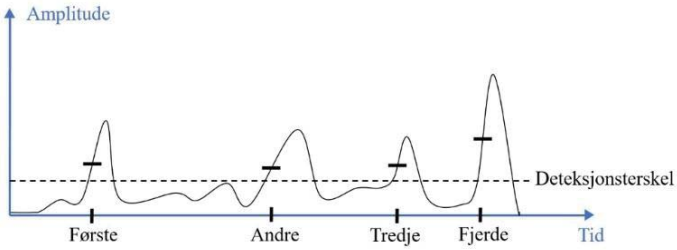
# Vedlegg 1 - Flybåren laserskanning




Figur 13: et FLS system.



Figur 14: illustrasjon av en punktsky (tverrprofil).




Figur 15: retursignalet blir registrert når den når en viss prosent av sin maksimale verdi, f.eks. 50 % (horisontale striper). Kun topper som når over deteksjonsgrensen blir registrert.

## Vedlegg 2 - Arealmetoden

Arealmetoden omfatter fem grunnleggende trinn:

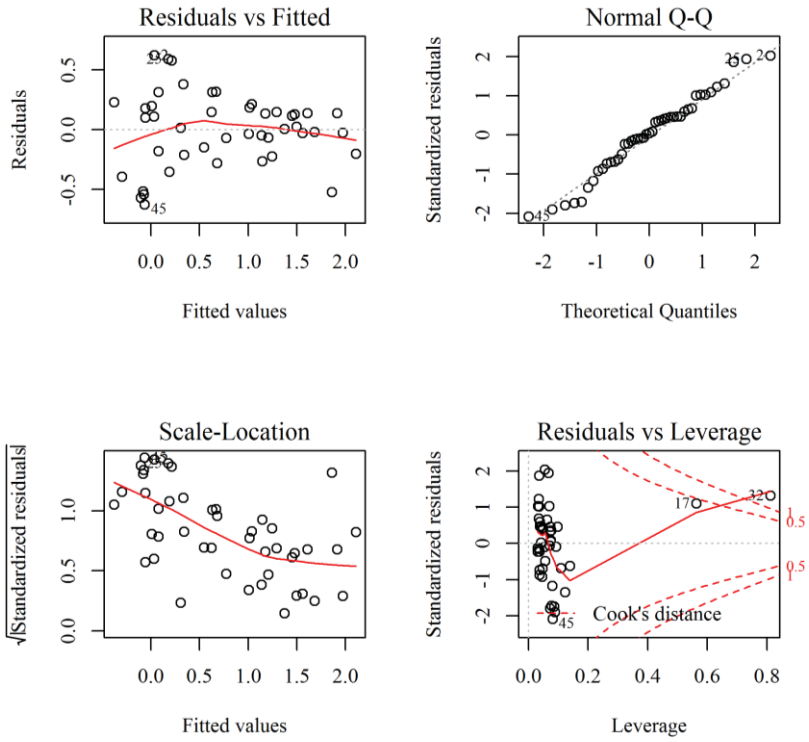
1. Bestandsavgrensning og tolkning: forvaltningsenhetene i takstområdet blir avgrenset ved hjelp av et ortofoto eller stereofotogrammetri. Egenskaper som dominerende treslag, bonitet og hogstklasse blir tolket manuelt for hvert bestand. Egenskapene blir som regel brukt til stratifisering av regresjonsmodellene.
2. Feltregistrering: prøveflater blir lagt ut gjennom takstområdet, mest ofte i henhold til stratifiserte, sannsynlighetsbaserte utvalgsteknikker der systematisk sampling med et tilfeldig startpunkt blir benyttet, ofte i kombinasjon med cluster sampling. Konvensjonelle feltobservasjoner blir registrert på stedfestede prøveflater.
3. Innsamling og prosessering av laserdata: laserdataene blir samlet inn for hele takstområdet. Området blir delt opp i gridceller der arealet av en celle tilsvarer arealet av en prøveflate. Ulike laservariabler som representerer punktskyen blir beregnet fra prøveflatene, vanligvis primært i form av tetthet- og høydebaserte laservariabler. Laserekko klassifiseres i kategoriene bakke eller vegetasjon ved hjelp av en automatisk algoritme som klassifiserer punktene etter naboskap mellom ekko. Ekkoene som blir klassifisert som bakketreff inngår i terrengmodellen. Terrengmodellen fungerer som et estimat på terrengoverflaten, og høyder til vegetasjonstreff blir beregnet med utgangspunkt i terrengmodellen som referanseverdi. Hvert ekko får da tildelt, i tillegg til x og y koordinater, en høyde i forhold til bakken. Laservariablene som representerer den romlige strukturen innen punktskyen blir ekstrahert for hver prøveflate, i første rekke høydepersentiler og tetthetsvariabler som karakteriserer frekvensen av ekkoene over en viss høydeterskel.
4. Modellering: det utføres en regresjonsanalyse der ulike skoglige egenskaper blir valgt som responsvariabler i modeller med ulike laservariabler som forklaringsvariabler. Det sørges for at et hensiktsmessig antall prøveflater blir fordelt utover strataene, at de mest signifikante forklaringsvariabler blir inkludert i modellene, og at modellformen passer til datasettet.
5. Estimering på bestandsnivå: modellene brukes for å predikere de ønskete skoglige egenskapene for hver gridcelle i takstområdet. Prediksjonene til de enkelte celler innen bestand blir aggregert for å fremskrive et estimat for hvert bestand.

