

ANALYSE AV DATA INNSAMLET MED FLYBÅREN LASERSKANNER I FIRE FORSKJELLIGE SKOGOMRÅDER I SØRØST NORGE

ANALYSING DATA COLLECTED WITH AIRBORNE LASER SCANNING FROM FOUR
DIFFERENT FOREST AREAS IN SOUTHEAST OF NORWAY

SANELA DZAKMIC

UNIVERSITETET FOR MILJØ- OG BIOVITENSKAP
INSTITUTT FOR NATURFORVALTNING
MASTEROPPGAVE 30 STP. 2009



FORORD

Bruken av flybåren laseskanner innenfor skogbruket er viet stor oppmerksomhet i de siste årene siden takstmetoden har vist meget lovende resultater. Laserens evne til å fange opp informasjonen i den varierende skogstrukturen gjør denne metoden svært aktuell. Fokus på muligheten av å utnytte laserdataene i flere områder som kan redusere kostnadene i planfirmaene førte til utformingen av problemstillingen.

Utarbeiding av denne oppgaven har vart tidskrevende, utfordrende, spennende, lærerikt og ikke minst gitt meg et godt grunnlag for videre arbeid med laserskanning innenfor skogbruket. Det sitter flest positive minner igjen, og jeg ser tilbake på dette som en positiv periode i livet mitt og min studietid på Ås.

Jeg vil først og fremst gi en stor takk til min veileder førsteamanuensis Terje Gobakken og min andreveileder Ole Martin Bollandsås for god hjelp med innspill, korrigeringer underveis og for alle data som er bitt stilt for rådighet. I tillegg vil jeg takke stipendiat Hans Ole Ørka for uvurderlig støtte med R og stipendiat Nadja Thieme for GIS -hjelp og moralsk støtte.

En takk må også rettes til mine kollegaer i fra tidligere Prevista AS for støtte og tålmodighet.

Til slutt vil jeg gi en stor takk til foreldrene mine, alle venner og kollegaer ved INA som har støttet og oppmuntret meg gjennom denne tiden!

SAMMENDRAG

Laserdata og feltdata fra totalt 639 georefererte prøveflater fra fire forskjellige områder lokaliserte i den sørlige delen av Norge ble benyttet i denne studien. Basert på disse dataene fra både ung og eldre skog ble sammenhengen mellom variabler avledet fra laserdataene og feltmålt volum modellert ved hjelp av regresjonsanalyse. Dataene var innsamlet som en del av operasjonelle lasertakster.

To forskjellige ALS instrumentene var brukt i disse fire forskjellige områdene. Punkttetthet varierte fra 0,8 til 1,1 m⁻² og gjennomsnittlig fotavtrykkdiameter varierte fra 20 til 36 cm.

Det er utledet variable relatert til både høyde og tetthet ut fra laserdataene. Disse ble brukt som uavhengige variabler i regresjonen. Til sammen var det lagret 62 regresjonsmodeller for volumestimering basert på forskjellig organisering av datamaterialet og for forskjellige treslagsandeler innenfor hvert område uavhengig av stratum og innenfor hvert stratum uavhengig av område.

Alle utvikla modeller hadde en god forklaringsgrad. Arealbasert metode var brukt til å få informasjon om volum (m³/ha) på bestands/kontrollflate-nivå. Valideringen av disse modellene blei utført mot uavhengige datasett fra hvert enkelt område.

Resultatene varierte for ulike modellbaser, det vil si ulik organisering av datamaterialet på grunnlag av område, stratum og treslag. Valideringsresultatene hadde et intervall på RMSE fra 10 % til 44 % og en systematisk feil fra 0,1 % til 35,5 %.

Resultatene fra valideringen ved bruk av ulike treslagsandelgrenser var veldig varierende som følge av at treslagsandel ikke blei betraktet som kontinuerlig variabel i de utvalgte modeller.

Statistiske effekter av forskjellige områder og strata på 8 regresjonsmodeller blei analysert med innføring av dummyvariabler. I analysene der dummyvariable representerte område hadde de signifikant effekt på regresjonsmodellene mens dummyvariablene som representerte stratum hadde ingen signifikant effekt på regresjonsmodeller.

Det ble funnet signifikante forskjeller mellom områdene. Dette medfører at man bør være varsom med sammenslåing av ulike takster ettersom påvirkninger av variabiliteten i skogen er knyttet til et område.

ABSTRACT

Laser data and field data from total 639 georeferenced field sample plots from four different areas located in the southern part of Norway was used in this study. Based on these data from young and mature forest empirical relationships between various metrics derived from the laser data and field measured volume was modeled using regression analysis. Sample plots had been measured as part of operational stand-based forest inventories.

Two different ALS instruments were used in these four different areas. The point density varied from 0.8 to 1.1 m⁻² and mean footprint diameter varied from 20 to 36 cm.

The regression analysis included independent variable related to canopy height and another related to canopy density. In total 62 regression models were calculated for volume estimation based on different data material structure and tree species composition for each area independent of strata and each stratum independent of area.

All of the developed models had good coefficients of determination. The models were used to predict volume (m³/ha) on stand/test plot level. The validation of these models was carried out by an independent dataset from each of the four areas.

Results obtained based on models developed using different subsets of the sample plots differed because the number of sample plots and stands used both for model development and validation varied. The validation results had a range of RMSE from 10% to 44% and biases from 0.1% to 35.5%.

The results from validation of models developed using different tree species composition had high variability since tree species composition were handled as discrete variable.

The statistical effect of area and stratum on 8 regression models was analyzed by adding a dummy variable to selected models. In analyses where dummy variables were representing area, they had a significant effect on the regression models while the dummy variables representing strata had no significant effects in the regression models.

It was revealed that district had a significant effect on the models. This entails that one should be careful with merging different inventories since influence of variations in forest is attached by district.

Innholdsfortegnelse

1	Innledning.....	1
2	Flybåren laser	3
2.1	ALS-komponenter	3
2.2	Laser pulser og ALS-typer	5
3	Operasjonell (kommersiell) lasertakst til skogbruksformål	7
3.1	De to ulike metoder av lasertakst	7
3.2	Stratifisering	8
4	Målsetting.....	9
5	Materiale.....	10
5.1	Studieområder.....	10
5.2	Datainnsamling - prøveflatadata.....	11
5.2.1	Beskrivende statistikk fra prøveflatene per område	12
5.2.2	Beskrivende statistikk fra valideringsdata per område.....	15
5.3	Beskrivelse av ALS-data	19
6	Metode.....	20
6.1	Regresjonsanalyse	20
6.2	Prediksjonsmodeller	20
6.2.1	Multiplikativ prediksjonsmodell	20
6.2.2	Beregninger av statistiske parametre fra laserdata	21
6.3	Dummyanalyser.....	22
6.4	Variabelutvelgelse	23
6.4.1	Statistiske utvelgelsesprosedyrer.....	24
6.4.2	Statistiske kriterier.....	24
6.4.3	Multikollinearitet.....	25
6.4.4	Determinasjonskoeffisienten (R^2)	26
6.5	Modellbaser	27

6.6	Estimering av volum til valideringsmaterialet.....	29
6.6.1	RMSE (root mean square error)	30
6.6.2	Bias (gjennomsnittlig differanse mellom predikert og observert volum)	30
7	Resultater.....	31
7.1	Ulike regresjonsmodeller.....	31
7.1.1	Modell for alle fire områder	31
7.1.2	Modell for hvert område uavhengig av stratum	31
7.1.3	Modell for hvert stratum uavhengig av område	32
7.1.4	Modell for hvert område per stratum	32
7.2	Validering av de ulike modellene	33
7.3	Dummyanalyser.....	35
7.4	Regresjonsmodeller for ulik treslagsandel	36
7.4.1	Modell per område uavhengig av stratum for treslagsandeler 0,5 og 0,8	36
7.4.2	Modell per stratum uavhengig av område for treslagsandeler 0,5 og 0,8	38
7.5	Validering av modeller basert på forskjellige treslagsfordelinger.....	39
8	Diskusjon.....	42
8.1	Ulike regresjonsmodeller.....	42
8.1.1	Modell for alle fire områder	43
8.1.2	Modell for hvert område uavhengig av stratum	44
8.1.3	Modell for hvert stratum uavhengig av område	45
8.1.4	Modell for hvert område per stratum	46
8.2	Dummy analyser.....	47
8.3	Generell effekt av treslagsandeler på regresjonsmodeller	50
9	Konklusjon	52
10	Referanser.....	53
11	Vedlegg	55

1 INNLEDNING

En god og rasjonell skogforvaltning avhenger i stor grad av hvor god informasjon vi har om ressursene. Hvis vi har god oversikt over mengde, egenskaper og lokalisering av de ulike ressursene, som i denne sammenheng stort sett er definert som trær, vil det være lettere å ta de rette beslutningene i forhold til å nå de mål vi har satt for skogbruksvirksomheten vår. Ressursoversikter er derfor et nødvendig hjelpemiddel. Slike ressursoversikter knyttet til skogbruksvirksomhet kalles skogbruksplaner, og som navnet indikerer er dette et hjelpemiddel knyttet til planleggingen av skogbruksvirksomheten. Tidshorisonten på slike planer kan være både kortsiktig og langsiktig, men planen oppdateres gjerne hvert 10 år.

En ideell skogbruksplan har nøyaktig informasjon og er produsert til lave kostnader. Nøyaktigheten til informasjonen i en skogbruksplan avhenger av mange faktorer og en kan grovt sett dele disse faktorene inn i to grupper: 1) innsamling og 2) bearbeiding av data. En effektiv datainnsamling, eller takst, er avhengig av god planlegging på forhånd. All informasjon er for eksempel ikke like viktig for alle typer skog, så bruk av forhåndsinformasjon som flybilder og eldre skogbruksplaner for å planlegge taksen kan være viktig. Informasjon for å bestemme riktig hogsttidspunkt vil for eksempel være viktigere i eldre skog enn i ung skog. Ulike målsettinger med skogsdriften vil også kunne påvirke hvordan taksten gjennomføres. En skogeier som ønsker å maksimere sine inntekter fra skogsdriften vil for eksempel ha andre informasjonsbehov enn en skogeier som ønsker å maksimere biologisk mangfold. Det vil være sentralt å velge en takstmetode som er effektiv. Skal taksten gjennomføres på et lite areal kan for eksempel en gitt takstmetode være det mest effektive, mens andre metoder kan være best dersom arealet som skal takseres er stort. Når det gjelder bearbeidingen av data vil det for eksempel være viktig å velge hensiktsmessige statistiske metoder og å dele datamaterialet inn i formålstjenlige enheter etter skogtype. En slik inndeling av datamaterialet kalles stratifisering.

Skogbruket er en arealkrevende næring og arealene er ofte svært heterogene, noe som medfører utfordringer som er spesielle for denne næringen. Skogen kan anses som en populasjon av mange enheter som har utstrekning over store arealer. Det vil være umulig å måle alle enheter (trær), og derfor er det nødvendig å måle kun et utvalg av enhetene fra populasjonen og basere våre konklusjoner på observasjoner fra dette utvalget. En slik tilnærming kalles sampling. Samplingen avhenger av hvilken informasjon vi vil fram til som igjen avhenger av hva slags populasjon vi har, dvs. om populasjonens variabilitet. Populasjonen må være klart definert og avgrenset. Beskrivelsen av populasjonen avhenger også av godt definerte parametere som karakteriserer og beskriver populasjonen på en god måte.

For planmessing utnytting av skogarealene er det rasjonelt med en inndeling av skogen i homogene behandlingsenheter, eller bestand. Et bestand er definert som et areal hvor skogen er ensartet. Dette gjelder både bonitet, alder og treslag. Dominerende metode for bestandsinndeling i planfirmaene i dag er bruk av flybilder, der vi får automatisk bestandsareal ved at stereoinstrument er koblet direkte til et digitalt kartsystem (GIS).

De mest vanlige skogtakseringsmetodene i Norge de senere år har vært fototakst med markkontroll og relaskopstakst. I løpet av det siste tiåret har imidlertid lasertakst befestet seg som den klart dominerende metoden som i dag brukes operativt av takstfirmaene. Forutsetningen for at bruk av laserskanningsdata skal være effektivt for skogtakseringsformål er at det finnes sterke statistiske sammenhenger mellom de skoglige parameterne vi er interessert i å måle (f.eks. volum) og laserdataene. Denne metoden har vist seg å være meget lovende med hensyn til nøyaktighet og kostnadseffektivitet. Likevel er det trolig fortsatt potensial for ytterligere reduksjon av kostnader samtidig som man opprettholder eller øker nøyaktigheten. Bedre nøyaktighet og kostnadsreduksjon i innsamling av informasjon gir økt netto nytte, som vil være en ideell situasjon.

Oppgaven er bygd opp ved først å finne statistiske sammenhenger mellom skoglige parameter og laserdataene for så å utvikle ulike regresjonsmodeller. Metoden baserer seg på den arealbaserte metode som er operativt tatt i bruk av noen takstfirmaer i dag, også kalt for to-steps prosedyre. Denne er utviklet av, *Næsset & Bjerknes (2001)* og *Næsset (2002)*. Det er utviklet flere modeller på grunnlag av sammenslåing av ulike prøveflatedata og laserdata fra fire forskjellige områder. Siden stående volum trolig er den viktigste skoglige variabel for en skogeier både når det gjelder skogtilstand, fremtidig investeringsbehov og økonomisk nytte, har jeg bestemt meg for bare å se på volum i denne oppgaven. Alle modeller er utviklet for estimering av volum (m^3/ha).

Det er brukt separate datasett til modellutvikling og validering. For modellutvikling er det brukt datasett til prøveflatetakst og til validering av modeller er det brukt datasett til kontrollflatetakst.

2 FLYBÅREN LASER

Airborne Laser Scanning (ALS) er måling som foregår ved hjelp av en aktiv sensor, dvs. en sensor som sender ut lys. Sensoren kan benyttes uavhengig av lysforhold, men er følsom for regn eller skyer. Teknologien er kjent som LiDAR (*Light Detection And Ranging*) og er uttrykket til avstandsmåling (Wehr & Lohr, 1999).

Bruken av LiDAR for flybåren topografisk overvåking begynte sent på 1970 tallet. Svakheten med de tidlige systemene var dårlig bestemmelse av flyets posisjon og orientering (Rayner og O'Neil et al., 1979).

Laser ble tatt i bruk i skogtakseringsforemål rundt 1980 tallet (Solodukhin et al., 1977; Nelson et al., 1984; Aldred and Bonnor, 1985; Maclean and Krabill, 1986). Det ble da brukt profilerende laser for å måle trehøyder, bestandstetthet, treslag og biomasse.

Tidlig på 1990-tallet skjedde en forbedring i navigasjonsteknologi, elektronisk miniatyrisering og metoder som førte til sterk praktisk utvikling av ALS-systemer og kommersialisering. I dag er laserskanning fullt operativt og i regulær bruk til mange formål.

De primære drivkreftene for satsingen på laserskanning er mange. Effektiv og nøyaktig metode for høydekartlegging, direkte anskaffelse av 3D informasjon av objekter (vegetasjon) og rask leveranse av laserdata er noen som kan nevnes.

2.1 ALS-KOMPONENTER

Moderne ALS-systemer inneholder følgende komponenter:

- Laserskanner, der måleprinsipper kan kort beskrives som at kontinuerlige strømmer av laserpulser blir sendt ut mot bakken (figur 2).
- *Global Positioning System* (GPS), er et system som muliggjør nøyaktig posisjonering. Selve posisjoneringen krever at man har minst en mottakerenhet som kan motta og tolke signalene satellittene sender ut. Til skogtakseringsforemål trengs det to GPS mottakere, en på flyet og en på bakken.
- *Inertial Navigation System* (INS) som inkluderer *Inertial Measuring Unit* (IMU) for orienteringsmålinger, det vil si vinklene flyet har i forhold til aksene i et koordinatsystem.

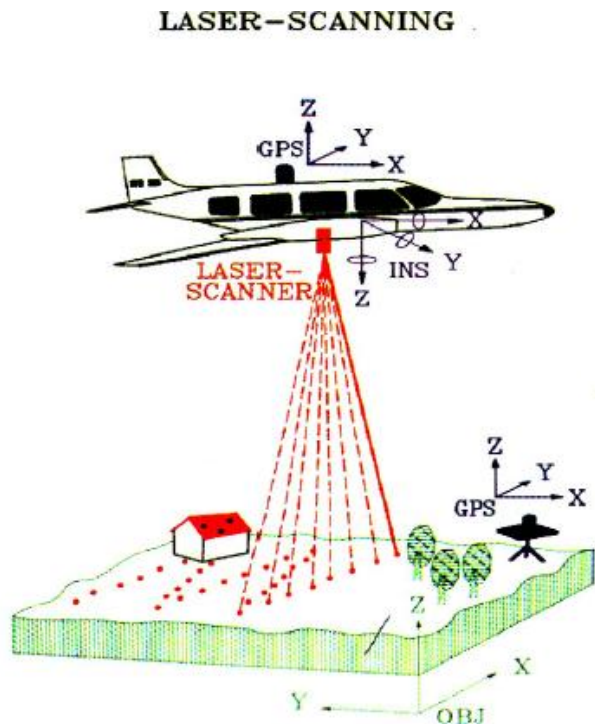
I tillegg til dette er det en kompleks softwareenhet som skal kontrollere alle disse komponentene, synkronisere tiden av målinger og lagre innsamlede data.

INS og laserskanner brukes til å få geo-refererte koordinater, dvs. posisjonbestemmelse av innsamlede laserdata. Sammenfattende innsamling og prosessering av IMU og GPS data er ofte integrert i en løsning som kalles POS (*Position and Orientation System*). Den andre GPS-stasjonen (referansemottakeren) er stilt i kjent bakkeposisjon og dens funksjon er å motta differensiell korleksjon for bedre posisjonering av flybane.

Etter flygningen blir flyets og bakkestasjonens GPS-data prosessert for å bestemme nøyaktig flybane. Disse dataene blir kombinert med laserregistrerte data, skannervinkel og IMU-målinger for bestemmelsen av nøyaktige horisontale og vertikale koordinater til hver enkelt puls som ble sendt ut fra skanneren. Koordinatene blir beregnet med interpolering i tid. For å transformere laserpunkter til koordinatsystem gjøres det bindinger mellom GPS og ønsket koordinatsystem WGS84(EUREF89), figur 1.

Så ut i fra dette beregnes det to viktige elementer:

1. Avstanden fra flyet ned til reflekterte objekt, som beregnes ut i fra medgått tid mellom utsending av laserpuls til en refleksjon blir oppfattet for hver enkelt puls.
2. Posisjonen i koordinatsystemet til laserpulsens som traff et objekt. Disse beregningene gjøres ut i fra flyets posisjon og orientering (vinkler) i koordinatsystem, og skannevinkelen.



FIGUR 1: ALS KOMPONENTER OG MÅLEPRINSIPPER SOURCE: TOPSCAN, GERMANY

Det er nødvendig med kalibrering av laseren hver gang den blir montert i flyet. Kalibreringer gjøres mot for eksempel bygninger som er kartlagt på forhånd og blir flydd over i forskjellige mønstre.

2.2 LASER PULSER OG ALS-TYPER

Laserlyset sendes ut med gitt spredning fra sensoren og kan best beskrives som en lyskjegle. På bakken vil denne lyskjegle ha et visst areal og er såkalt fotavtrykk (*footprint*). Fotavtrykk størrelse på bakken avhenger av åpningsvinkel og flyhøyde.

Avhengig av fotavtrykkets størrelse på bakken deles ALS systemer og laser pulsene i:

- Lite fotavtrykk (*Small footprint*), har diameter på bakken $< 1\text{m}$.
- Stort fotavtrykk (*Large footprint*), har diameter på bakken $> 5\text{m}$.

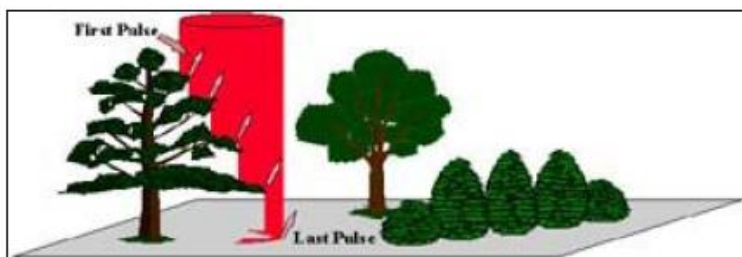
Et lite fotavtrykk er mest egnet til kartlegging av skogressurser. Det er fordi det har evnen til å registrere mer variasjon. Det er også fordi at vi har større punkttetthet, og flere pulsreturer enn ved et stort fotavtrykk.

Det finnes to typer ALS-systemer: Kontinuerlig bølgeform (*continuous waveform*) som registrerer hele fordelingen av retur energien fra der pulsen treffer et objekt, noe som krever mye lagringskapasitet, og *discrete pulse* systemer som registrer et avgrensa antall pulsreturer, figur 2. Til skogbrukelige anvendelser blir det ofte registret første og siste retur for hver puls. Dermed oppnås to datasett, sistepulsdatasett og førstepulsdatasett, figur 2.

Sistereturdatasett brukes til konstruksjon av digital terrengmodell (DTM) som er resultat av triangulering dvs. trekantmodell (TIN). Det brukes spesielle filtreringsalgoritmer som filtrerer ut vegetasjon og bygninger og som glatter terrenget. Denne filtreringen påvirker nøyaktigheten til DTM noe, spesielt i bratte og ujevne terreng og der evnen til å trenge gjennom vegetasjonen (penetreringsgraden) er lav. Punkttetthet, flyhøyde, og skanningsvinkel har også innflytelse på kvaliteten av DTM. For brukeren som anskaffer seg data til DTM er det viktig å kjenne dato, pulstetthet, flygningshøyde og skannevinkel. Disse bakkemodellene ses på som bakkessanne, men en må ta i betraktning mulige feil som kan overføres videre til vegetasjonsberegningene.

Bestemmelse av bakkenivå er essensielt i skogbrukens anvendelser av ALS. Første steget er nesten alltid å bygge terrengmodell og å subtrahere laserpulshøyder for å beregne høyder over bakkenivå. Styrken av reflektert signal varierer med objektenes reflekterende egenskaper og penetreringsgraden. Ikke alle pulser gir multiple returer. Et eksempel på dette er i tett skog der det er lav penetrasjonsgrad gjennom kroner og hvor vi ikke får registrert nok antall siste returpulser, vis man tenker seg at første retur utgjør pulsen som treffer kronen på et tre og siste retur når den treffer bakken. Også ujevnt terreng kan gi feil i DTM (*Kraus & Pfeifer, 1998*).

Økning i pulsfrekvens (repetisjonsrate) uten økning i fotavtrykkstørrelse vil gi flere første returer på hver prøveflate, ved i tillegg å øke samplingsarealet ved å fly flere striper vil dette forbedre målingene og laserestimer i skog.



FIGUR 2: FØRSTE OG SISTE RETUR REGISTRERINGER, SOURCE: TOPSCAN, GERMANY

Det finnes både skannende og profilerende (*profiling*) laser. Skannere samler inn data langs en korridor på bakken siden laserpulsene blir fordelt i en viss bredde som avhenger av skannevinkel og fly høyde, dvs. sprer laserpulsene på tvers av flyets retning. Laserpulsene blir spredd ut til et mønster av punkt ved hjelp av speil i sveipemekanismen.

Innsamling av data i profilerings laser skjer i en rett strekning som har bredde lik fotavtrykket av laserpulsene. Profilerende laser er typisk brukt for å dekke store områder.

Til operasjonell (kommersiell) lasertakst i Norge brukes det som oftest skannende laser med *discrete pulse systemer* med fotavtrykkstørrelse mellom 20 til 30 centimeter som oppnås med skanningsvinkler og flyhøyder som resulterer med pulstetthet fra < 1 til 5 pulser per kvadratmeter (Næsset 2002, 2004a; Bollandsås & Næsset, 2007).

Vertikal nøyaktighet avhenger av bakkekonfigurasjonen, målt horisontal nøyaktighet og flyhøyden. Med lavere flyhøyde kan vi oppnå bedre nøyaktighet.

3 OPERASJONELL (KOMMERSIELL) LASERTAKST TIL SKOGBRUKSFORMÅL

Laserskanning har vært i operasjonell bruk i skogbruket siden 2002 (*Næsset 2004, Næsset, 2007*). Det er utviklet to laserskanningsmetoder for å få informasjon om egenskaper ved skog: arealbasert metode og enkelttremetoden. En oppsummering av de to metodene finnes i *Hyyppä et al. (2006)*.

3.1 DE TO ULIKE METODER AV LASERTAKST

I arealbasert metode utledes alle skoglige variable per arealenhet. Denne metoden er brukt i oppgaven. Metoden er utviklet til operativt bruk av *Næsset (2002)* som en prosedyre i to steg. Prosedyren går ut på innsamling av data fra et visst antall koordinatfestete prøveflater fordelt på ulike strata. Variable som måles kan være brysthøyde diameter (d_{bh}), treantall, trehøyde osv. Prøveflatene blir lagt ut systematisk innenfor det laserskannede området. Deretter kobles disse prøveflatene sammen med laservariablene for å utlede et empirisk forhold. De utviklede modellene benyttes til prediksjon av ønskede skogvariable for hele skogarealet dekket av laser. Dette skogarealet blir først delt inn i celler (grid) som har likt areal som prøveflaten og som er lagt ut over hele området. Regresjonsanalyse er den mest vanlige estimeringsteknikken brukt i arealbasert metode, men det finnes også andre metoder, som for eksempel ikke-parametrisk tilnærming. Fordeler med arealbasert metode er at den er lett integrerbar med vanlig skogtaksering, metoden er godt utprøvd og dokumentert for beregninger av skogvariable, og laserskanningsdataene er billigere på grunn av lavere oppløsning enn enkelttremetoden.

I enkelttremetoden er hovedenheten et enkelttre. Alle ønskede skoglige variable utledes ut i fra innsamlet informasjon om det enkelte tre slik at bestandsdata fremkommer som summen av alle individuelle trær innenfor et gitt areal. Metoden trenger mer informasjon om fysiske egenskaper og sammenhenger hos det enkelte tre, som for eksempel forholda mellom kronevolum og kronediameter, for å skille ut enkelte trea. Denne type av klassifikasjoner kan baseres på kombinasjon av digitale bilder og laser. Etableringen av bestandsgrenser gjøres med utgangspunkt i bonitets- og høydevariasjon samt treslag og tetthet. Ulempen med denne metoden er høyere kostnader på grunn av kravet til høyere punkttetthet og mer kompleks bearbeiding av data.

Arealbasert metoden trenger prøveflater som er posisjonsbestemte for kalibrering av bakkesanne parametere å kunne utvikle regresjons modeller. Målet med GPS målinger er å posisjonsbestemme prøveflatesentrum slik at innsamlede data fra flaten kan posisjoneres i forhold til laserdataene. Derfor er det viktig med nøyaktig posisjonering av prøveflater. Nøyaktigheten av GPS-målingene avhenger av blant annet av kvaliteten på mottakeren, signalstyrken, fysiske hindringer og måleteknikk.

Differensiell Global Positioning System (GPS) er brukt til å bestemme posisjonen av flatenes sentrum. For å kunne bestemme posisjonen til flaten i (x, y) må det være tilgjengelig minst 4 satellitter, og det blir målt en avstand fra satellittene til mottakeren.

Kostnadene knyttet til laser avgjøres i første rekke av kravet til nøyaktigheten i sluttproduktet men også takstområdets størrelse, det vil si antall striper og stripekilometer, har noe å sei.

3.2 STRATIFISERING

Målet med stratifisering er å eliminere den delen av variabiliteten som er mellom strata og som kan være betydelig. Stratifisering har bare hensikt hvis variabiliteten mellom stratum er stor. Med mindre variabilitet innenfor stratum trenger vi mindre feltobservasjoner (prøveflater), og for en gitt nøyaktighet blir dermed kostnadene mindre.

Bestandsinndeling kan gjennomføres på grunnlag av rekke forskjellige egenskaper, for eksempel produksjonsforhold, jordbunnsforhold, geologi, vegetasjonsforhold osv.. Vanlige stratifiseringskriterier brukt i Norge er hogstklasse, bonitet (H_{40}), tetthet og treslag. Hogstklasse beskriver bestandets utviklingstrinn ut i fra alder og bonitet. Høydeboniteten (H_{40}) er definert som overhøyden (i meter) ved brysthøydealder 40 år. Bonitetskurver brukes til å bestemme høydeboniteten (H_{40}) (*Tveite, 1977; Braastad 1980*).

Inndelingen av antall strata kan skje til en viss grense, dvs. inndeling av populasjonen i flere strata vil ikke påvirke minskning av variasjonen innenfor strata. Det vil også ha påvirkning av antall flater per stratum, få flater i hvert stratum gir stor usikkerhet.

For å få ”optimal” antall av strata må vi ha god informasjon om strukturen til populasjonen/sampel til de skoglige variable som vi undersøker. Fordelingen av prøveflater til hvert enkelt stratum, dvs. fordeling av samplingsenheter, kan skje på to måter:

- Proporsjonal allokering, der samplingsenheter fordeles proporsjonelt med størrelsen til stratum.
- Optimal allokering, slik at antall av feilen til samplingsenheter blir minst mulig med hensyn til variasjonen innen de enkelte strata.

4 MÅLSETTING

Hovedmålet med denne oppgave var å se om man kan effektivisere feltdatainnsamlingen ved å benytte data fra ulike områder og stratifisere på alternative måter ved lasertakst. Dataene som er benyttet stammer fra fire forskjellige områder som er lasertaksert tidligere. Det ble utviklet ulike modeller som skal undersøke forholdet mellom laserregistreringer og feltmålt volum mellom områder. Jeg har videre slått sammen datasettene og delt materiale inn i ulike grupper eller strata for å undersøke effekten av å øke prøveflateantallet ved utvikling av forskjellige modeller, derfor er det også utviklet separate modeller for hvert enkelt område per stratum til å illustrere denne effekten.

Det blei foretatt også en stratifisering på bakgrunn av treslagsandelen i hvert område for å se om det kan bidra til nøyaktigere estimering av volum. Ulike grenser for treslagsfordeling ble benyttet for å undersøke om mer homogene bestand med hensyn på treslag vil medføre mer nøyaktig estimert volum gitt en modellbase. Med modellbase menes ulike måter å inndelegge datasettet før selve modelleringen startet.

5 MATERIALE

Materiale består av innsamlede felt data fra prøveflater og kontrollflater samt laserdata fra flybårenlaserskanning. Analysene i denne studien er gjennomført på data fra operasjonelle områdetakster fra fire forskjellige områder.

Datasettene er tidligere benyttet i separate studier (*Næsset, 2002, 2004a, 2004b, 2007*) og samlet i en studie av *Næsset & Gobakken (2008)*. Den stratifisering som de benyttet vil videre i oppgaven bli referert som originalstratifisering.

5.1 STUDIEOMRÅDER

Alle fire studieområder ligger i den sørlige delen av Norge. Høyden over havet for disse fire områdene varierte fra 70 til 900 m.o.h. (Tabell 1).

TABELL 1: OMRÅDER OG DENS GEOGRAFISKE KARAKTERISTIKKER

Område	Geografiske koordinater	Høyde over havet (m)
Hole	60°1'N, 10°20'E	240-540
Våler	59°30'N, 10°55'E	70-120
Nordre Land	60°50'N, 10°85'E	140-900
Krødsherad	60°10'N, 10°85'E	130-660

Hole og Krødsherad ligger i Buskerud fylke som strekker seg fra kysten i sør til Handagervidda i nord og har store variasjoner i klima og topografi. Hole og Krødsherad har store variasjoner i landskapet med bratte skråninger og bratte fjell.

Våler ligger i Østfold fylke og klimaet blir ofte karakterisert som sub-oseanisk, det vil si en overgangstype mellom et typisk kystklima (oseanisk) og et innlandsklima (kontinentalt) med milde vintre og høy årlig nedbørsmengde. De ytre deler av fylket har et relativt mildt kystklima, mens de indre deler er mer preget av innlandsklima. Våler er ganske flat med svake skråninger.

Nordre Land ligger i Oppland fylke og har et kontinentalt klima som karakteriseres med kalde vintrer og tørr klima.

Det dominerende treslaget i dataene fra disse områdene var gran (*Picea abies* (L.) Karst.) og furu (*Pinus sylvestris* L.)