# Vedlegg 3 - Transformasjoner

Log- transformasjon av responsvariabelen:

| | ***log(Hd)*** |
|---|---|
| *Konstantledd* | 0.09 ± 0.11 |
| *Hmax* | 0.30 ± 0.03 *** |
| *stdevHmax* | -0.55 ± 0.07 *** |
| *stdevHcv* | 0.55 ± 0.40 |
| $R^2$ | 0.84 |
| $R^2_{adj}$ | 0.83 |
| p-modell | < 2.20E-16 |
| n | 45 |

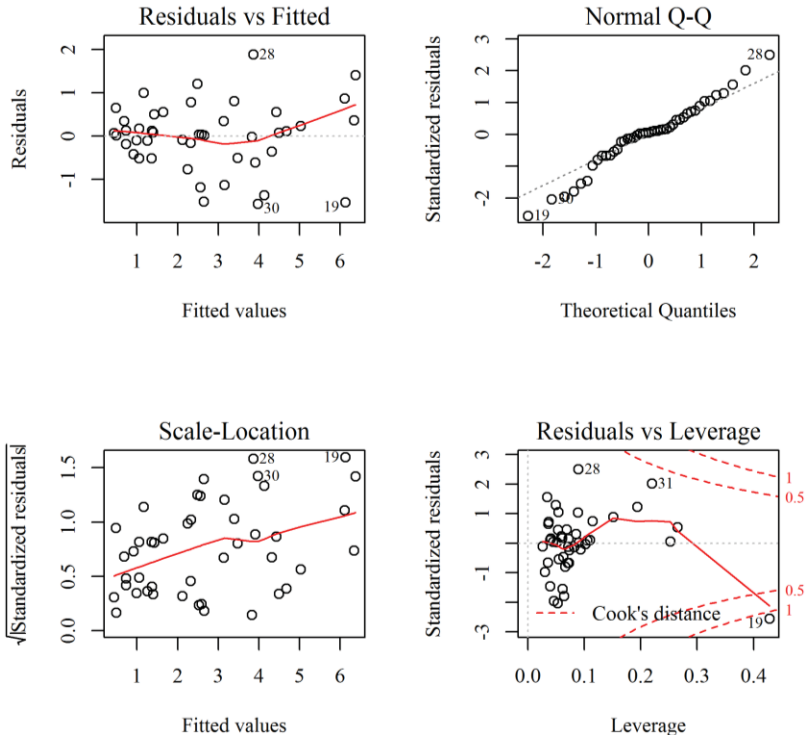
Signif. koder: 0 '*** 0.1 ' '




Log- transformasjon av alle prediktorvariabler:

| | ***Hd*** |
|---|---|
| *Konstantledd* | 5.58 ± 0.86 *** |
| *logH40* | 1.05 ± 0.24 *** |
| *logH10* | 1.23 ± 0.35 *** |
| *logstdevD0* | 0.42 ± 0.22 . |
| $R^2$ | 0.84 |
| $R^2_{adj}$ | 0.82 |
| p-modell | 4.30E-16 |
| n | 45 |

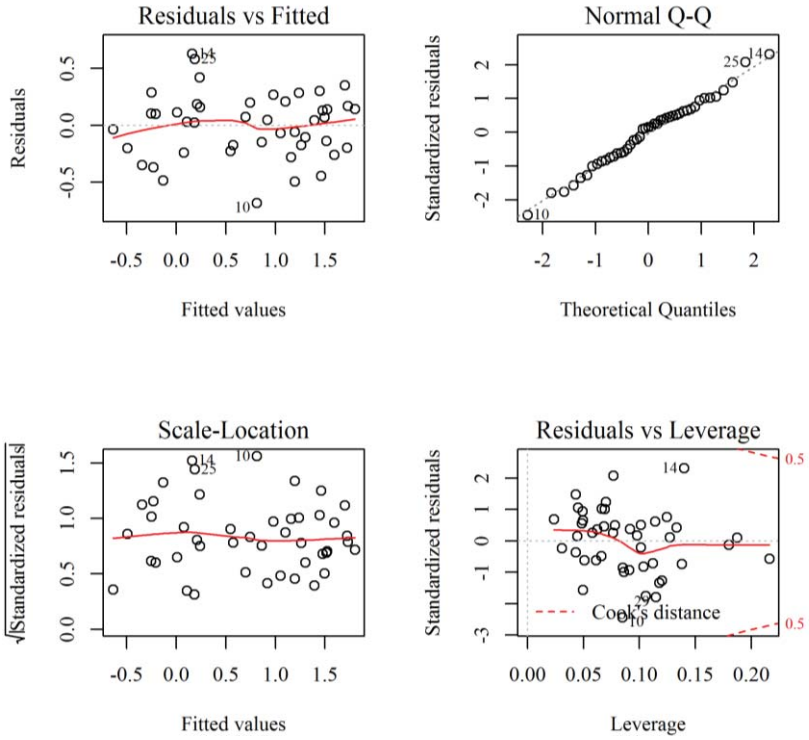
Signif. koder: 0 '***' 0.05 '.' 0.1 ' '



Log-log transformasjon:

| | ***log(Hd)*** |
|---|---|
| *Konstantledd* | 0.87 ± 0.25 ** |
| *logD0* | 0.47 ± 0.10 *** |
| *logH50* | 0.41 ± 0.08 *** |
| *logstdevD7* | -0-16 ± 0.78 * |
| $R^2$ | 0.84 |
| $R^2_{adj}$ | 0.82 |
| p-modell | 4.30E-16 |
| n | 45 |