5.2 DATAINNSAMLING - PRØVEFLATEDATA

Datainnsamlingen av prøveflatetakst har foregått i hvert område for seg. Prøveflatene var lagt ut i systematisk nett og hadde et fast areal avhengig av det stratum de tilhørte. Flatene ble plottet på kart og oppsøkt i marka. Denne typen av takst er vanlig i kommersiell lasertakst basert på arealbasert metode. Stratifisering er gjort ut i fra informasjon om treslagsfordeling fra flybilder.

På grunnlag av bestandsinndeling ble prøveflatene i utgangspunktet delt inn i tre strata definert som:

- Stratum I: ungskog (Hogstklasse=III)
- Stratum II: eldre skog (Hogstklasse>III) på dårlig bonitet ($H_{40} \leq 11$)
- Stratum III: eldre skog (Hogstklasse>III) på god bonitet ($H_{40} > 11$)

Denne stratifiseringen vil videre bli referert som original stratifisering.

Prøveflatene hadde sirkulær form og arealet varierte fra 200m² til 400m² avhengig av området (Tabell 2). På prøveflatene blir det registrert sedvanlige skoglige variable. Det vil si at det ble registrert diameter i brysthøyde (d_{bh}). Registreringer av nedre diametergrense for klava trær i diameter i brysthøyde (d_{bh}) varierte fra $d_{bh} > 4$ cm og $d_{bh} > 10$ cm avhengig av område og stratum. Trehøyder ble målt til utvalgte prøvetrær utvalgt med relaskop. Vertex hypsometer ble brukt til å bestemme høyden på prøvetrærne. Registreringer ble gjort med tanke på nøyaktig estimering av middelhøyde, overhøyde, middeldiameter, grunnflate, treantall og volum. Alle prøveflater ble posisjonsbestemt ved differensiel GPS og GLONASS. Loggingen av data varte cirka 15 til 30 min for hver enkelt flate.

Volumet til hvert enkelt prøvetre ble beregnet med volumfunksjoner (*Braastad 1966, Brantseg 1967, Vestjordet 1967*) der inngangsvariablene var diameter og høyde. For resten av trærne ble høyden beregnet gjennom diameter og høyde forhold. Volumet per prøveflate var beregnet som summen av hvert enkelt tre volum for trær med $d_{bh} > 4$ cm og $d_{bh} > 10$ cm. Deretter er volumet beregnet om til volum m³/ha.

TABELL 2: BESKRIVELSE AV TAKSTOMRÅDENE

Område	Takst år	Størrelsen av taksten (km ²)	Antall prøveflater	Prøveflate størrelse (m ²)	Min. klavediameter d_{bh} (cm)
Hole	2005	65	115	250	10
Våler	1998	10	143	300 ^A , 400 ^B	4 ^A , 10 ^B
Nordre Land	2003-2004	49	265	250	4
Krødsherad	2001	65	116	232,9	10

^A - Ungskog

^B - Eldre skog

5.2.1 BESKRIVENDE STATISTIKK FRA PRØVEFLATENE PER OMRÅDE

I Tabell 3 er det gitt beskrivende statistikk i fra feltmålingene til prøveflatene. Til sammen er det registrert 639 prøveflater i disse fire områdene. For alle områder uten stratifisering er standardavvik på feltmålt volum (m^3/ha) 472,6. Uten stratifisering for områdene hver for seg hadde Krødsherad den største standardavvik på 459,6 (m^3/ha) og Hole hadde minste på 292 (m^3/ha).

Det er forholdsvis stor variasjon i volum (m^3/ha) innenfor definerte strata, kalt for originale strata. Krødsherad hadde den største standardavvik på volum i stratum I med 119,6 (m^3/ha). Hole hadde største standardavvik i stratum II med 108,7 (m^3/ha), mens i stratum III hadde Nordre Land det største standardavviket på 131,8 (m^3/ha).

TABELL 3: BESKRIVENDE STATISTIKK FOR OBSERVERT VOLUM FRA FELTMÅLINGENE ETTER FORSKJELLIG ORGANISERING AV FELTDATAMATERIALE

Område	Stratum	Volum (m^3/ha)				Antall prøveflater (n)
		Min.	Maks.	Gj.snitt	SD	
Alle områder		8,8	677,1	342,9	472,6	639
Hole		17	430	223,5	292	115
Våler		35,3	548,5	291,9	362,9	143
Nordre Land		8,8	623,7	316,2	434,8	265
Krødsherad		27	677	352	459,6	116
	I	8,8	623,8	316,3	434,9	204
	II	22,4	448	235,2	300,9	196
	III	67,8	677,1	372,4	430,8	239
Hole	I	14	230	104	52,9	35
	II	32	430	156,8	108,7	31
	III	95	419	248,8	71,2	49
Våler	I	35	442	170,6	86,6	56
	II	56	262	148,6	50,4	36
	III	88	549	264,4	100,7	51
Nordre Land	I	9	413	114,7	81,8	74
	II	22	415	183,5	89,5	91
	III	68	624	312,5	131,8	100
Krødsherad	I	27	624	173,5	119,6	39
	II	30	448	174,1	93	38
	III	78	677	323,8	124,9	39

Med utgangspunkt i den originale stratifisering er det her foretatt en ytterligere stratifisering i forhold til treslagsandeler. De originale bonitets- og hogstklassegrensene er altså beholdt, men i tillegg er ulike grader av treslagsdominans grunnlag for stratifisering.

Hvert av de originale strataene er dermed delt i tre på grunnlag av om det er grandominert, furudominert eller blandingsskog. Dette vil si at for eksempel grandominert er skog der volumandelen gran av totalt volum er større enn en gitt grense. Grenseverdiene 0,5 og 0,8 er benyttet i analysene.

- Gran: innslaget av gran på kontrollflaten er $\geq 0,5$ og 0,8 blir prøveflaten definert som gran.
- Furu: innslaget av furu på kontrollflaten er $\geq 0,5$ og 0,8 blir prøveflaten definert som furu.
- Blanding: alt annet som er $< 0,5$ og $0,8$ for gran og furu blir satt som blanding.

I Hole er volum (m^3/ha) for gran den samme for treslagsandel 0,5 og 0,8, for treslagsandelen 0,8 har vi en prøveflate som ble klassifisert som furu og resten blei blanding (Tabell 4).

Våler og Krødsherad har samme situasjon. Nordre Land har mindre gjennomsnittlig volum (m^3/ha) for treslagsandelen 0,8, dette kommer av at flere prøveflater er klassifisert som furu og blanding (Tabell 4).

I stratum I ser vi økning i gjennomsnittlig volum (m^3/ha) for treslagsandel 0,8. I stratum II og III er det i treslagsanden 0,8 økning i volum (m^3/ha) bare for furu og blanding, mens volum (m^3/ha) av gran er litt mindre (Tabell 4).

TABELL 4: PRØVEFLATEVOLUM VED FORSKJELLIG INNDELING ETTER TRESLAGSANDEL (0,5 OG 0,8), OMRÅDE OG STRATIFISERING

Område	Stratum	Treslag	Volum(m^3/ha)								Antall prøveflater (n)	
			Min.		Maks.		Gj.snitt		SD			
			0,5	0,8	0,5	0,8	0,5	0,8	0,5	0,8	0,5	0,8
Hole		Gran	17	17	430	430	223,5	223,5	292	292	107	89
		Furu	26	26	87	26	56,5	26	43,1		4	1
		Blanding	18	18	130	312	74	165	79,2	207,9	4	25
Våler		Gran	66	161	548,5	548,5	307,2	354,8	341,2	274	74	36
		Furu	35,3	38,8	308,5	230,3	171,9	134,6	193,2	135,4	56	20
		Blanding	90,8	35,3	411,9	452,6	251,3	244	227,1	295,1	13	87
Nordre Land		Gran	8,8	8,8	623,7	619,2	316,2	314	434,8	431,6	205	171
		Furu	41,4	42,6	501,4	501,4	271,4	272	325,3	324,4	55	42
		Blanding	25,9	25,9	102,1	623,7	64	324,8	53,9	422,7	5	52
Krødsherad		Gran	35	35	623,8	623,8	329,4	329,4	416,3	416,3	55	36
		Furu	27,5	29,7	677,1	677,1	352,3	353,4	459,3	457,8	52	35
		Blanding	61	27,5	329	480,2	195	253,8	189,5	320,1	9	45
	I	Gran	8,8	8,8	623,8	623,8	316,3	316,3	434,9	434,9	153	104
		Furu	26	26	375	375	200,5	200,5	246,8	246,8	33	10
		Blanding	18	18	241	480,2	129,5	249,1	157,7	326,8	18	90
	II	Gran	22,4	430	448	430	235,2	226,2	300,9	288,2	108	86
		Furu	29,7	29,7	363,5	363,5	196,6	196,6	236	236	83	58
		Blanding	90,8	35	170,1	448	130,4	241,5	56,1	292	5	52
	III	Gran	79,5	79,5	623,7	619,2	351,6	349,4	384,8	381,6	180	142
		Furu	87,5	87,5	677,1	677,1	382,3	382,3	416,9	416,9	51	30
		Blanding	67,8	67,8	411,9	623,7	239,8	345,8	243,3	393,1	4	67

Tabell 5 viser prosentvis treslagsfordeling for hvert enkelt område per stratum. Gran er det dominerende treslag i Hole og Nordre Land i alle tre strata. I Våler og Krødsherad er gran dominerende i stratum I og III, mens i stratum II er furu dominerende.

TABELL 5: PROSENTVIS TRESLAGSFORDELING FOR PRØVEFLATENE PER OMRÅDE OG PER STRATUM

Område	Stratum	Gran %	Furu %	Lauv %
Hole	I	80,0	2,7	17,3
	II	82,4	9,5	8,1
	III	96,7	0,5	2,8
Våler	I	52,2	34,3	13,5
	II	30,7	64,5	4,8
	III	66,8	24,8	8,5
Nordre Land	I	81,3	8,5	10,2
	II	71,0	25,6	3,4
	III	70,4	27,0	2,6
Krødsherad	I	61,9	21,1	17,0
	II	19,6	75,4	5,0
	III	56,3	31,5	12,3

5.2.2 BESKRIVENDE STATISTIKK FRA VALIDERINGSDATA PER OMRÅDE

For å bestemme egnethet til modellene ble det registrert uavhengige data for hvert område. Disse dataene ble innsamlet på et større areal enn prøveflater og er omtalt som kontrollflater. Feltarbeidet og beregningen av volum ble utført på samme måte som for prøveflatene. Utgangspunktet for utvelgelsen av kontrollflatene er forskjellige for hvert enkelt område, det vil si at de er stratifisert på en annen måte i forhold til stratifiseringsgrunnlaget som jeg har benyttet i oppgaven. Stratifiseringen som jeg har foretatt blir kalt her for original stratifisering. Stratifisering av kontrollflatene er identisk til den gjort av *Næsset & Gobakken (2008)* (Tabell 7):

- Stratum I: ungskog (Hogstklasse=III)
- Stratum II: moden skog (Hogstklasse>III) på dårlig bonitet ($H_{40} \leq 11$)
- Stratum III: moden skog (Hogstklasse>III) på god bonitet ($H_{40} > 11$)

Det var forskjellig antall kontrollflater per område i forhold til publiserte artikler som har brukt det samme valideringsmaterialet, dette kommer av ulike stratifiseringsgrunnlag.

For å få en oversikt over ulike prosent treslagsandeler (Tabell 6) blei bonitetsklassene delt inn etter følgende kriterier:

- Lav bonitet: $H_{40} \leq 8$
- Middels bonitet: $11 \leq H_{40} \leq 14$
- Høy bonitet: $H_{40} \geq 17$

For alle områder går bonitetsintervallet fra bonitet 8 til bonitet 20.

Det er synlig i fra tabell 6 at gran er det dominerende treslag på gode boniteter, dette er fordi gran ofte er planta på gode boniteter.

TABELL 6: PROSENTVIS TRESLAGSFORDELING I FORHOLD TIL BONITET

Bonitet	Gran %	Furu %	Lauv %	Sum
Lav bonitet	3,6	7,2	1,0	11,8
Middels bonitet	35,7	21,9	5,1	62,7
Høy bonitet	20,3	3,0	2,2	25,5
Sum	59,6	32,1	8,3	100

- Hole:

Takseringen av prøveflatene og kontrollflatene ble utført sommeren 2005. Kontrollflatene hadde størrelse på 1000m^2 . Kontrollflatene ble delt i 4 kvadranter slik at størrelsen til hver kvadrant var 250m^2 som er likt prøveflatearealet.

Det er definert to strata for innsamlingen av data for kontrollflatene i følge *Næsset (2007)*:

- Stratum I: modne bestand med høy eller middels bonitet dominert av gran.
- Stratum II: modne bestand med lav bonitet dominert av gran og furu.

Etter original stratifisering er det til sammen 39 kontrollflater, i stratum III er det 24 og i stratum II er det 15 kontrollflater (Tabell 7). I kontrollflatedatasettet var det en kontrollflate som gikk inn som stratum I (kontrollflate nr 152). Denne kontrollflaten er ikke tatt med videre i valideringen.

- Våler:

Dataene fra Våler innbefattet til sammen 61 bestand som også er beskrevet i *Næsset (2002)*. Felt data var innsamlet sommeren 1998. Bestandsgrensene var skilt ut via fototolking slik at hver bestand representerer ulike kombinasjoner av aldersklasser, bonitetsklasser og treslagsfordeling. Stratum I har gjennomsnittlig bestandsareal på 2,46 ha., mens stratum II og stratum III har gjennomsnittlig bestandsareal på henholdsvis 6,26 ha. og 3,34 ha.. Felldata for kontrollbestand i Våler var innsamlet på flater som var systematisk fordelt i kvadratisk forband innenfor hvert bestand. I stratum I var størrelsen på flatene 100m^2 ($d_{bh}>4\text{cm}$), og i stratum II var størrelsen på flatene 200m^2 ($d_{bh}>10\text{cm}$). Flest flater var det i stratum I (Tabell 7).

- Nordre Land:

Datainnsamlingen på kontrollflatene fulgte samme opplegg som i Hole (*Næsset, 2004b*). Kontrollflatene har areal på 1000m^2 . I stratum A var antall kontroll flater 21 og i stratum B var antall kontroll flater 18 ifølge artikkelen *Næsset (2004b)*. Arealet til kontrollflatene ble delt i fire like deler slik av $\frac{1}{4}$ av kontrollflaten utgjør et areal på 250m^2 . Etter original stratifisering er det i stratum III 34 kontrollflater og i stratum II 5 (Tabell 7). Kontrollflate nr 423 er tatt ut fra analysen for den tilhørte stratum I ifølge original stratifisering.

- Krødsherad:

Det totale arealet til kontrollflaten er tilnærmet $61 \times 61\text{m}$ (3721m^2), dvs. arealstørrelsen på de 57 kontrollflatene varierte fra 3121m^2 til 4219m^2 med gjennomsnitt på 3736m^2 . Innsamling av felldata ble gjennomført i sommeren 2001.

Kontrollflatene var stratifiserte etter forhåndsdefinerte strata, *Næsset (2004a)*.

- Stratum I: ungskog (Hogstklasse=III)
- Stratum II: modensskog (Hogstklasse>III) på dårlig bonitet ($H_{40} \leq 11$)
- Stratum III: modensskog (Hogstklasse>III) på god bonitet ($H_{40} > 11$)

Denne stratifisering er lik originalstratifiseringen som jeg har brukt.

I stratum I var alle trær med $d_{bh}>4\text{cm}$ registrert, mens i stratum II og III var alle trær med $d_{bh}>10\text{cm}$ registrert.

TABELL 7: BESKRIVENDE STATISTIKK FOR OBSERVERT VOLUM FRA VALIDERINGSMATERIALET

Område	Treslagsandel	Stratum	Treslag	Antall kontroll flater	Volum(m^3/ha)		
					Min.	Maks.	Gj.snitt
Hole	Original stratifisering	II		15	168,9	314,8	233,2
		III		24	40,2	310,7	146
Våler	Original stratifisering	I		22	100,1	345,8	158,3
		II		19	90,8	257,6	137,2
		III		20	138,8	410,9	280,9
Nordre Land	Original stratifisering	II		5	68,1	244,7	156,4
		III		34	68,7	513,2	291
Krødsherad	Original stratifisering	I		18	64	329,6	212,5
		II		19	83	292,7	162,7
		III		19	186	378,9	287
Hole	0,5	II	Gran	14	40,2	310,7	149,3
			Furu	1	97,7		97,7
			Blanding	0			
		III	Gran	24	168,9	314,8	233,2
			Furu	0			
			Blanding	0			
Våler	0,5	I	Gran	9	130,2	345,8	197,4
			Furu	8	100,1	190,3	139,9
			Blanding	5	104,9	164,5	131,6
		II	Gran	2	187,7	188,7	188,1
			Furu	15	90,8	257,6	131,9
			Blanding	2	140,3	145,3	143,5
		III	Gran	18	148,7	410,9	288,6
			Furu	0			
			Blanding	2	138,8	283,4	211,7
Nordre Land	0,5	II	Gran	0			
			Furu	5	68	245	149,4
			Blanding	0			
		III	Gran	21	69	513	335,1
			Furu	13	143	374	255
			Blanding	0			
Krødsherad	0,5	I	Gran	10	95,8	329,6	244,4
			Furu	4	64	272,7	176,8
			Blanding	4	93,2	255,3	165
		II	Gran	6	104,4	292,7	219,3
			Furu	12	83	236,4	140,5
			Blanding	1	93,2		93,2
		III	Gran	13	186	378,9	299,9
			Furu	5	225,9	312,2	259,9
			Blanding	1	270,3		270,3

Hole	0,8	II	Gran	11	40,2	310,7	127,3
			Furu	0			
			Blanding	4	97,7	307,5	197
		III	Gran	20	168,9	314,8	233
			Furu	0			
			Blanding	4	170,7	294,6	234
Våler	0,8	I	Gran	2	319,4	345,8	332,2
			Furu	1	124,9		124,9
			Blanding	19	100,1	243,1	150,1
		II	Gran	0			
			Furu	4	90,8	158,8	119,4
			Blanding	15	104	257,6	142,3
		III	Gran	9	234,2	410,9	317,3
			Furu	0			
			Blanding	11	138,8	373,4	247,7
Nordre Land	0,8	II	Gran	0			
			Furu	2	68	145	103
			Blanding	2	193	245	219
		III	Gran	16	69	513	353,9
			Furu	8	143	362	239
			Blanding	10	157	374	277,6
Krødsherad	0,8	I	Gran	6	244,7	329,6	295,8
			Furu	2	152,7	272,7	222
			Blanding	10	64	255,3	162
		II	Gran	3	231,2	292,7	254,2
			Furu	8	83	195,1	133,8
			Blanding	8	93,2	288,2	158,2
		III	Gran	11	190,4	378,9	314,8
			Furu	4	234,8	312,2	268,4
			Blanding	4	186	270,3	226,3

5.3 BESKRIVELSE AV ALS-DATA

Instrumentene ALTM 1210 og ALTM 1233 er produsert av Optech Inc.. Innsamling av laserdata har foregått i sommersesong ("blad - på" vegetasjonsperioden) i perioden fra juni 1999 til september 2004 (Tabell 8).

Gjennomsnittlig flyhøyde varierte fra 650 til 1200 m.o.h., gjennomsnittlig fotavtrykk varierte fra 20 til 36 cm, og gjennomsnittlig punkttetthet varierte mellom 0,8 til 1,1 m⁻². Maksimalt prosessert skanningsvinkel varierte fra 14 til 17 grader.

TABELL 8: ALS INSTRUMENT SPESIFIKASJONER

Område	Instrument	Gjennomsnittlig flyhøyde ^a	Gjennomsnittlig fotavtrykk (cm)	Gjennomsnittlig punkttetthet (m ⁻²)	Max prosesserte skanningsvinkel (deg)
Hole	ALTM 1233	1200	36	0,8	17
Våler	ALTM 1210	700	21	1,1	17
Nordre Land	ALTM 1233	800	24	1,1	14
Krødsherad	ALTM 1210	650	20	1	16

^a Meter over bakken

For å kunne konstruere terrengflaten må pulsene som ikke hører til terrengflaten filtreres vekk. Denne filtreringen skjer med bruk av dataprogram som bruker en eller flere algoritmer. Filtreringen utføres som oftest av leverandør som i dette tilfelle var Blom Geomatics, Norge. Pulsene var registrert i TIN i følge deres koordinater. Høydeverdi for første pulsretur var deretter beregnet for hvert enkelt punkt med lineærinterpolering fra TIN. Det er nødvendig å definere hvilke returer som kan klassifiseres som vegetasjonstreff for videre beregninger av skoglige variable. Derfor er det satt en grenseverdi for å avgjøre hva som er vegetasjonstreff. Returer som har høydeverdi ≥ 2 m ble betraktet som vegetasjonstreff (Tabell 9).

Fra tabell 9 ser vi at i Nordre Land er det en retur med ekstrem høydeverdi (44,3m). Denne returen tilhører prøveflate nr 401 og ble ikke tatt med videre i beregningene.

TABELL 9: FØRSTERETURER KLASSIFISERT SOM VEGETASJONSTREFF

Områder	Min. (m)	Maks.(m)	Gjennomsnitt(m)	S.D (m)
Hole	2,0	28,3	10,1	4,8
Våler	2,0	29,8	11,0	5,0
Nordre Land	2,0	44,3	11,2	5,5
Krødsherad	2,0	31,4	11,1	5,7

6 METODE

6.1 REGRESJONSANALYSE

Regresjonsanalyser er en statistisk analyseteknikk for å bestemme kvantitative sammenhenger mellom ulike variable, samt retningen og styrken på sammenhengen. En regresjon beskriver og evaluerer altså forholdet mellom en avhengig variabel (Y) og en eller flere forklarende variable (X), som benevnes uavhengige variable. Med regresjon prøver vi altså å forklare endringer i en variabel ut i fra endringer i en eller flere andre variable. For arealbasert metode er det mest vanlig å utlede skoglige parametre med nettopp regresjonsanalyse. For å koble sammen uavhengige variable (X) som er utledet fra ALS-datasett kreves det et visst antall prøveflater hvor det er registrert en egenskap som settes som den forklarte variabelen (Y). Så undersøkes det hvilke av ALS -variablene og deres samspill som best forklarer variasjonen i Y. I fra ALS-data kan vi utlede mange uavhengige variable og derfor blir det brukt multipl regressjonsanalyse.

6.2 PREDIKSJONSMODELLER

6.2.1 MULTIPLIKATIV PREDIKSJONSMODELL

Hvis uavhengige variable er uavhengige seg i mellom, og hver av dem har innvirkning på den avhengige variabelen, adderes innvirkningen og vi snakker om additiv innvirkning. Et eksempel på dette er bonitet og alder. Bonitet er uavhengig av alder og motsatt, men samtidig vil bonitet og alder ha innvirkning på volumet som den avhengige variabelen.

Hvis de uavhengige variablene er korrelerte seg i mellom seg, det vil si at den ene variabelen til en viss grad kan forklare og erstatte den andre. Et eksempel på dette er diameteren og høyden på tre, jo større diameter jo høyere er trærne og omvendt, begge er sterkt korrelerte med hverandre, men gir forklaring på variasjonen på volumet og kan dermed inntil en viss grad erstatte hverandre. I praksis blir vi for det meste møtt med substantiv innvirkning av variablene, men de blir behandlet som additive på grunn av enklere tolkning.

Ofte er det slik at det trengs flere variable som kan forklare den avhengige variabelen. En ønsker derfor å se på alle de forklarende variablene sammen og vurdere den totale effekten disse har på den avhengige.

Dette er multiple regresjonsmodeller. Den generelle formen for eksponentielle multiplikative modeller er uttrykt som:

$$Y_i = b_0 X_{i_1}^{b_1} X_{i_2}^{b_2} \dots X_{i_p}^{b_p} \varepsilon_i; \quad i=1,2,\dots, n \quad (1.1)$$

b 'er = ukjente parametre

ε = er tilfeldig feil (residual) som representerer restvariasjon. Vanligvis antatt til å være normalfordelt med gjennomsnitt null og variansen σ^2 .

$$\varepsilon \sim N(0, \sigma^2)$$

Y = avhengig variabel eller responsvariabel

X = uavhengige variable eller forklaringsvariable

Når vi plotter avhengig variabel (Y) mot uavhengig variabel (X) blir ikke alltid kurven rettlinjert (lineær). Dette forholdet kan være kurvet (graf 1, Vedlegg). Fordi det vil være matematisk vanskelig å finne parametrene ($b_0, b_1, \dots, b_p$) til funksjonen 1.1, transformeres modellen til lineær form. I praksis er det som oftest brukt logaritmisk transformering, men det kan også brukes transformeringer som $Y = \sqrt{N}$ eller $Y = \frac{1}{x}$. For å få lineær kurve blir det gjort transformering av X og Y i ln form.

$$\ln Y_i = \ln b_0 + b_1 \ln X_{i_1} + b_2 \ln X_{i_2} + \dots + b_p \ln X_{i_p} + \ln \varepsilon_i \quad (1.2)$$

6.2.2 BEREKNINGER AV STATISTISKE PARAMETRE FRA LASERDATA

For bestandsvariablene er informasjonen om høyde og tetthet essensielt. Kort kan dette forholdet beskrives som:

$$V\left(m^3/ha\right) = f\{\text{laserutleda variable}\}$$

I dette arbeidet er det beregnet både høyde og tetthets laservariable fra første returdatasett. Beregningene er basert på arbeidet til Næsset (2002, 2004a) og Næsset & Gobakken (2008).

- Høyderelaterte variable:

Først ble alle laserpulser $< 2m$. filtrert bort. Grunnen til dette er å utelukke refleksjoner fra bakke, steiner, greiner osv. Alle laservariable ble utledet per prøveflate og de er:

- Gjennomsnitt (Dz_{avg}): Gjennomsnittlig verdi av høyder.
- Standartavviket (Dz_{sd}): Gir et mål på dispersjon av målte høyder. Den blir utregnet som:

$$Dz_{sd} = \frac{1}{n} \sum_{i=1}^n (Dz - Dz_{avg}) \quad (1.3)$$

- Variasjonskoeffisienten (Dz_{cv}) beregnes som:

$$Dz_{cv} = \frac{Dz_{sd}}{Dz_{avg}} * 100 \quad (1.4)$$

- Høydepersentiler ($P_0, \dots, P_{90}$): Alle målte verdier var sortert etter høydeverdier og delt i 10 persentiler. Denne variabelen viser til høydeposisjonen av målte verdi.

- Tetthetsrelaterte variable:

Tetthetsrelaterte variable gir uttrykk på ulike penetrasjonsgrad. Det er beregnet tetthetsvariabel som gir en vertikal beskrivelse av høyde lag til laserpulsene.

$$F_0 = \frac{\left(n_{>(\frac{max}{10*1})}\right)}{N}, F_{10} = \frac{\left(n_{>(\frac{max}{10*2})}\right)}{N}, \dots, F_{90} = \frac{\left(n_{>(\frac{max}{10*9})}\right)}{N} \quad (1.5)$$

Laservariablene, forklaringsvariablene, som blir brukt til regresjonsanalysen er:

- Persentiler: $P_0, P_{10} \dots P_{90}$
- Fraksjoner: $F_0, F_{10} \dots F_{90}$
- Variasjonskoeffisienten: Dz_{cv}
- Gjennomsnittlig delta Z av vegetasjonstreff: Dz_{avg}

Til sammen er det 22 laservariable (*Næsset & Gobakken, 2008*) som er satt som uavhengige variable mot avhengig variabel (volum [$V(m^3/ha)$]). Den endelige multiplikative modellen ser slik ut:

$$Y = \beta_0 P_0^{\beta_1} * P_{10}^{\beta_2} * \dots P_{90}^{\beta_{10}} * F_0^{\beta_{11}} * F_{10}^{\beta_{12}} * \dots F_{90}^{\beta_{20}} * Dz_{cv}^{\beta_{21}} * Dz_{avg}^{\beta_{22}} \quad (1.6)$$

Her er det brukt logaritmisk transformering (grunntall e). Den transformerte multiplikative modellen har formen:

$$\ln Y = \ln \beta_0 + \beta_1 \ln P_0 + \beta_2 \ln P_{10} + \dots \beta_{10} \ln P_{90} + \beta_{11} \ln F_0 + \beta_{12} \ln F_{10} + \dots \beta_{20} \ln F_{90} + \beta_{21} \ln Dz_{cv} + \beta_{22} \ln Dz_{avg} + \varepsilon \quad (1.7)$$

Ved bruken av slik modell får vi estimerte verdier for den transformerte avhengige variabelen ($\ln Y$) og ikke til den originale avhengige variabelen (Y). Med antilogaritmen av estimerte verdier i $\ln$ form ($\ln \hat{Y}$) får vi estimerte verdier i original form ($\hat{Y}_i$). Denne fremgangsmåten blir ikke helt riktig fordi den underestimerer kostnadsleddet, så det legges til en korreksjonsfaktor knyttet til kostleddet β_0 .

$$\text{Korreksjonsfaktoren er: } (1/2 \sigma^2), \quad (1.8)$$

der σ^2 er den estimerte variansen til modellen. Korreksjonen i uttrykket (1.8) er anbefalt av *Flewelling & Pienaar (1981)* for moderate til store datasett.

6.3 DUMMYANALYSER

Med multiplikative regresjonsmodeller kan vi gjøre kvalitative eller kategoriske analyser. Dette gjøres med at vi setter inn dummyvariable som er definert som kvalitative uavhengige variable i modellen. De beskriver nærvær eller fravær av en kvalitet eller effekt. Kvalitative variable i motsetning til kvantitative variable, kan ikke måles på numerisk skala, derfor trenges det omkoding. Kodingen av dummyvariable er oftest i binærnivåer, dvs. enten 0 eller 1. Antall dummyvariable som skal innføres i regresjonsmodell er en dummy mindre enn antall kvaliteter. En generell formel kan skrives som $D_x = n - 1$, der D_x er antall dummy variable og n er antall kvaliteter for eksempel stratum.

6.4 VARIABELUTVELGELSE

I mange tilfeller vil de registrerte variablene i en tekst være mange. Spesielt i en lasertekst hvor en teoretisk sett kan utlede hundrevis av variable fra laserpunktene. Hver eneste av dem er en potensiell prediktor av den avhengige variabelen (Y). En modell som inneholder et veldig stort antall av uavhengige variable vil gjøre modellen upraktisk å bruke. Hvis vi har med alt for mange variable overtilpasser vi modellen. Konsekvensene av det er:

- Modellen er ikke god for prediksjon på nye data
- Modellen er for innviklet, modell "noise" (støy)

Også for få variable kan skape problemer gjennom at vi utelater variable som kan hjelpe med forklaringen av responsen (Y), da underpasser vi modellen. Noen av konsekvensene er:

- Modellen er ikke god for prediksjon av nye data
- Estimer av feilvariansen er for stor

Det er utviklet flere statistiske metoder (utvalgsprosedyrer) og kriterier som kan redusere antall variable for å unngå overpasning og underpasning. Målet er å forklare data på en enklest mulig måte. I statistikken er det en regel som gjelder når en velger mellom modeller med cirka den samme forklaringsgrad og det er alltid å velge modellen med færrest parametere.

I oppgaven ble det brukt både stegviseprosedyre og kriterier basert prosedyre for utvelgelse av variable.

6.4.1 STATISTISKE UTVELGELSESPROSEDYRER

Utvelgelsen av uavhengige variable og analyser er utført i R (*R Development Core Team 2007*).

Det er brukt ”*stepwise both directions*” regresjons prosedyre, som er en modifikasjon av ”*forward selection*” og ”*backward elimination*”. Det skal være et gitt signifikansnivå for at variablene blir selektert (*t*-verdier) i spesifisert α nivå. Stegvis regresjonsprosedyre kan brukes når vi skal bestemme hvilke av et stort potensielt antall av uavhengige variable som skal brukes i modelloppbygningsprosessen.

- *Forward selection*: begynner med inklusjon av en variabel, så legger man til den beste tilpassende variabelen, osv, til det ikke er muligheter for forbedringer.
- *Backward selection*: starter med den mest komplekse modellen og så reduserer suksessivt modellen.
- *Both directions*: er en kombinasjon av ”*backward*” og ”*forward*” strategi. Etter hver fjerning av en variabel, analyserer en om inkludering av tidligere fjernet variabel kan forbedre modellen.

6.4.2 STATISTISKE KRITERIER

Kriteriet brukt her får bestemme om en variabel skal være med er *Akaike Information Criterion* (AIC). AIC baserer seg på maksimum *likelihood* (sannsynlighetsmaksimering på norsk) som uttrykker sannsynligheten for å observere de data som faktisk er observert og benyttet i estimeringen, som funksjon av de ukjente parametrene i modellen (*Heldal, 2006*).