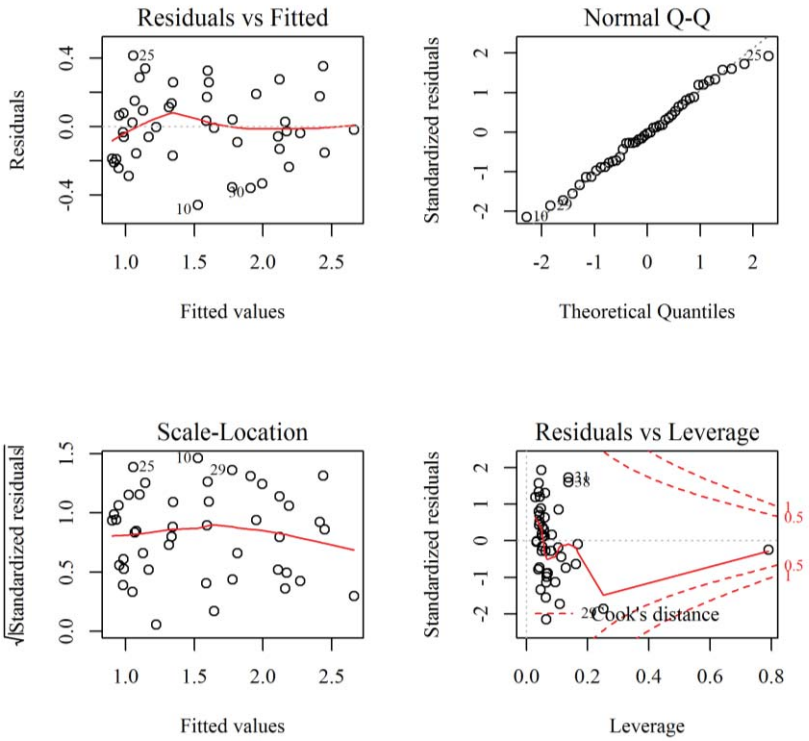
Signif. koder: 0 '***' 0.001 '**' 0.01 '*'




Sqrt- transformasjon av responsvariabelen:

| | ***sqrt(Hd)*** |
|---|---|
| *Konstantledd* | 0.80 ± 0.06 *** |
| *D0* | 1.46 ± 0.36 *** |
| *H50* | 0.35 ± 0.05 *** |
| *D5* | -1.60 ± 0.51 ** |
| $R^2$ | 0.86 |
| $R^2_{adj}$ | 0.85 |
| p-modell | < 2.20E-16 |
| n | 45 |

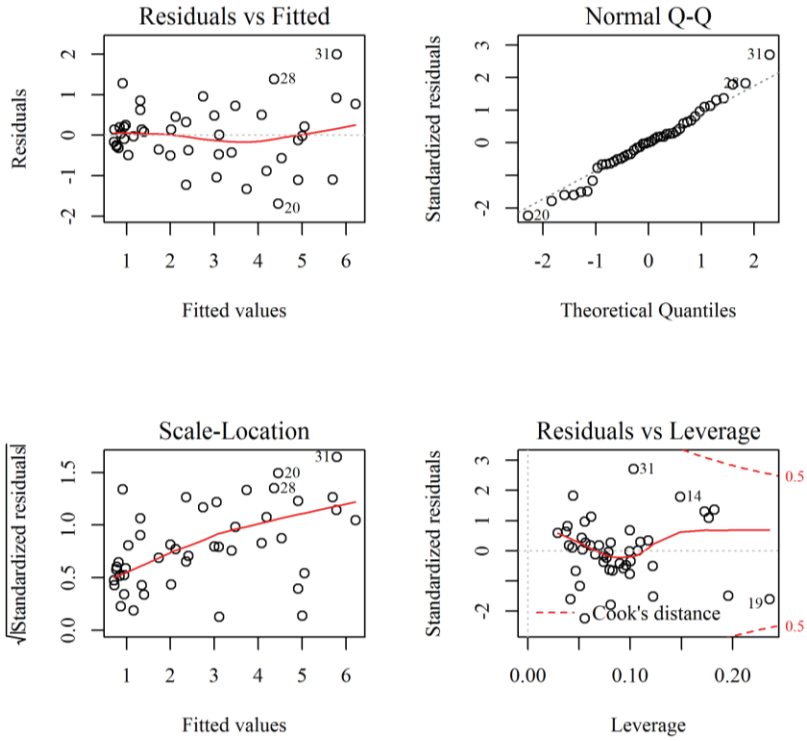
Signif. koder: 0 '***' 0.001 '**'



Sqrt- transformasjon av alle prediktorvariabler:

| | ***Hd*** |
|---|---|
| *Konstantledd* | -0.80 ± 0.32 * |
| *sqrtH40* | 3.09 ± 0.47 *** |
| *sqrtD0* | 2.65 ± 1.18 * |
| *sqrtstdevD7* | -5.45 ± 2.76 . |
| R2 | 0.85 |
| $R^2_{adj}$ | 0.83 |
| p-modell | < 2.20E-16 |
| n | 45 |