Målet er å finne en modell som best forklarer data med et minimum antall frie parametre. Å tilpasse en modell kan ses som optimering av to motstridende mål:

- Modelltilpasning, dvs. lage tilpasningen til dataene så god som mulig: *likelihood* skal være stor
- Modellkompleksitet, dvs. lage modellen enkel: bruke så få uavhengige variable som mulig (*p*)

AIC definerer mål som kombinerer begge disse aspekter. AIC er definert som -2 ganger $\log$ *likelihood* pluss 2 ganger antall av parametre. Modell med minimal AIC prøver å finne optimal kompromiss mellom modelltilpasning og modellkompleksitet.

For å velge ut den beste modellen vil en ideelt prøve ut modeller med alle mulige kombinasjoner av prediktorene.

$$AIC = 2\log L(b) + 2(p + 1) \quad (1.9)$$

L = maksimal *likelihood* funksjon

b = maksimumlikelihoodestimatet

p = antall parametre

Eller

$$AIC = n * (\ln((2 * \pi * RSS) / n) + 1) + 2 * p \quad (2.0)$$

RSS = residual sum of squares $\sum (\tilde{Y}_i - Y_i)^2$

p = antall parametre

n = antall observasjoner

Ved bruk av modellseleksjon bør en være klar over at AIC tenderer å foreslå for med for mange variable. I oppgaven ble utvelgelsen derfor ”kjørt” to ganger.

6.4.3 MULTIKOLLINEARITET

Korrelasjonen mellom to variable måler graden av lineær sammenheng de to variablene i mellom. Til forskjell fra regresjonsanalyse kan man her ikke komme frem hvilken variabel som er avhengig og hvilken som er uavhengig.

Dersom korrelasjonen mellom variablene ikke eksisterer vil det ikke ha betydning for de andre variablenes koeffisienter om man legger til eller trekker fra en variabel. Svak korrelasjon mellom variablene vil ha liten betydning i regresjonens presisjon. I de fleste tilfeller i praksis vil det eksistere korrelasjon mellom de forklarende variablene. Dersom denne korrelasjonen er sterk så eksisterer det multikollinearitet.

Multikollinearitet er hvis to eller flere variable er høyt korrelert slik at de måler samme fenomen (muligens på forskjellig måte), det vil si at en variabel fanger opp mye av effekten av en eller flere andre forklaringsvariable, og det ikke er mulig å skille ut den separate effekten av hver enkelt variabel.

Hvis vi ignorerer multikollinearitet vil R^2 bli høy, viktige variable kan se ut som ikke-signifikante og de individuelle koeffisientene vil ha høye standardfeil som vil resultere i store konfidensintervall for parameterne, slik at signifikanstester gir gal konklusjon og inferens. Det er best å fjerne overflødige variable for dermed å eliminere multikollinearitet.

Undersøking av korrelasjoner er det første steg ved en vurdering av multikollinearitet. Det ble deretter utført en prinsipal komponentanalyse basert på Pearson korrelasjons matrise for variablene som ble foreslått i modellen. Ut i fra matrisen ble egenverdier (λ_i) (*eigenvalues*) beregnet.

Condition number (κ) var beregnet som kvadratroten av største *eigenvalue* delt på minste *eigenvalue*.

$$\kappa = \sqrt{\frac{\lambda_{\max}}{\lambda_{\min}}} \quad (2.1)$$

Verdien til *condition number* (κ) var brukt som indikasjon på kollinearitet, Næsset (2002).

- $\kappa > 30 \Rightarrow$ *sterk grad av kollinearitet*
- $\kappa = 10 - 30 \Rightarrow$ *moderat grad av kollinearitet*
- $\kappa < 10 \Rightarrow$ *svært liten grad av kollinearitet*

6.4.4 DETERMINASJONSKOEFFISIENTEN (R^2)

Determinasjonskoeffisienten (R^2) er et mål på hvor godt modellen passer til datasettet som modellen er utviklet fra. For å kunne bruke determinasjonskoeffisienten som mål på hvor god modellen er brukte jeg alternativ R_a^2 (*adjusted multiple coefficient of determination*) som tar i betraktning antall observasjoner og antall av parametre i modellen og har den fordel at den justerer for kompleksitet. R_a^2 vil alltid være mindre enn R^2 . Grunnen til dette er at R_a^2 er en forventningsrett ”utgave” av determinasjonskoeffisienten.

En skal være varsom til å bedømme modelltilpasning bare på denne verdien.

$$R_a^2 = 1 - \left[\frac{(n-1)}{n-(k+1)} \right] (1 - R^2) \quad (2.2)$$

n= antall observasjoner

k= antall uavhengige variabler

6.5 MODELLBASER

Med modellbaser menes ulike måter å inndelegge datasettet før modelleringen startet. For original stratifisering er det utviklet 20 modeller til sammen.

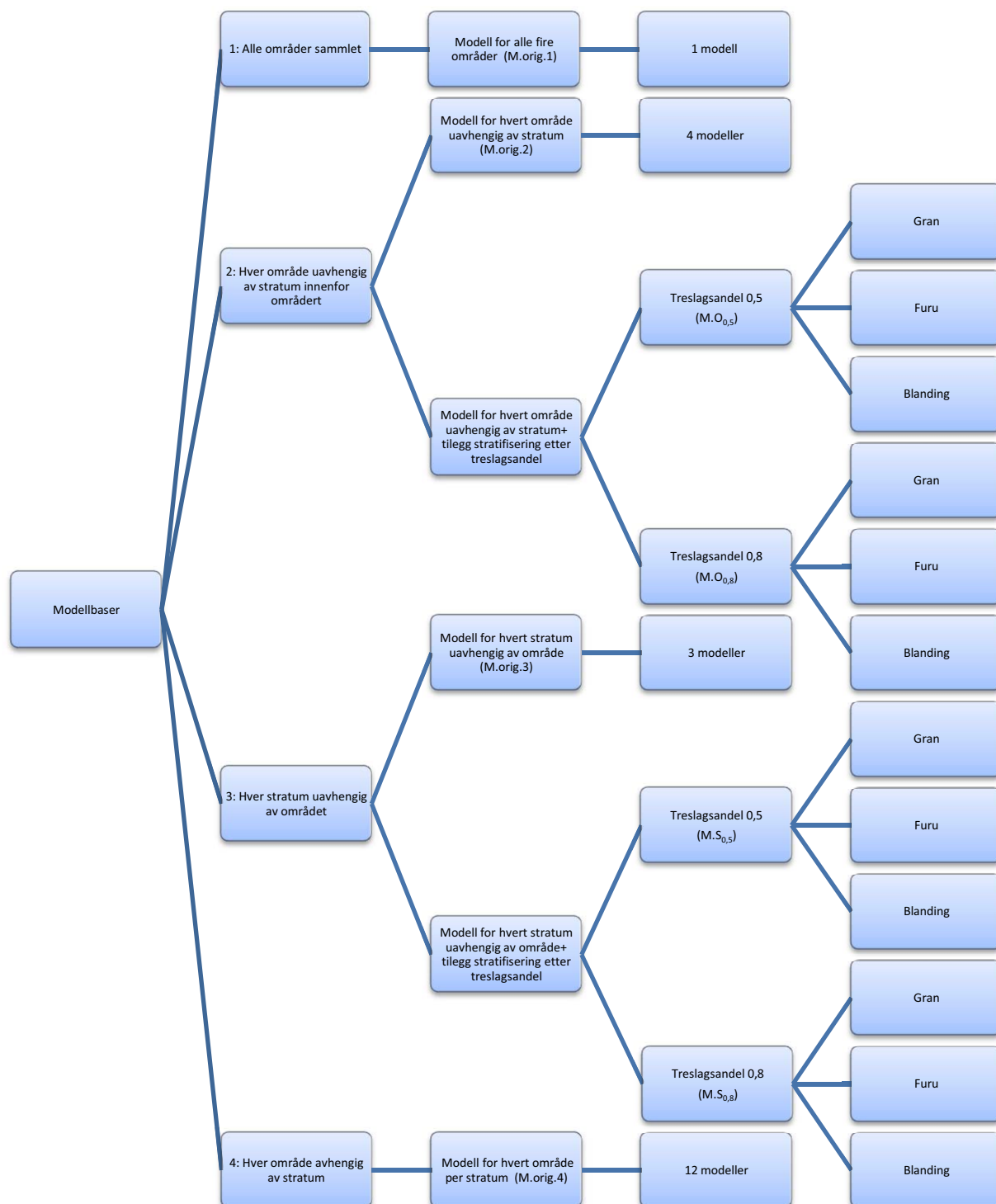
Følgende 4 modellbaser ble benyttet (figur 3):

- *M.orig.1* - Modell alle fire områder, ingen stratifiseringskriterier
- *M.orig.2* - Modell per område uavhengig av stratum, stratifiseringskriterie er modell på områdenivå
- *M.orig.3* - Modell per stratum uavhengig av område, stratifiseringskriterie er stratum
- *M.orig.4* - Modell per område per stratum, stratifiseringskriterie er område og stratum innenfor område

Der M.= modell; orig.= originalstratifisering; 1=modellbase nr. 1

I tillegg til modellbasene overfor blei det ytterligere stratifisert i forhold til treslagsandel. Dvs. treslagsandeler er definerte som relativ andel av gran, furu eller blanding på 0,5 (50 %) og 0,8 (80 %) i datasettet til hvert av områdene. Det er utviklet flere modeller for denne typen av stratifisering (figur 3).

For utvikling av alle modeller er det lagt vekt på å få så høy forklaringsgrad som mulig, med betingelser av at alvorlig multikollinearitet ikke skal forekomme og at residuaene har en fornuftig fordeling.



FIGUR 3: MODELL OPPBYGNING

M.= modell; O= område; S= stratum; 0,5= treslagsandel. Eksempel: M.O_{0,5-0,8}- modeller per område uavhengig av stratum for treslagsandel 0,5 og 0,8 og M.S_{0,5-0,8}- modeller per stratum uavhengig av området for treslagsandel 0,5 og 0,8.

6.6 ESTIMERING AV VOLUM TIL VALIDERINGSMATERIALET

Grunnlaget for den arealbaserte metode er å utlede ønskede skogvariablene per arealenhet. Egnethet til modellene ble validert mot det uavhengige valideringsmaterialet (kontrollflatene) fra de fire områdene. Kontrollflatene ble delt inn i kvadratiske celler med celledørrelse lik prøveflatearealet så langt dette lot seg gjøre. Deretter ble volum per hektar predikert for hver celle. Disse predikerte volumene for hver celle ble summert for hver enkelt kontrollflate ved å arealveie predikert volum for de enkelte celledelene som inngikk i hver kontrollflate.

6.6.1 RMSE (ROOT MEAN SQUARE ERROR)

RMSE er det *estimerte* standardavviket til feilene og gir et mål på nøyaktigheten til modellen.

$$RMSE = \sqrt{\frac{1}{n} \times \sum_{i=1}^n (E_i - O_i)^2} \quad (2.3)$$

n= antall observasjoner

O= observert volum pr. ha.

E= estimert volum pr. ha.

6.6.2 BIAS (GJENNOMSNITTLIG DIFFERANSE MELLOM PREDIKERT OG OBSERVERT VOLUM)

Middelverdier for differansen mellom predikert og observert volum (m³/ha) er regnet ut ved:

$$D_i = V_E - V_o \quad (2.4)$$

O= observert volum pr. ha.

E= estimert volum pr. ha.

Dette er uttrykk for *bias*.

$$\bar{D} = \frac{1}{n} \sum_{i=1}^n D_i \quad (2.5)$$

Dette viser om det er systematisk feil i differansen mellom predikert og observert volum.

Både *bias* og RMSE til stående volum ble brukt som kriterier i vurderingen av de ulike utviklede modellene.

7 RESULTATER

Det vil først bli presentert en modell der alle prøveflatene er benyttet til å lage en modell uten å ta hensyn til stratifisering og at dataene kom fra forskjellig område. Deretter vil separate modeller for de ulike områder og separate modeller for de ulike originale strata bli presentert. Valideringen av disse modellene vil så bli presentert. Utvikla modeller for ulik treslagsandel fordeling vil bli presentert i eget kapittel og resultater fra valideringen følger deretter. Alle resultater er presentert (beregnet) i m^3/ha .

7.1 ULIKE REGRESJONSMODELLER

For å få innsikt i hvor stor innvirkning de ulike strata og områder hadde på estimeringen av volum, ble det utviklet flere regresjonsmodeller avhengig av hvordan datamaterialet blir organisert. R_a^2 og RMSE er oppgitt på estimeringsform dvs. i ln-skala.

7.1.1 MODELL FOR ALLE FIRE OMRÅDER

Modellen for alle fire områder (graf 2, vedlegg) ble utviklet ut i fra alle prøveflatene samlet, uavhengig av hvilke strata og område de tilhørte. Antall prøveflater til utvikling av denne modellen (n) er 639 og $R_a^2 = 0,88$ (Tabell 10).

TABELL 10: MODELL FOR ALLE FIRE OMRÅDER

Variablene i modellen	Modelltype	R_a^2	RMSE	κ	n
$\ln P_{50}; \ln P_{90}; \ln F_0$	M.orig.1	0,88	0,23	5,6	639

M.orig.1- Modell for alle fire områder

n- antall prøveflater

7.1.2 MODELL FOR HVERT OMRÅDE UAVHENGIG AV STRATUM

Eventuelle forskjeller mellom områder ble forsøkt analysert å utvikle separate modeller for de ulike områdene (graf 3, vedlegg).

Ulike variable inngår i de forskjellige modellene, men forklaringsevnen er forholdsvis lik. Det vil si at R_a^2 verdier er høyest i Våler og Nordre Land på 0,91 hver for seg. Laveste R_a^2 er i Krødsherad på 0,87 (Tabell 15). *Condition number* (κ) er lavest i modellen til Hole ($\kappa=1,6$) og høyest i modell til Krødsherad ($\kappa=7,9$) (Tabell 11).

TABELL 11: MODELL FOR HVERT OMRÅDE UAVHENGIG AV STRATUM

Område	Variablene i modellen	Modelltype	R_a^2	RMSE	κ	n
Hole (geo.id. I)	$\ln P_{90}; \ln F_{20}$	M.orig.2	0,89	0,22	1,6	115
Våler (geo.id. II)	$\ln DZ_{\text{avg}}; \ln P_{30}; \ln F_0$	M.orig.2	0,91	0,15	7,5	143
Nordre Land (geo.id. III)	$\ln DZ_{\text{avg}}; \ln P_{20}; \ln F_0$	M.orig.2	0,91	0,21	6,4	265
Krødsherad (geo.id. IV)	$\ln P_{20}; \ln P_{80}; \ln F_{50}; \ln F_0$	M.orig.2	0,87	0,23	7,9	116

M.orig.2- Modell for hvert område uavhengig av stratum

7.1.3 MODELL FOR HVERT STRATUM UAVHENGIG AV OMRÅDE

Effekter av stratifisering ble nedenfor undersøkt ved å utvikle separate modeller uten å ta hensyn til at dataene kom fra fire områder. Forskjellige uavhengige variable inngår i modellene, men variabelen $\ln F_0$ inngår i alle tre modellene. R_a^2 er høyest i stratum I og lavest i stratum III (Tabell 12) (graf 4, vedlegg).

TABELL 12: MODELL FOR HVERT STRATUM UAVHENGIG AV OMRÅDE

	Variablene i modellen	Modelltype	R_a^2	RMSE	κ	n
Stratum I	$\ln DZ_{avg}; \ln F_0$	M.orig.3	0,88	0,23	2,2	204
Stratum II	$\ln P_{90}; \ln F_0$	M.orig.3	0,81	0,24	1,5	196
Stratum III	$\ln DZ_{avg}; \ln P_{30}; \ln F_0$	M.orig.3	0,74	0,21	6,6	239

M.orig.3- Modell for hvert stratum uavhengig av område

7.1.4 MODELL FOR HVERT OMRÅDE PER STRATUM

Nedenfor presenteres separate modeller for område og stratum (Tabell 13). For de ulike modellene inngår fra to til fire ulike uavhengige variable avledet fra laserdata. Hole har høyeste R_a^2 i stratum II med 0,92. Våler, Nordre Land og Krødsherad har høyest R_a^2 i stratum I med verdier 0,93; 0,92 og 0,84 (graf nr. 5,6,7,8, vedlegg).