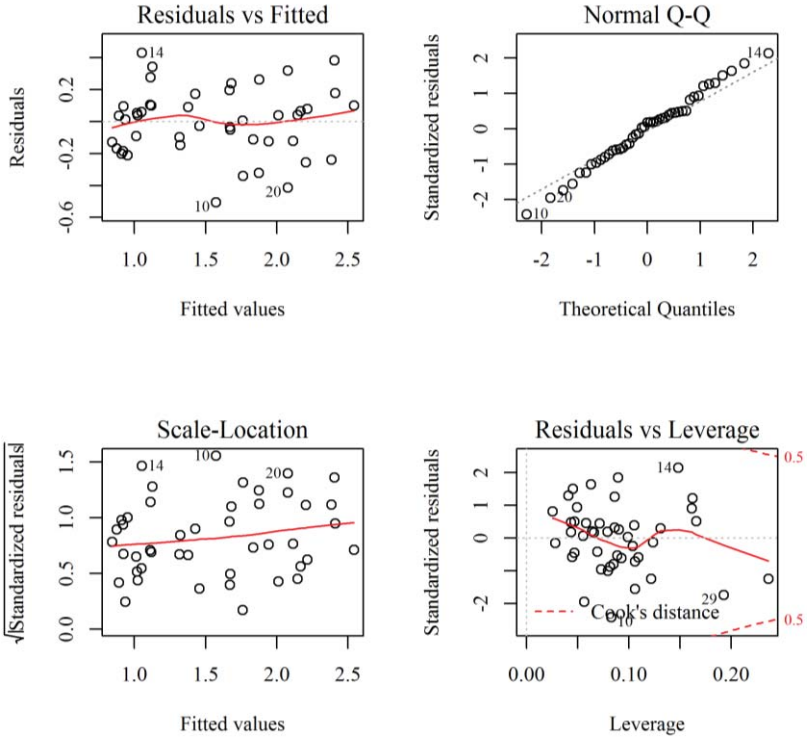
Signif. koder: 0 '***' 0.01 '*' 0.05 '.'




Sqrt- transformasjon av respons- og prediktorvariablene:

| | ***sqrt(Hd)*** |
|---|---|
| *Konstantledd* | -0.42 ± 0.09 *** |
| *sqrtH50* | 0.72 ± 0.12 *** |
| *sqrtD0* | 1.16 ± 0.32 *** |
| *sqrtstdevD7* | -1.58 ± 0.78 . |
| $R^2$ | 0.86 |
| $R^2_{adj}$ | 0.85 |
| p-modell | < 2.20E-16 |
| n | 45 |

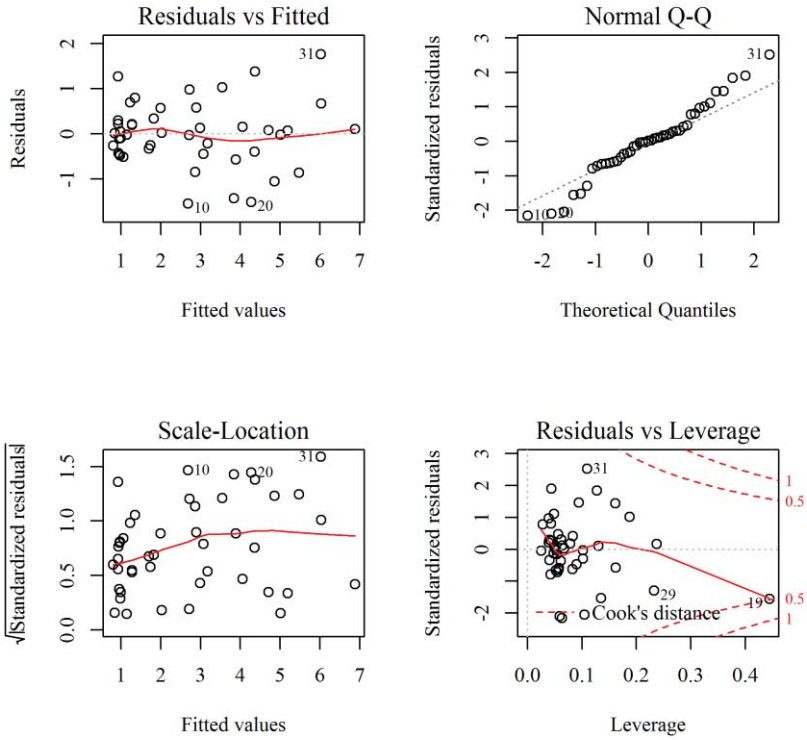
Signif. koder: 0 '***' 0.05 '.'



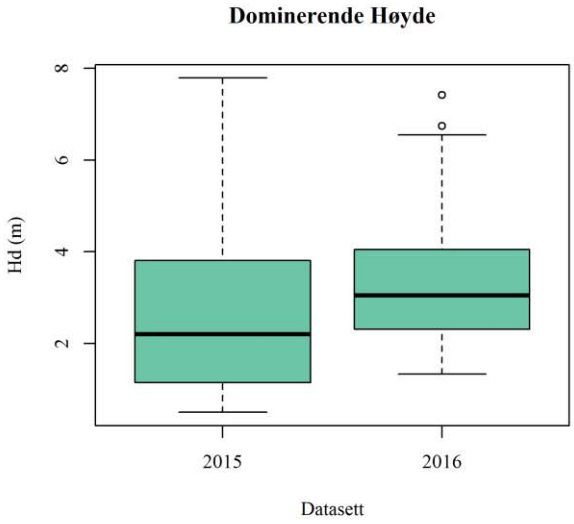
Ingen transformasjon:

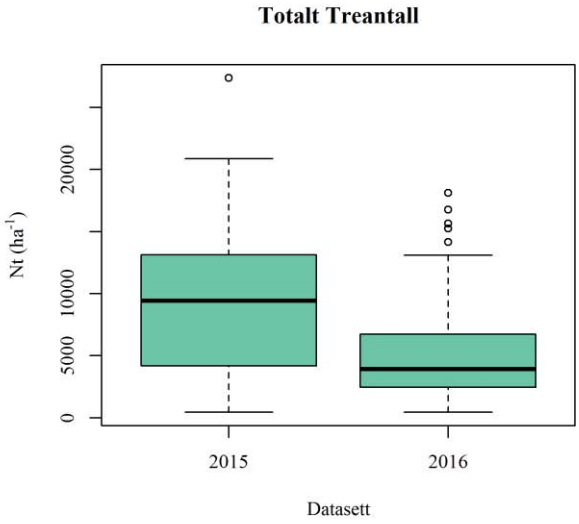
| | ***Hd*** |
|---|---|
| *Konstantledd* | -0.66 ± 0.19 ** |
| *D0* | 2.78 ± 0.95 ** |
| *H50* | 1.18 ± 0.15 *** |
| *stdevD7* | -17.99 ± 7.14 * |
| $R^2$ | 0.86 |
| $R^2_{adj}$ | 0.84 |
| p-modell | < 2.20E-16 |
| n | 45 |