TABELL 13: MODELL FOR HVERT OMRÅDE PER STRATUM

Område	Stratum	Variablene i modellen	Modell type	R_a^2	RMSE	κ	n
Hole	I	$\ln DZ_{CV}; \ln P_0; \ln F_{30}$	M.orig.4	0,80	0,26	2,9	35
	II	$\ln P_{90}; \ln F_{20}$	M.orig.4	0,92	0,21	2,1	31
	III	$\ln P_{80}; \ln F_{40}$	M.orig.4	0,75	0,15	1,3	49
Våler	I	$\ln DZ_{avg}; \ln F_0$	M.orig.4	0,93	0,14	2,1	56
	II	$\ln DZ_{avg}; \ln P_{30}; \ln F_{60}$	M.orig.4	0,71	0,18	5,4	36
	III	$\ln P_{80}; \ln F_{60}; \ln F_0$	M.orig.4	0,88	0,14	2,3	51
Nordre Land	I	$\ln DZ_{avg}; \ln P_{10}; \ln F_{10}$	M.orig.4	0,92	0,21	5,7	74
	II	$\ln P_{80}; \ln F_0$	M.orig.4	0,82	0,22	1,4	91
	III	$\ln DZ_{avg}; \ln P_{20}; \ln F_0$	M.orig.4	0,80	0,21	5,1	100
Krødsherad	I	$\ln DZ_{avg}; \ln F_{30}$	M.orig.4	0,84	0,26	3,1	39
	II	$\ln P_{80}; \ln F_{60}$	M.orig.4	0,79	0,25	2,5	38
	III	$\ln DZ_{avg}; \ln P_{20}; \ln F_{90}; \ln F_{80}$	M.orig.4	0,66	0,25	6,9	39

M.orig.4- Modell for hvert område per stratum

7.2 VALIDERING AV DE ULIKE MODELLENE

Modellene ble validert mot uavhengige valideringsdatasett fra de samme områdene for hvert enkelt stratum. Valideringen er organisert i forhold til valideringsresultater til ulike modellbaser:

- *M.orig.1* - Modell alle fire områder
- *M.orig.2* - Modell per område uavhengig av stratum
- *M.orig.3* - Modell per stratum uavhengig av område
- *M.orig.4* - Modell per område per stratum

Fra tabell 14 kan vi se at i Hole har modellbase *M.orig.2* laveste RMSE både i stratum II og III. I Våler har modellbase *M.orig.3* gitt laveste RMSE i stratum I og III, mens i stratum II har modellbase *M.orig.2* lavest RMSE. Nordre Land har både i stratum II og III laveste RMSE for modellbase *M.orig.1*. Til slutt for Krødsherad har stratum I lavest RMSE gitt modellbase *M.orig.3* mens i stratum II og III er det tilfelle med modellbase *M.orig.2*.

TABELL 14: RMSE (m^3/ha) TIL MODELLENE. RMSE I PROSENT AV FELTMÅLT VOLUM ER ANGITT I PARENTES

Område	Stratum	Antall kontrollflater	Volum (m^3/ha)			
			M.orig.1	M.orig.2	M.orig.3	M.orig.4
Hole	II	15	25,0 (17 %)	20,1 (14 %)	28,7 (20 %)	23,6 (16 %)
	III	24	30,4 (13 %)	29,4 (13 %)	31,9 (14 %)	23,5 (10 %)
Våler	I	22	28,1 (17 %)	27,9 (17 %)	26,5 (16 %)	27,7 (17 %)
	II	19	34,6 (23 %)	16,5 (11 %)	36,6 (24 %)	20,7 (14 %)
	III	20	33,3 (12 %)	36,8 (13 %)	32,7 (12 %)	35,3 (13 %)
Nordre Land	II	5	45,3 (30 %)	57,9 (39 %)	47,4 (32 %)	66,5 (44 %)
	III	34	53,1 (17 %)	54,0 (18 %)	54,7 (18 %)	55,7 (18 %)
Krødsherad	I	18	30,5 (13 %)	31,5 (15 %)	26,7 (13 %)	27,3 (13 %)
	II	19	20,9 (13 %)	16,0 (10 %)	24,3 (15 %)	26,8 (17 %)
	III	19	39,9 (14 %)	37,4 (13 %)	46,1 (16 %)	51,3 (18 %)

I tabell 15 ser vi, for Hole har modellbase *M.orig.2* i stratum II og III gitt den minste systematiske feil på 3,5 % og 0,1 %. Største systematiske feilen er i dette området med modellbasen *M.orig.4* i stratum II og *M.orig.3* i stratum III.

Våler har minste systematiske feilen i stratum I på 0,8 % med modellbasen *M.orig.3* og største med modellbasen *M.orig.2*. For stratum II modellbasen *M.orig.4* har minste systematiske feilen på -1,2 %. I dette stratumet har modellbasene *M.orig.1* og *M.orig.3* vist signifikant forskjell mellom predikert og observert volum. Stratum III har modellbasen *M.orig.1* med minste systematiske feilen på 1,7 % og største systematiske feilen er med modellbasen *M.orig.2* (-6,8 %).

Nordre Land har den minste systematiske feilen på 20,1 % i stratum II og -3,0 % i stratum III med modellbasen *M.orig.3* i begge tilfelle. Krødsherad har den minste systematiske feilen i stratum II (-1,3 %) og i stratum III (3,2 %) med samme modellbasen *M.orig.2*. I stratum I er det modellbasen *M.orig.4* med minste systematiske feilen på -1,6 %. I stratum II har modellbasene *M.orig.1* og *M.orig.3* gitt signifikant forskjell mellom predikert og observert volum, stratum III har her signifikant forskjeller mellom predikert og observert volum med modellbasen *M.orig.3*.

TABELL 15: BIAS TIL MODELLENE (m^3/ha). BIAS I PROSENT AV FELTMÅLT VOLUM ER ANGITT I PARENTES

Område	Stratum	Antall kontrollflater	Volum (m^3/ha)			
			M.orig.1	M.orig.2	M.orig.3	M.orig.4
Hole	II	15	6,8 (4,7 %)	5,1 (3,5 %)	11,7 (0,5 %)	11,0 (7,6 %)
	III	24	-1,9 (-0,8 %)	0,3 (0,1 %)	-7,5 (-3,2 %)	1,3 (-0,6 %)
Våler	I	22	3,3 (2,0 %)	-11,4 (-6,9 %)	1,3 (0,8 %)	-10,7 (-6,5 %)
	II	19	30,7*** (20 %)	7,2 (4,7 %)	31,6*** (20,6 %)	-1,8 (-1,2 %)
	III	20	4,7 (1,7 %)	-18,8 (-6,8 %)	5,4 (1,9 %)	-15,9 (-5,7 %)
Nordre Land	II	5	33,3 (22,3 %)	45,9 (30,7 %)	30,0 (20,1 %)	53,1 (35,5 %)
	III	34	-10,4 (-3,4 %)	15,8 (5,2 %)	-9,0 (-3,0 %)	16,8 (5,5 %)
Krødsherad	I	18	5,7 (2,7 %)	-9,7 (-4,6 %)	3,2 (1,5 %)	-2,7 (-1,3 %)
	II	19	15,5*** (9,6 %)	-2,5 (-1,6 %)	17,9*** (11,0 %)	-5,7 (-3,5 %)
	III	19	21,9 (7,6 %)	9,0 (3,2 %)	30,2*** (10,5 %)	-12,6 (-4,4 %)

Signifikansnivå: ***p < 0.001, **p < 0.01, *p < 0.05; ns: ikke signifikant (p > 0.05)

7.3 DUMMYANALYSER

For å undersøke effekten av strata og område på regresjonsmodellene ble det brukt dummyvariable som representerte stratum (Ds1=1 hvis stratum II; Ds2=1 hvis stratum III) og område (Do1=1 hvis område Våler; Do2=1 hvis område Krødsherad; Do3=1 hvis område Hole), tabell 16.

Dummyvariable blei satt inn i tre forskjellige modellbaser:

- Modell alle fire områder (*M.orig.1*):
- Modell per område uavhengig av stratum (*M.orig.2*):
- Modell per stratum uavhengig av område (*M.orig.3*):

TABELL 16: P-VERDIER FOR DUMMYVARIABLE

Modellbasen	Områder/Stratum	Variablene i modellen
M.orig.1	Alle fire områder	Ds1ns; Ds2ns
	Alle fire områder	Do1***; Do2***; Do3***
M.orig.2	Hole	Ds1ns; Ds2ns
	Våler	Ds1ns; Ds2ns
	Nordre Land	Ds1ns; Ds2ns
	Krødsherad	Ds1ns; Ds2ns
M.orig.3	I	Do1***; Do2***; Do3***
	II	Do1***; Do2***; Do3**
	III	Do1***; Do2***; Do3ns

Signifikansnivå: ***p < 0.001, **p < 0.01, *p < 0.05; ns: ikke signifikant (p > 0.05)

Med denne typen av analyse er det bare konstantledd som vurderes.

Ved å sette inn dummy variablene i de utvalgte modellene (Tabell 16) viste det seg at for modell basisen *M.orig.1* der dummy variablene representerte stratum hadde stratum ingen signifikant effekt på modellen. Med utgangspunkt i den samme modellbasen blei dummyvariablene deretter satt til å representere område og det har vist seg at områdene hadde signifikant effekt på modellen.

For modellbasen *M.orig.2* representerte dummyvariablene stratum. Det viste seg at stratum hadde ingen signifikant effekt på disse modeller basert på alle flatene innenfor et område.

I modellbasen *M.orig.3* representerte dummyvariablene områdene. I alle strata har det vist seg at område hadde signifikant effekt på modellene, med en viss unntak i modellbasen *M.orig.3* for stratum III der Hole hadde ingen signifikant effekt.

7.4 REGRESJONSMODELLER FOR ULIK TRESLAGSANDEL

For å kunne utvikle modeller må vi ha minst 22 prøveflater. Dette kommer av at vi har 22 laservariabeler tilgjengelig for utvikling av modeller. Med tilleggsstratifisering på grunnlag av ulike treslagsandeler blei det altfor få prøveflater til utvikling av modeller per område og stratum. Derfor blei det utviklet bare to modellbaser på grunnlag av forskjellig treslagsandel:

- Modell per område uavhengig av stratum for treslagsandel 0,5 og 0,8 ($M.O_{0,5-0,8}$)
- Modell per stratum uavhengig av område for treslagsandel 0,5 og 0,8 ($M.S_{0,5-0,8}$)

7.4.1 MODELL PER OMRÅDE UAVHENGIG AV STRATUM FOR TRESLAGSANDLER 0,5 OG 0,8

I Hole er R_a^2 høyere i modellen med treslagsandelen 0,8 for gran sammenlignet med modellen til treslagsandelen 0,5 (Tabell 17). I dette området var det ikke nok prøveflater til å utvikle modell for furu både med treslagsandelen 0,5 og 0,8 og for skog med treslagsblanding der en treslagsandel på 0,5 ble benyttet for å definere gran og furu flater.

For granmodellen i Våler (Tabell 17) er R_a^2 høyere (0,90) med treslagsandelen 0,5 enn med treslagsandelen 0,8. For utvikling av furumodell var det bare nok prøveflater for treslagsandelen 0,5 og for blanding var det nok prøveflater bare i treslagsandelen 0,8.

Nordre Land (Tabell 17) har modellen for gran med treslagsandelen 0,8 noe høyere R_a^2 med færre prøveflater til utvikling av modellen ($n=171$) og lavere κ (2,4) enn modellen med treslagsandelen 0,5. For furu modellen er det omvendt, her er det modell med treslagsandelen 0,5 ($R_a^2=0,84$) som har noe høyere R_a^2 enn modellen med treslagsandelen 0,8. For blanding var det bare treslagsandelen 0,8 som hadde nok flater til å utvikle modell.

Krødsherad (Tabell 17) hadde bedre R_a^2 for treslagsandelen 0,8 både for gran og furu og med mindre prøveflater enn med treslagsandelen 0,5 for utvikling av modeller. Det var laget modell til blanding bare for treslagsandel 0,8 siden den hadde nok prøveflater ($n=45$).

TABELL 17: MODELL PER OMRÅDE UAVHENGIG AV STRATUM FOR TRESLAGSANDELER 0,5 OG 0,8

Område	Treslag	Treslagsfordelingsgrense	Variablene i modellen	Modell type	R_a^2	RMSE	κ	n
Hole	Gran	0,5	$\ln DZ_{avg}; \ln DZ_{CV}; \ln F_{20}$	M.O _{0,5}	0,90	0,19	2,7	107
		0,8	$\ln DZ_{avg}; \ln DZ_{CV}; \ln F_{20}$	M.O _{0,8}	0,92	0,17	2,8	89
	Furu	0,5						4
		0,8						1
	Blanding	0,5						4
		0,8	$\ln P_{80}; \ln F_{10}$	M.O _{0,8}	0,86	0,24	1,8	25
Våler	Gran	0,5	$\ln P_{80}; \ln F_{40}$	M.O _{0,5}	0,90	0,14	1,3	74
		0,8	$\ln P_{30}; \ln P_{40}; \ln F_0$	M.O _{0,8}	0,74	0,15	8,9	36
	Furu	0,5	$\ln P_{40}; \ln F_0$	M.O _{0,5}	0,91	0,14	1	56
		0,8						20
	Blanding	0,5						13
		0,8	$\ln DZ_{avg}; \ln F_{10}$	M.O _{0,8}	0,92	0,14	1,4	87
Nordre Land	Gran	0,5	$\ln DZ_{avg}; \ln P_{20}; \ln F_0$	M.O _{0,5}	0,95	0,17	6,7	205
		0,8	$\ln P_{60}; \ln F_0$	M.O _{0,8}	0,95	0,17	2,4	171
	Furu	0,5	$\ln P_{80}; \ln F_{20}$	M.O _{0,5}	0,84	0,23	1,2	55
		0,8	$\ln P_{80}; \ln F_{10}$	M.O _{0,8}	0,82	0,22	1	42
	Blanding	0,5						5
		0,8	$\ln P_{40}; \ln F_0$	M.O _{0,8}	0,88	0,26	2,6	52
Krødsherad	Gran	0,5	$\ln DZ_{avg}; \ln F_0$	M.O _{0,5}	0,92	0,17	2,3	55
		0,8	$\ln DZ_{avg}; \ln F_0$	M.O _{0,8}	0,94	0,16	2,4	36
	Furu	0,5	$\ln P_{80}; \ln F_{30}$	M.O _{0,5}	0,89	0,22	2,4	52
		0,8	$\ln P_{80}; \ln F_0$	M.O _{0,8}	0,90	0,2	2,1	35
	Blanding	0,5						9
		0,8	$\ln P_{80}; \ln F_0$	M.O _{0,8}	0,79	0,28	2	45

M.O_{0,5-0,8}- Modeller per område uavhengig av stratum for ulike treslagsandeler

7.4.2 MODELL PER STRATUM UAVHENGIG AV OMRÅDE FOR TRESLAGSANDELER 0,5 OG 0,8

Ut i fra tabell 18 ser vi i stratum I har modellen for gran gitt høyere R_a^2 for treslagsandelen 0,8 enn for treslagsandelen 0,5 med færre prøveflater. Furumodellen i dette stratumet var det bare nok prøveflater for treslagsandelen 0,5 og for blandingsmodellen var det nok prøveflater i treslagsandelen 0,8.

Stratum II i fra samme tabellen er R_a^2 høyere for treslagsandelen 0,8 for granmodellen og furumodellen. I dette stratumet for utvikling av blandingsmodellen for treslagsandelen 0,5 var det ikke nok prøveflater, så det er her utviklet modell bare for treslagsandelen 0,8.

Stratum III har granmodell med treslagsandelen 0,8 høyere R_a^2 enn treslagsandelen 0,5. For furumodellen er situasjonen den motsatte der treslagsandel 0,5 har høyere R_a^2 enn treslagsandelen 0,8. Blandingsmodellen i dette stratumet var utviklet bare for treslagsandelen 0,8.

TABELL 18: MODELL PER STRATUM UAVHENGIG AV OMRÅDE FOR ULIKE TRESLAGSANDELER 0,5 OG 0,8

Stratum	Treslag	Treslagsforde- lingsgrense	Variablene i modellen	Modell type	R_a^2	RMSE	κ	N
Stratum I	Gran	0,5	$\ln P_{60}; \ln F_0$	M.S _{0,5}	0,92	0,2	2,6	153
		0,8	$\ln P_{40}; \ln F_{10}$	M.S _{0,8}	0,93	0,19	2,8	104
	Furu	0,5	$\ln DZ_{avg}; \ln F_{40}$	M.S _{0,5}	0,81	0,19	2,3	33
		0,8						10
	Blanding	0,5						18
		0,8	$\ln DZ_{avg}; \ln F_0$	M.S _{0,8}	0,87	0,22	2,6	90
Stratum II	Gran	0,5	$\ln P_{90}; \ln F_0$	M.S _{0,5}	0,85	0,22	1,7	108
		0,8	$\ln P_{50}; \ln P_{90}; \ln F_0$	M.S _{0,8}	0,9	0,19	4,2	86
	Furu	0,5	$\ln P_0; \ln P_{50}; \ln F_{10}; \ln F_0$	M.S _{0,5}	0,85	0,19	9,8	83
		0,8	$\ln P_{90}; \ln F_{30}$	M.S _{0,8}	0,85	0,19	1,1	58
	Blanding	0,5						5
		0,8	$\ln DZ_{avg}; \ln DZ_{CV}; \ln F_{50}$	M.S _{0,8}	0,8	0,23		52
Stratum III	Gran	0,5	$\ln DZ_{avg}; \ln P_{30}; \ln F_{10}$	M.S _{0,5}	0,79	0,19	7,3	180
		0,8	$\ln DZ_{avg}; \ln F_{10}$	M.S _{0,8}	0,8	0,18	1,8	142
	Furu	0,5	$\ln P_{70}; \ln F_{40}$	M.S _{0,5}	0,81	0,19	1,7	51
		0,8	$\ln DZ_{avg}; \ln F_{40}$	M.S _{0,8}	0,79	0,21	1,9	30
	Blanding	0,5						4
		0,8	$\ln P_{60}; \ln F_0$	M.S _{0,8}	0,77	0,22	1,8	67

M.S_{0,5-0,8}- modeller per stratum uavhengig av området for ulike treslagsandeler

7.5 VALIDERING AV MODELLER BASERT PÅ FORSKJELLIGE TRESLAGSFORDELINGER

Valideringen av de ulike modeller har foregått mot uavhengig datasett for hvert område per stratum.

- $M.O_{0,5-0,8}$ - modeller per område uavhengig av stratum for treslagsandeler 0,5 og 0,8
- $M.S_{0,5-0,8}$ - modeller per stratum uavhengig av området for treslagsandeler 0,5 og 0,8

Ut i fra tabellen 19 og 20 ser vi at i Hole i stratum II har modellbase $M.O_{0,5}$ for gran minste bias på 5,7 % og for blandingsmodellen er $M.O_{0,8}$ med minste bias (-15,6 %). Det samme gjelder for RMSE verdiene i tabell 20 for Hole, stratum II, gran- og blandingsmodell. I stratum III i samme området har granmodellen med modellbasen $M.O_{0,8}$ gitt minste bias på 0,6 % (Tabell 19) og RMSE på 11 % (Tabell 20). Den samme modellbasen har også for blandingsmodellen gitt minste bias på -9,8 % og RMSE 18 %.

Våler: har for granmodellen i stratum I minste bias for modellbasen $M.S_{0,5}$ (-0,3 %) og RMSE er 12 %. I det samme stratumet for furumodellen har modellbasen $M.S_{0,5}$ gitt laveste bias (-4,3 %) og RMSE (15 %). Blandingsmodellen i dette stratumet har vist seg å være best med modellbasen $M.O_{0,8}$ som ga minste bias på -5,9 % men her var RMSE mindre med modellbase $M.S_{0,8}$ som var 15%.

I stratum II i dette området var granmodellen med modellbasen $M.O_{0,5}$ med minste bias (8,7 %) og RMSE (9%). Furumodellen var også i dette stratumet best med modellbasen $M.O_{0,5}$ der bias hadde en verdi på 1,2 % og RMSE på 11%. Blandingsmodellen i dette stratumet med modellbasen $M.O_{0,8}$ hadde minste bias på 0,2 % og RMSE på 8%.

Stratum III i dette området for granmodellen ga modellbasen $M.O_{0,5}$ minste bias på -5,1% og en RMSE på 12%. I dette stratumet blandingsmodell med modellbasen $M.S_{0,8}$ hadde minste bias med -3,3% og en RMSE på 11%.

Nordre Land: i stratum II har furumodellen med $M.S_{0,8}$ minste bias på 9,3 % og RMSE på 22%. Den samme modellbasen $M.S_{0,8}$ ga også minste bias (-3,2 %) og RMSE (4%) for blandingsmodellen i dette stratumet. Det var ingen granmodeller i dette stratumet.

I stratum III i dette området var for granmodellen med modellbasen $M.O_{0,8}$ med minste bias (2,6 %) og RMSE (12 %). Furumodellen i dette stratumet var med modellbasen $M.O_{0,5}$ med minste bias (3,2 %) mens RMSE for denne modellen var mindre for modellbasen $M.S_{0,5}$. Blandingsmodellen i dette stratumet var modellbasen $M.S_{0,8}$ med minste bias (-1,3 %) og en RMSE på 6 %.

Krødsherad: har i stratum I for granmodell minste bias på - 1,7 % med modellbasen $M.O_{0,5}$ men her har modellbasen $M.O_{0,8}$ en lavere RMSE (6 %). Furumodellen i dette stratumet har lavest bias (-3,3 %) og RMSE (11 %) med modellbasen $M.S_{0,5}$. Blandingsmodellen i dette stratumet har lavest bias (-0,3 %) med modellbasen $M.O_{0,8}$ men her er RMSE lavere med modellbasen $M.S_{0,8}$ (14 %).

Stratum II i dette området har granmodell laveste bias på 1,1 % med modellbasen $M.O_{0,5}$ men her er RMSE lavere for modellbasen $M.O_{0,8}$. Furumodellen i dette stratum har lavest bias (2,3 %) og RMSE (10 %) med modellbasen $M.S_{0,8}$. Blandingsmodellen her har lavest bias (-5,2 %) og RMSE (13 %) med modellbasen $M.O_{0,8}$.

Stratum III i dette området har for granmodellen minste bias (4,1 %) og RMSE (12 %) med modellbasen $M.O_{0,5}$. Furumodellen har minste bias (7,9 %) og RMSE (11 %) med modellbase $M.S_{0,5}$. Mens blandingsmodell har minste bias (5,8 %) og RMSE (12 %) med modellbasen $M.O_{0,8}$.

TABELL 19: BIAS TIL MODELLENE MED TRESLAGSANDEL 0,5 OG 0,8 (m^3/ha). BIAS I PROSENT AV FELTMÅLT VOLUM ER ANGITT I PARENTES

Område	Stratum	Treslag	M.O _{0,5}	M.O _{0,8}	M.S _{0,5}	M.S _{0,8}
Hole	II	Gran	8,5 (5,7 %)	13 (10,2 %)	23,3*(15,6 %)	33,3**(26,1 %)
		Furu				
		Blanding		-30,7(-15,6 %)		-42,1(-21,4 %)
	III	Gran	3,0 (1,2 %)	1,3(0,6 %)	-1,6(-0,7 %)	3,2 (1,4 %)
		Furu				
		Blanding		-22,9 (-9,8 %)		-36,4(-15,6 %)
Våler	I	Gran	-18,3*(-8,8 %)	-42,4(-12,7 %)	-0,6(-0,3 %)	-16,3(-4,9 %)
		Furu	-13,5(-10,1 %)		-5,7 (- 4,3 %)	
		Blanding		-8,8(-5,9 %)		-9,2(-6,2 %)
	II	Gran	16,3*(8,7 %)		56,7(30,1 %)	
		Furu	1,8(1,2 %)		7,7(5,1 %)	18,8 (15,4 %)
		Blanding		0,3(0,2 %)		2,0 (1,3 %)
	III	Gran	-14,7(-5,1 %)	-19,6(-6,2 %)	18,6*(6,5 %)	21,6(6,8 %)
		Furu				
		Blanding		-16,2(-6,6 %)		-8,2(-3,3 %)
Nordre Land	II	Gran				
		Furu	18,3 (12,2 %)	10,8(10,4 %)	19,5(13,0 %)	9,6 (9,3 %)
		Blanding		54,7(25,0 %)		-7,1 (-3,2 %)
	III	Gran	19,7(5,9 %)	9,1(2,6 %)	-18,8(-5,6 %)	-25,7*(-7,3 %)
		Furu	8,1(3,2 %)	36,4*** (15,2 %)	9,8(3,8 %)	16,4(6,9 %)
		Blanding		10,9(3,9 %)		-3,6(-1,3 %)
Krødsherad	I	Gran	-4,2(-1,7 %)	5,5(1,9 %)	7,7(3,1 %)	16,4(5,6 %)
		Furu	-22,5(-12,8 %)	-29,8(- 14,0 %)	-5,8 (-3,3 %)	
		Blanding		-0,6(-0,3 %)		-1,3(-0,8 %)
	II	Gran	2,4(1,1 %)	3,4(1,3 %)	39,1*(17,8 %)	62,5*(24,6 %)
		Furu	4,9(3,5 %)	9,8(7,4 %)	4,3 (3,1 %)	3,1(2,3 %)
		Blanding		-8,2(-5,2 %)		-8,5(-5,4 %)
	III	Gran	12,2 (4,1 %)	25,7*(8,1 %)	31,7*(10,6 %)	41,7**(13,2 %)
		Furu	31,7(12,1 %)	37,1(13,7 %)	20,6(7,9 %)	33,1(12,3 %)
		Blanding		13,1(5,8 %)		13,5(6,0 %)

Signifikansnivå: ***p< 0.001, **p< 0.01, *p< 0.05; ns: ikke signifikant (p> 0.05)

TABELL 20: RMSE TIL MODELLENE MED TRESLAGSANDEL 0,5 OG 0,8 (m^3/ha). RMSE I PROSENT AV FELTMÅLT VOLUM ER ANGITT I PARENTES.

Område	Stratum	Treslag	M.O _{0,5}	M.O _{0,8}	M.S _{0,5}	M.S _{0,8}
Hole	II	Gran	21,76 (15 %)	23,72 (19 %)	37,32 (25 %)	44,22(35 %)
		Furu				
		Blanding		32,35(16 %)		50,5 (26 %)
	III	Gran	29,82 (13 %)	25,59(11 %)	31,9(14 %)	28,66(12 %)
		Furu				
		Blanding		41,71 (18 %)		50,84(22 %)
Våler	I	Gran	27,51(13 %)	48,27(15 %)	24,42(12 %)	29,39(9 %)
		Furu	24,06(18 %)		19,46(15 %)	
		Blanding		22,7(15 %)		22,21(15 %)
	II	Gran	16,35(9 %)		59(31 %)	
		Furu	15,9(11 %)		19,98(13 %)	24,33(20 %)
		Blanding		13,18(8 %)		23,63(15 %)
	III	Gran	34,93(12 %)	37,83(12 %)	36,85(13 %)	40,69(13 %)
		Furu				
		Blanding		29,53(12 %)		26,4(11 %)
Nordre Land	II	Gran				
		Furu	28,91(19 %)	23,89(23 %)	36,02(24 %)	22,67(22 %)
		Blanding		57,18(26 %)		8,28(4 %)
	III	Gran	50,29(15 %)	41,6(12 %)	50,8(15 %)	46,3(13 %)
		Furu	27,37(11 %)	39,59(17 %)	22,83(9 %)	31,84(13 %)
		Blanding		23,93(9 %)		15,65(6 %)
Krødsherad	I	Gran	17,38(7 %)	16,88(6 %)	21,82(9 %)	22,91(8 %)
		Furu	32,85(19 %)	46,25(22 %)	19,82(11 %)	
		Blanding		28,34(18 %)		23,26(14 %)
	II	Gran	13,02(6 %)	8,27(3 %)	42,89(20 %)	64,53(25 %)
		Furu	17,47(13 %)	17,47(13 %)	17,16(12 %)	13,48(10 %)
		Blanding		19,79(13 %)		26,25(17 %)
	III	Gran	35,28(12 %)	41,83(13 %)	44,09(15 %)	53,01(17 %)
		Furu	41,23(16 %)	44,94(17 %)	27,54(11 %)	38,14(14 %)
		Blanding		25,99(12 %)		27,14(12 %)

8 DISKUSJON

Feltarbeid med måling av prøveflater er essensielt dersom man skal taksere skog ved hjelp av laser og arealmetoden. Det er viktig og få gode sammenhenger mellom prøveflatedata og laserdata, men feltarbeid og prøveflatetaksering er tidkrevende og dyrt. En effektiv datainnsamling er derfor viktig. I regioner der klimatiske forhold er like og trærne har tilnærmet samme forhold mellom kroneform, kronetetthet og stammevolum, kan det være mulig til å utnytte prøveflater fra nærliggende områder.

I denne oppgaven har jeg derfor undersøkt om det er forskjell i forholdet mellom laserregistreringer og feltmålt volum mellom områder, ved at jeg laget separate modeller for de ulike områder. Videre er datasettene slått sammen og materialet delt inn i ulike grupper eller strata for å undersøke effekten av å øke prøveflateantallet ved utvikling av modellene. For å illustrere denne effekten er det også utviklet separate modeller for hvert enkelt område og stratum. Dataene benyttet i denne oppgaven er tidligere publisert i ulike separate studier og noen av resultatene er derfor tilnærmet de samme som i de publiserte arbeidene. Noen flater er også tatt ut, da man har funnet ulike feil i registreringene etter at arbeidet med de separate studiene ble publisert. Detaljert informasjon om områdene: Våler, Krødsherad og Nordre Land finnes i *Næsset (2002,2004a, d)*

8.1 ULIKE REGRESJONSMODELLER

Ved utvikling av regresjonsmodeller er det viktig å ha et tilstrekkelig antall observasjoner. Dersom man har få observasjoner kan man få en dårlig modell som ikke kan benyttes til prediksjon. I denne oppgaven er alle modellene evaluert ved å se på R_a^2 , siden den gir en ytterligere dimensjon enn R^2 og tar hensyn til kompleksitet i modellen.

I et datamateriale med prøveflater vil det tilnærmet alltid være noen ekstreme flater, dvs. flater der det finnes for eksempel observasjoner med bare et tre (eller veldig få trær). Disse vil ha stor innflytelse på modellens forløp og ”styre” funksjonene. En bør selvfølgelig være forsiktig med å ta ut disse ekstreme observasjonene fordi de er en del av våre data som skal representere populasjonen, men man må vurdere om spesielle flater enkelte ganger skal tas ut av ulike årsaker.

8.1.1 MODELL FOR ALLE FIRE OMRÅDER

Med denne modellbasen (*M.orig.1*) kan vi se hvor mye jevnfordeling av prøveflatene mellom områdene har betydelig regresjonsmodellen. Dersom vi setter prøveflateantallet i prosent så ser vi at Nordre Land gjekk inn med 41,5 % av prøveflater, Hole med 18 % av prøveflater, Våler med 22,4 % og Krødsherad med 18,2 %. Standardavviket på feltmålt volum var 472,3 (m³/ha) og gjennomsnittlig volum var 342,9 (m³/ha) for denne modellbasen på prøveflatene.

I modellbasen (*M.orig.1*) for alle fire områder var det til sammen 639 prøveflater tilgjengelige for utvikling av modellen, med en forklaringsgrad (R_a^2) på hele 88 % (Tabell 10). Dette viser at siden vi har såpass mange observasjoner vil vi alltid klare å tilpasse modellen ganske godt til datasettet. Når det gjelder *conditon number* i denne modellen, som er brukt som indikasjon på kollinearitet, så ligger den på et rimelig nivå ($k=5,61$) og det vil si at vi ikke har problemer med korrelerte forklaringsvariabler.

Ut vi fra valideringen hadde denne modellbasen vist seg å gi lavest RMSE (Tabell 14) i Nordre Land både i stratum II og III. Grunnen til dette kan være at hele 41,5 % av det totale prøveflateantallet tilhørte dette området. En annen ting som kan diskuteres her, knyttet til denne lave RMSE i stratum II og III i valideringen, er at i stratum II er antallet kontrollflater veldig lavt (Tabell 14). Det var bare 5 kontrollflater i dette stratumet. Med så høyt antall prøveflater som stammer fra dette området, og med så lavt antall kontrollflater fra samme område, er det stor sannsynlighet for at tilpassningen til det uavhengige datamateriale er god. I stratum III er kontrollflateantallet høyere ($n=34$) og det ga likevel laveste RMSE i forhold til andre modellbaser, men den var ikke mye lavere. Den systematiske feilen (Tabell 15) var ganske høy her for stratum II og lå på 33,3 m³/ha (22,3 %). For stratum III lå den på -10,4 m³/ha (-3,4 %). Så i første tilfelle var det en overestimering på hele 22,3 % og i andre tilfelle var det en underestimering på 3,4 %.