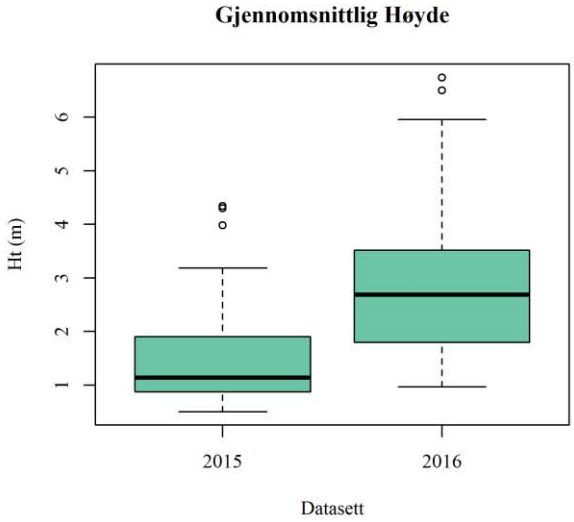
Signif. koder: 0 '***' 0.001 '**' 0.01 '*'

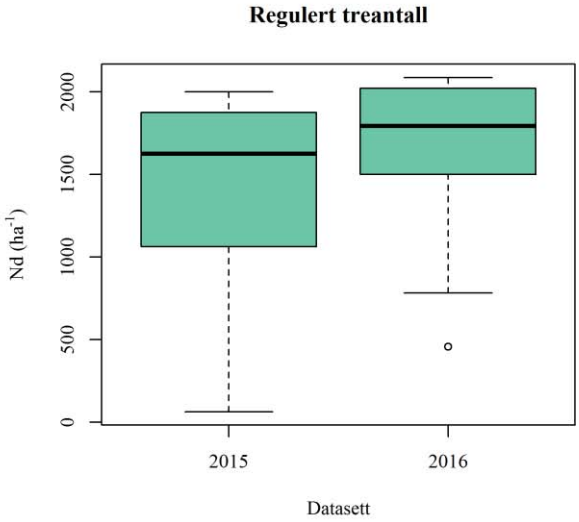


# Vedlegg 4 - Boksdiagram for 2015- og 2016-datasettene

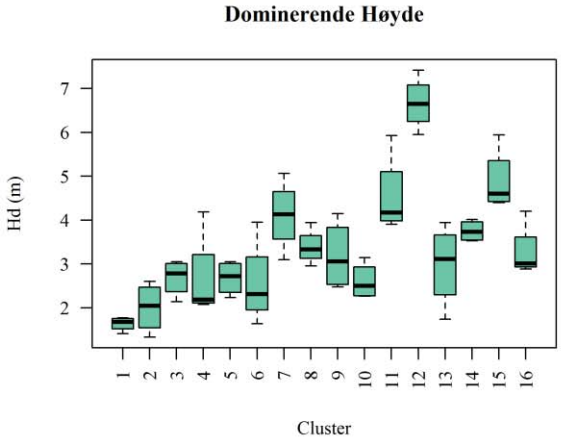


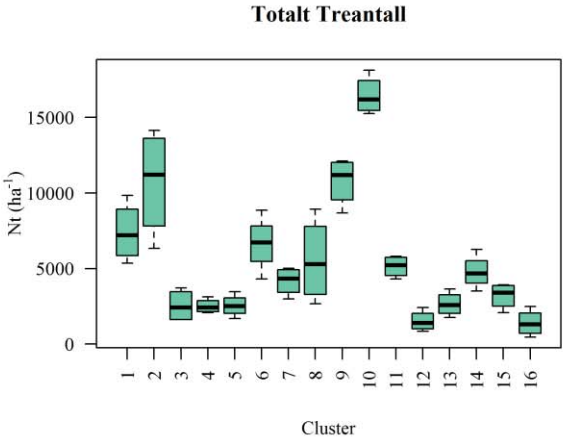



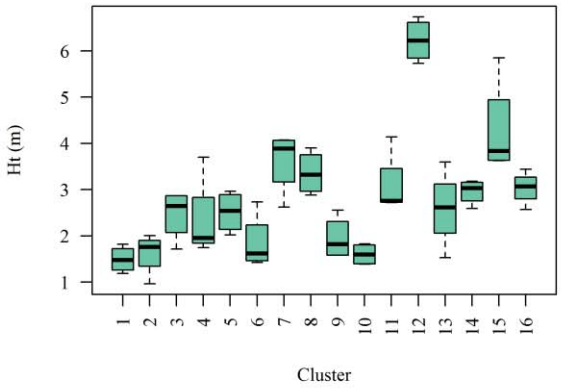



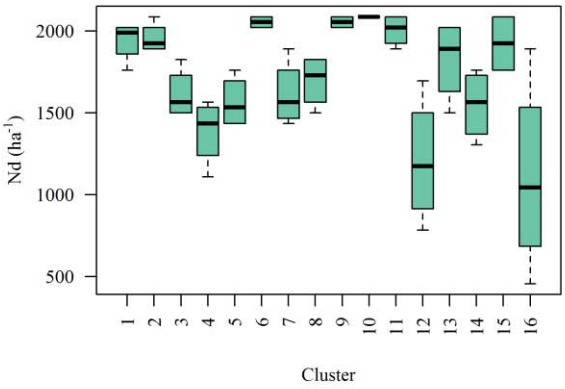



# Vedlegg 5 - Boksdiagram for clustrene



# Vedlegg 6 - Bonitetsklasser

Antall observasjoner per bonitetsklasse for begge datasett

| | 2015 | | 2016 | |
|---|---|---|---|---|
| **Bonitet** | Antall | Prosent | Antall | Prosent |
| 11 | 3 | 7 | 4 | 6 |
| 14 | 10 | 22 | 12 | 18 |
| 17 | 27 | 60 | 40 | 63 |
| 20 | 5 | 11 | 8 | 13 |
| Total | 45 | 100 | 64 | 100 |

# Vedlegg 7 - Resultatoversikt validering

| | | | | | Differanse mellom predikert og feltmålt verdi | | | |
|---|---|---|---|---|---|---|---|---|
| | Respons | RMSE | RMSE% | $\bar{y}$ | min | max | gjennomsnitt | standardavvik |
| Kryssvalidering av lineære regresjonsmodeller (n = 45) | *Hd* (m) | 0.81 | 29.8 | 2.72 | -2.29 | 1.79 | 0.00 | 0.82 |
| | *Nt* ($ha^{-1}$) | 6390 | 66.6 | 9588 | -14943 | 14890 | 0.02 | 6462 |
| | *Ht* (m) | 0.59 | 37.1 | 1.60 | -1.58 | 1.20 | 0.01 | 0.60 |
| | *Nd* ($ha^{-1}$) | 486 | 34.4 | 1415 | -1186 | 1063 | -48 | 490 |
| Validering med 2016-datasettet (n = 64) | *Hd* (m) | 0.74 | 27.2 | 2.72 | -2.71 | 0.83 | -0.33 | 0.67 |
| | *Nt* ($ha^{-1}$) | 4606 | 48.0 | 9588 | -10818 | 10157 | 1021 | 4526 |
| | *Ht* (m) | 1.05 | 66.1 | 1.60 | -4.08 | 1.25 | -0.26 | 1.03 |
| | *Nd* ($ha^{-1}$) | 428 | 30.2 | 1415 | -1456 | 852 | -96 | 420 |
| Validering med 2016-datasettet, aggregert (n = 16) | *Hd* (m) | 0.57 | 21.0 | 2.72 | -1.77 | 0.41 | -0.33 | 0.48 |
| | *Nt* ($ha^{-1}$) | 4154 | 43.3 | 9588 | -8893 | 7788 | 1021 | 4159 |
| | *Ht* (m) | 0.93 | 58.2 | 1.60 | -2.91 | 1.13 | -0.26 | 0.92 |
| | *Nd* ($ha^{-1}$) | 367 | 25.9 | 1415 | -719 | 622 | -96 | 365 |
| Kryssvalidering av mixed effects modellene (n = 64) | *Hd* (m) | 0.74 | 22.0 | 3.35 | -2.49 | 1.78 | 0.00 | 0.74 |
| | *Nt* ($ha^{-1}$) | 3412 | 62.1 | 5498 | -9100 | 8825 | -12.6 | 3439 |
| | *Ht* (m) | 0.90 | 32.0 | 2.81 | -2.74 | 1.48 | 0.00 | 0.91 |
| | *Nd* ($ha^{-1}$) | 327 | 19.0 | 1724 | -773 | 779 | -10.20 | 330 |
| Validering med 2015-datasettet (n = 45) | *Hd* (m) | 0.92 | 27.5 | 3.35 | -1.65 | 3.06 | 0.42 | 0.87 |
| | *Nt* ($ha^{-1}$) | 7059 | 128.4 | 5498 | -14642 | 18271 | -752 | 7057 |
| | *Ht* (m) | 0.64 | 22.7 | 2.81 | -1.18 | 1.83 | 0.35 | 0.59 |
| | *Nd* ($ha^{-1}$) | 474 | 27.5 | 1724 | -637 | 1273 | 136 | 476 |

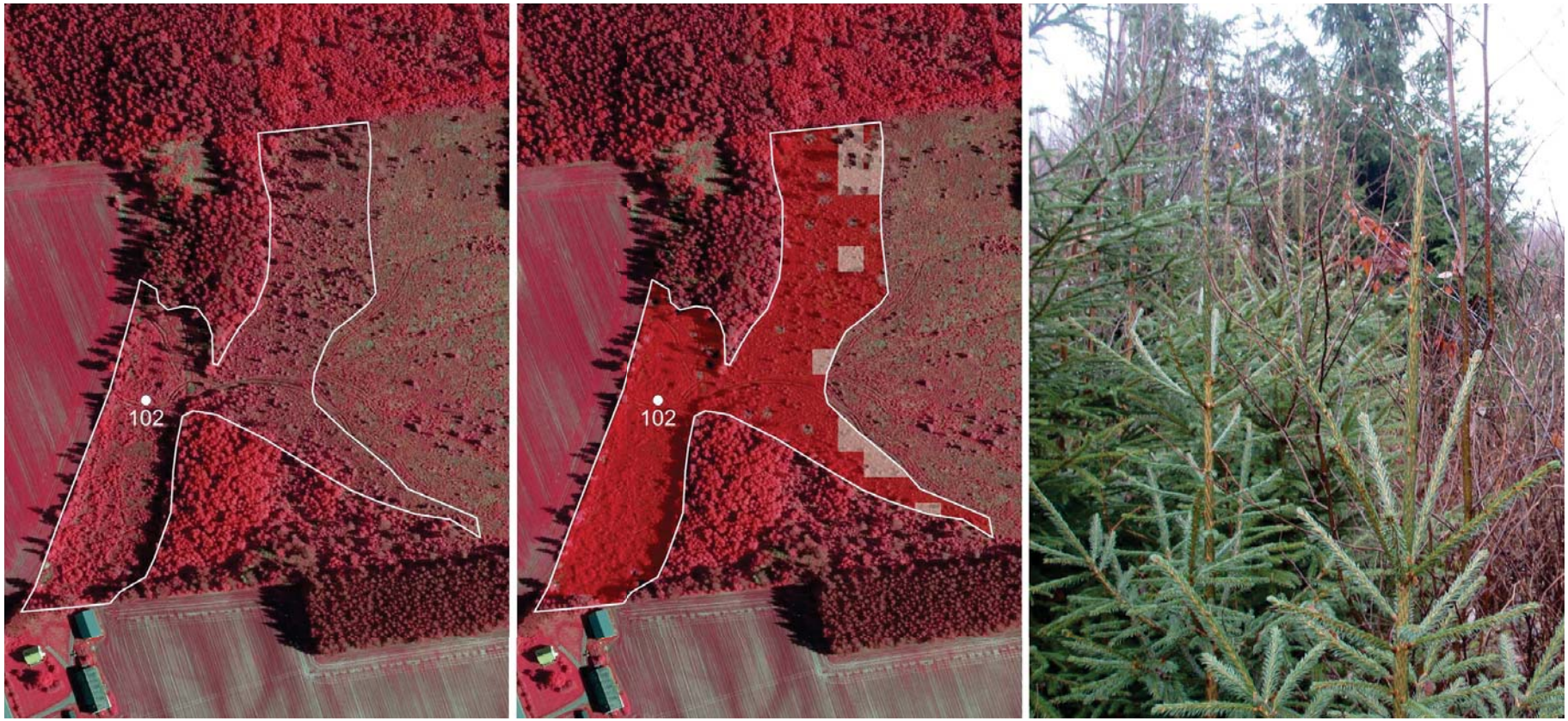
Der $\bar{y}$ = gjennomsnittlig feltmålt verdi

# Vedlegg 8 - Prediksjon med observerte verdier

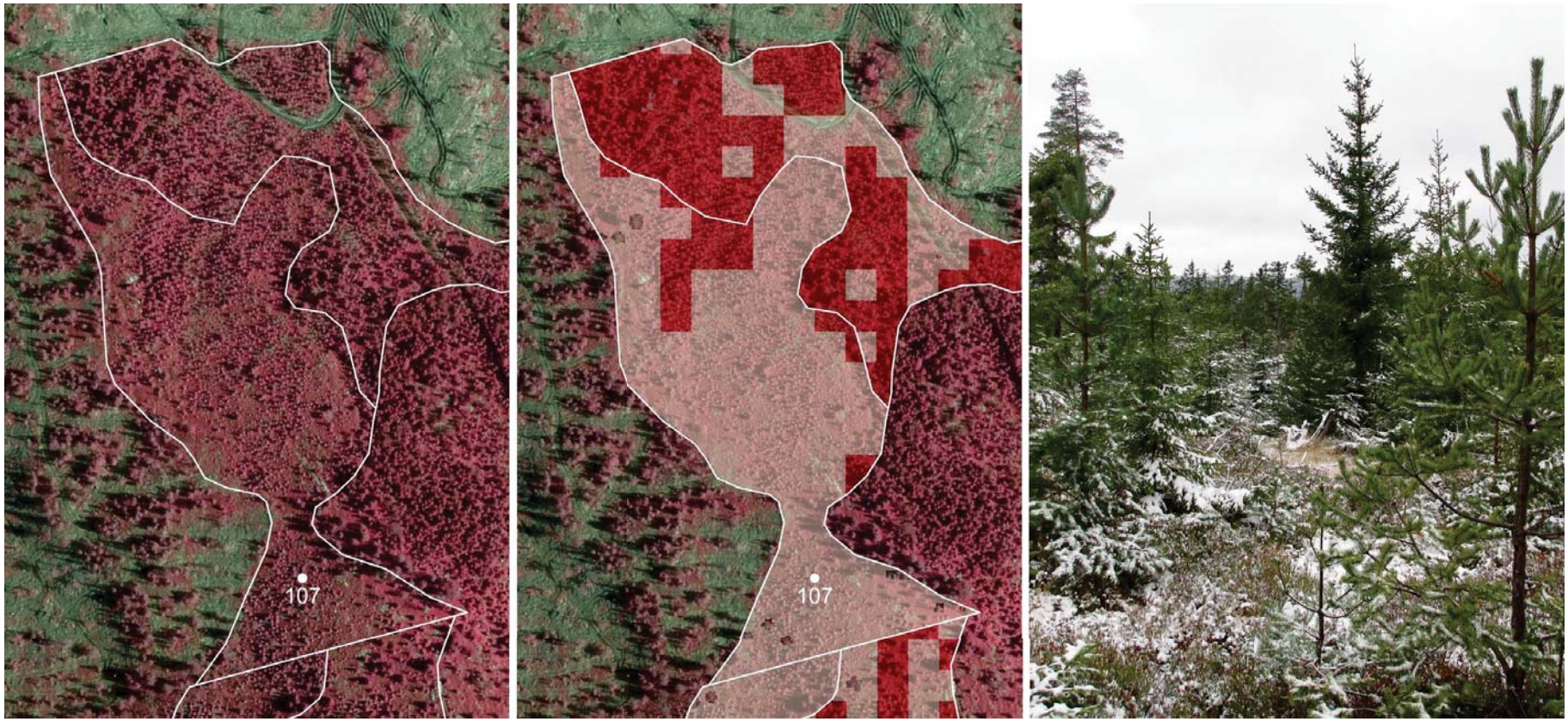
Feilmatrise for klassifikasjon ved validering av den logistiske regresjonsmodellen der observerte verdier for totalt treantall og dominerende høyde ble benyttet både som "training" og "testing data"

| Modell | Feltobservasjoner | | | |
|---|---|---|---|---|
| | Ikke behov | Behov | Total | User's accuracy |
| Ikke behov | 17 | 4 | 21 | 0.81 |
| Behov | 7 | 54 | 61 | 0.89 |
| Total | 24 | 58 | 82 | |
| Producer's accuracy: | 0.71 | 0.93 | | |
| Samlet nøyaktighet: | 0.87 | | | |
| Kappa: | 0.66 | | | |

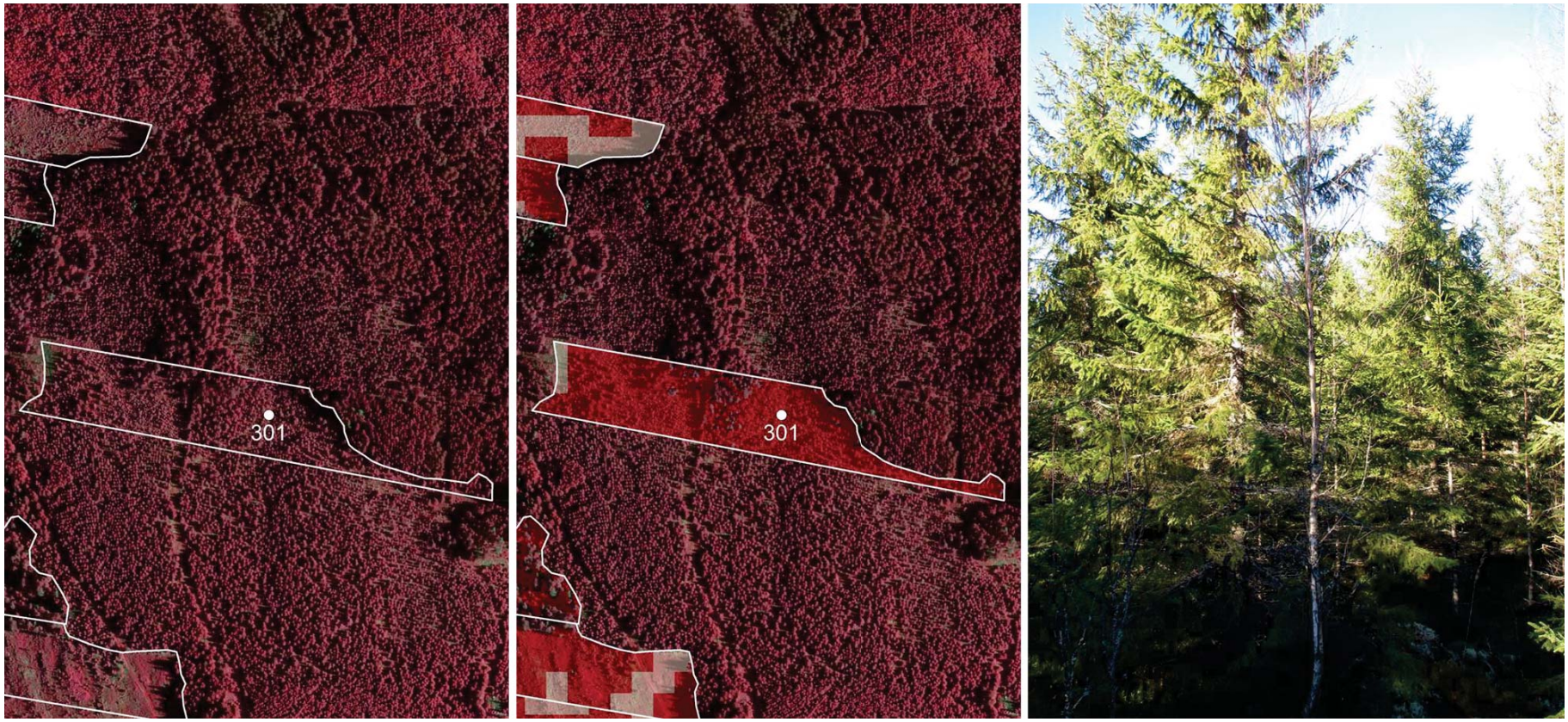
## Vedlegg 9 - Eksempler på kartlegging av behov for ungskogpleie




Venstre: et avgrenset ungskogbestand. Midterst: behov for ungskogpleie predikert av den logistiske modellen. Rød farge indikerer behov for ungskogpleie og hvit farge indikerer ingen behov. Høyre: skogtilstanden ved observasjon nr. 102 i felt (2016).



Venstre: et avgrenset ungskogbestand. Midterst: behov for ungskogpleie predikert av den logistiske modellen. Rød farge indikerer behov for ungskogpleie og hvit farge indikerer ingen behov. Høyre: skogtilstanden ved observasjon nr. 107 i felt (2016).



Venstre: et avgrenset ungskogbestand. Midterst: behov for ungskogpleie predikert av den logistiske modellen. Rød farge indikerer behov for ungskogpleie og hvit farge indikerer ingen behov. Høyre: skogtilstanden ved observasjon nr. 301 i felt (2016).


**Norges miljø- og biovitenskapelig universitet**
Noregs miljø- og biovitskapelege universitet
Norwegian University of Life Sciences

Postboks 5003
NO-1432 Ås
Norway