Ved sammenligning av resultatene fra denne modellbasen i forhold til de andre modellbasene finner vi forholdsvis lite systematiske feil for Nordre Land. Det er likevel en tendens til at systematiske feil i denne modellbasen er nest størst i stratum II, dette gjelder for Våler og Krødsherad.

I Våler var det 20 % systematiske feil i stratum II for denne modellbasen og det var signifikant forskjell mellom predikert og observert volum. I Krødsherad i samme stratumet har denne modellbasen vist signifikant forskjell mellom predikert og observert volum, og systematiske feil var 9,6 %. Igjen kan dette ha sammenheng med at vi i disse to områdene hadde det største antallet av kontrollflater, og et relativt lavt antall prøveflater til modellutvikling.

Den tilfeldige feilen (RMSE) gir uttrykk for den variable feilen, over- eller underestimering, som forekommer i det enkelte bestand i tillegg til den generelle nivåfeilen. Siden det er en så pass stor systematisk feil så kan vi forvente også stor tilfeldig feil.

8.1.2 MODELL FOR HVERT OMRÅDE UAVHENGIG AV STRATUM

Ut i fra denne modellbasen (*M.orig.2*) (Tabell 11) kan vi se at jo flere prøveflater vi hadde per område, jo bedre R_a^2 og lavere RMSE fikk vi ved utvikling av modellen. Vi ser at *condition number* er høyere med færre prøveflater og med større standardavvik på volum per område.

I Krødsherad er *conditon number* 7,9 og R_a^2 er ganske lav (87 %), dette kan forklares med at det er ganske stor variasjon i prøveflatevolum i dette området ($SD=459,6 \text{ m}^3/\text{ha}$) og at det er ganske få prøveflateantall ($n=116$) i forhold til andre områder.

R_a^2 -verdier i modellbasene er høyest i Våler og Nordre Land, på 0,91 hver for seg. Nordre Land har en stor standardavvik på feltmålt volum på prøveflatene for denne modellbasen ($SD= 434,8 \text{ m}^3/\text{ha}$), men har også det største prøveflateantallet til modelloppbygning. Dette gjør at vi her får høy forklaringsgrad på modellen og at *condition number* er høyere, fordi det trengs flere laservariabler til å bygge opp modellen.

Våler har det nest største prøveflateantallet til denne modellbasen ($n=143$) og standardavviket i volum var på $362,3 \text{ m}^3/\text{ha}$ som er forholdsvis lavt i forhold til standardavviket på volum i Krødsherad og Nordre Land. RMSE (0,15) i dette området var også lavest av alle tre områder.

Hole hadde R_a^2 i modellbasen på 0,89 og hadde det minste prøveflateantallet til oppbygning av denne modellen med 115 prøveflater. Også *condition number* til denne modellen var lavest (1,6). Dette kan forklares med at skogen i dette området var mer ensartet (Tabell 5) enn i de andre områdene ($SD=291 \text{ m}^3/\text{ha}$).

Når det gjelder valideringsresultater så har denne modellbasen gitt lavest RMSE (Tabell 14) og systematisk feil (Tabell 15) i Hole både i stratum II og III. Dette kan forklares med at vi økte prøveflateantallet for utvikling av denne modellbasen, mens variasjonen innen stratum II og III er liten siden disse kontrollflatene var grandominerte. I stratum III i valideringsdatasettet er standardavviket på volum er $37,82 \text{ m}^3/\text{ha}$, som også kan ha innvirkning.

Dette modellbasen har også vist seg å gi minst systematisk feil i Krødsherad i stratum II og III (Tabell 15) og vi kan komme med samme påstand som i tilfellet over, at ved å øke prøveflateantallet vil vi påvirke nøyaktigheten til stratum II og III, der det ellers er liten forklaringsgrad med *M.orig.4*.

Modellbasen *M.orig.2* kan sies å være bedre enn *M.orig. 3* siden det ikke er noen signifikante forskjeller mellom predikert og observert volum der. Grunnen til dette kan være at det er brukt samme flysensor, topografi og stratum innenfor områder har mindre påvirkning i modellbasen *M.orig.2* enn *M.orig.3*. Dette har også kommet frem i flere artikler av Næsset (et al. 2005; 2007)

8.1.3 MODELL FOR HVERT STRATUM UAVHENGIG AV OMRÅDE

For modellbasen per stratum uavhengig av områdene (*M.orig.3*) ser vi at vi har mest prøveflater til utvikling av modellen i stratum III og minst i stratum II. Prosentvis andel av prøveflater tilknyttet områdene for oppbygning av disse modellene er forskjellig. Generelt kan en si at Nordre Land har flest flater i alle tre modellene, spesielt i stratum II-modellen der nesten halvparten av prøveflaten er fra Nordre Land (46 %).

Dette kan til en viss grad påvirke og forklare valideringsresultatene. I Nordre Land så har denne modellbasen gitt minste systematisk feil både i stratum II (20,1 %) og III (-3,0 %) i forhold til andre modellbasene. Det er egentlig ganske stor systematisk feil i stratum II i dette området (20,1 %) og den systematiske feilen er også ganske stor i de andre områdene i dette stratumet for samme modellbasene unntatt Hole.

Denne modellbasen har gitt minste bias for Våler i stratum I (0,8 %) der dette området hadde den nest største andelen av prøveflater (27 %) til modelloppbygning. Størst andel av prøveflatene i dette stratumet er i fra Nordre Land med 36 %.

Denne modellbasen ga også tre tilfeller med signifikante forskjeller mellom predikert og observert volum. Dette skjedde i Våler i stratum II, som kanskje kan forklares med at Våler, for utvikling av denne modellen, hadde den nest minste andelen av prøveflater (18 %) i forhold til antall kontrollflater. Også i Krødsherad hadde denne modellbasen signifikant forskjell både i stratum II og i stratum III. Dette kan kanskje også forklares med at i stratum III for utvikling av modeller, hadde dette området minst prøveflater, bare 16 %. Våler skiller seg topografisk mest ut i fra de andre tre områdene siden området er ganske flatt. I Våler er det også forholdsvis få prøveflater. Dette påvirker valideringsresultater for denne modellbasen i dette området. Dette har også skjedd i Krødsherad der det både i stratum II og III er signifikante forskjeller i predikert volum og det er store systematiske feil.

Det kan trekkes frem at både for Krødsherad og Våler er det brukt forskjellige laserinstrument. Med en "uheldig" kombinasjon av ujevne prøveflateandeler, forskjellige instrumenter og forskjellige distrikter som disse to områdene hører til, kan vi ikke forvente å få en god modell ved å slå sammen ulike områder på stratumnivå. Forskjellige topografiske og typologiske forskjeller mellom områdene, det faktiske antallet prøveflater hvert område hadde til disponering og ikke minst ulike laserdata fra disse områdene, har større innvirkning på denne modellbasen.

I artikkelen (*Næsset & Bollandsås & Gobakken, 2005*) har de trukket frem at det ikke bare er forholdet mellom biofysiske egenskaper og laserdata som har betydning, men også laserdata i seg sjøl. Med dette viser de til at ved å kombinere de ulike datasettene oppnådd fra forskjellige flyhøyder og sensor, påvirker det fotavtrykkstørrelsen, som kan lede til alvorlig reduksjon av den nøyaktigheten som kan forventes fra prediksjonene på bestandsnivå.

8.1.4 MODELL FOR HVERT OMRÅDE PER STRATUM

Denne modellbasen hadde til utgangspunktet å utvikle modeller på grunnlag av stratum med de tilhørende prøveflatene innenfor hvert enkelt område. Alle modeller utviklet har rimelig høy forklaringsgrad (Tabell 13) og resultatene samsvarer bra med *Næsset* (*Næsset, 2002; Næsset et al., 2004; Næsset, 2004c; Næsset, 2007*) sine undersøkelser basert på de samme datasettene. Avvik fra Næssets resultater skyldes ulike digresjon etter min organisering av datamateriale.

En god modelltilpasning i hogstklasse 3 (dvs. stratum I) er vanlig *Næsset (2002)*. Dette kan skyldes at kroneformen på yngre skog ikke er ulik for gran og furu og at denne skogen ellers er forholdsvis ensartet.

I valideringen var det bare Krødsherad i stratum I som hadde minste systematiske feilen (-1,3 %) med denne modellen. Nordre Land hadde den største systematiske feilen for stratum II (35,5 %) og tilfeldige feilen (RMSE) på 44 %, disse var ikke signifikante. Dette kan forklares med at det bare var fem kontrollflater som var brukt i valideringen av denne modellen.

For andre områder lå den tilfeldige feilen på et rimelig nivå, med et intervall fra 10 % til 18 % for forskjellige områder i forskjellige stratum.

8.2 DUMMY ANALYSER

For modellbasene *M.orig.1* og *M.orig.3* der dummyvariablene som representerte område (Tabell 16) har vist signifikant effekt på modellen. Det kan være flere årsaker til dette resultatet.

Næsset & Gobakken (2008) har i sin studie også gjort blant anna en analyse av områdeeffekten på de utvalgte regresjonsmodeller. Områdeeffekten blei analysert ved å utvide utvalgte regresjonsmodeller for biomassen over bakken og biomassen under bakken med dummyvariabler som representerte de ti områdene i studiet ($A_2, A_3, \dots, A_{10}$).

Siden jeg har fire av disse områdene vil jeg presentere resultatene for biomassen over bakken i disse områdene som er omtalt i denne artikkelen. De satt dummyvariabel lik 0 for Våler, A_2 lik 1 for Krødsherad, A_3 lik 1 for Nordre Land og A_6 lik 1 for Hole.

Semi-elasticity (semi-elastisitet) var deretter brukt får å undersøke effekten av forskjellige områder ved å forandre på en enhet i den utvalgte variabelen (for eksempel A_2) mens alle andre variabler i modellen er konstante (uforandret). Med dette vil en få prosent forandring for den avhengige variabelen som i dette tilfelle var biomasse over eller under bakke.

Det har vist seg med dette at for biomassen over bakke var det dannet tre distriktsgrupper av områder der områder innenfor hver distriktsgruppe ikke var signifikant forskjellige. Dette kan påvise at blandingen av de forskjellige områder kan skje til en viss grad, dvs. at de forskjellige områdene tilhører en viss distriktsgruppe som igjen avhenger av de forskjellige topografiske, klimatiske osv forskjellene. Så det er dannet en grense på områdene der vi kan slå sammen ulike takser på et visst distriktsnivå.

I min oppgave har det vist seg at jeg har tre distriktsgrupper der Krødsherad og Hole dannet en gruppe, Nordre Land har dannet andre gruppen og Våler dannet tredje gruppen. Det er også viktig å merke seg at Krødsherad og Hole geografisk sett ligger ikke veldig langt i fra hverandre. Så ut i fra min dummyanalyse (Tabell 16) har det vist seg for disse forskjellige modellbaser (*M.orig.1* og *M.orig.3*) at det er signifikant forskjell mellom disse grupper dermed er det signifikant forskjell mellom disse områder.

Hvis vi ser på predikert volum ut i fra modellbasene *M.orig.1* og *M.orig.3* der alle laservariablene er like for alle fire områder og for hvert stratum, har distriktsgruppen en (Krødsherad og Hole områdene) og distriktsgruppen tre (Våler) signifikant forskjell mellom predikert og observert volum i stratum II og III. At denne signifikante forskjellen kommer ikke fram i distrikt 2 (Nordre Land) kan det forklares med den ujevne fordelingen av prøveflate antallet som gjekk inn i oppbygningen av modellbasen (42 %) og antallet kontrollflater til valideringen (bare 9 % i stratum II, dvs. 5 kontrollflater, og 35 % i stratum III, dvs. 34 kontrollflater). Ut i fra dette ser vi hvor viktig det er å ha prøveflatene fordelt i hvert fall likt mellom de forskjellige distriktene dersom en skal slå sammen laserdataene i fra forskjellige takster fra vidt forskjellige områder.

Så for å kontrollere effekten av de forskjellige geografiske regioner og de komponentene som følger med dem, som for eksempel bonitet, høyde over havet osv. vil vi klare også å kontrollere de forskjellige skogtyper siden deres utvikling avhenger sterkt ut i fra geografiske region. Forskjeller mellom disse geografiske regioner som påvirker utviklingen av stammeform, krone form, treslagsandel fordeling (sammensetning) osv. er også karakteristikk til hvert enkelt bestand og i større grad skogtypen som kan då kontrolleres med stratifisering og ved bruk av lokale prøveflater tilknyttet innenfor hvert enkelt stratum.

Det er stor diversitet i topografiske forhold og landskap mellom disse fire områder og ut i fra dette blei det analysert også om forskjellige laserinstrumentene som blei brukt i disse områdene kunne også forårsake forskjeller mellom ALS variabler. *Næsset & Gobakken (2008)* har skrevet at effekter av forskjellige områder kommer mer til uttrykk enn effekten av ALS instrumenter. Med forskjellige områder menes at faktorer som bestandsstruktur, stammeform og kroneform er signifikant forskjellige fra et område til et anna, då er det mest sannsynlig at variabiliteten fra estimerte regresjonsmodeller vil øke hvis slik bred variasjon av prøveflater er sammenslått i et estimeringsdatasett. Dette stemmer sannsynligvis også i mitt tilfelle siden laserinstrumentene brukt i artikkelen for disse fire områder er de samme og dummy analysen har vist at det er signifikant forskjell mellom områdene.

Næsset (2004 c) fant at siste retur data var mer påvirket av forskjellige flyhøyder enn første retur data. Derfor er det benyttet bare første retur data i oppgaven siden de er funnet å være mer stabile i forhold til forskjellige fly høyder og instrumenter.

I tidligere studier av *Næsset (2007)* er det også gjort dummyanalyser for å påvise effekt av områdene. Det har i denne studien vist seg at områdene ikke hadde noen signifikant effekt på de utvalgte skogvariablene, unntatt på dominant høyde. Det som er viktig å merke seg her og som også har kommet frem i diskusjon overfor, er at disse analysene ble gjort for to områder som sannsynligvis tilhørte samme distrikt, dvs. områdene lå ikke langt i fra hverandre. I denne artikkelen har det også kommet fram at tidligere forskning har vist at regresjonsmodeller utviklet over to ulike områder ser ut til å opptre ganske godt i begge områder hvis antall prøveflater er forholdsvis jevnt fordelt mellom områdene (*Næsset et al., 2005*).

Siden områder har signifikant effekt på regresjonsmodellen kan vi konkludere med at vi ikke kan bruke samme modeller i ulike områder som hører til forskjellige distriktsgrupper og dermed kan vi ikke spare på penger ved og ikke utføre noe takst men bare flygningen av området. Med dette kan vi då enkelt se ved å utføre for eksempel dummyanalysen om det er forskjeller mellom områder og hvor kan en "sette" grensen for eventuell sammenslåing av forskjellige laserdata fra forskjellige områder.

Dummyanalysen som blei gjort for modellbasen *M.orig.2* (Tabell 16) der dummyvariablene representerte stratum var ikke fleksibel nok til å fange opp variasjonen mellom stratum som tilhørte hvert enkelt område. Grunnen til dette kan være den store variasjonen i volum (m^3/ha) (Tabell 3) i hvert enkelt stratum på hvert enkelt område, slik at denne typen av analyser klarer ikke å fange opp variasjonen mellom sjølve strata.

Også i modellbasen *M.orig.1* var det satt dummyvariabler som skal representere stratum. Her var det også ingen signifikant effekt av stratum på denne modellen.

Det som kunne ha vart gjort her er å behandle for eksempel hogstklasse og område som dummyvariabel og bonitet som kontinuerlig variabel for å kunne vise effekten av disse faktorer på modellen. En slik behandling av variablene var gjort i *Næsset & Gobakken (2008)*.

8.3 GENERELL EFFEKT AV TRESLAGSANDELER PÅ REGRESJONSMODELLER

For å få innblikk i hva jeg kunne ha gjort i min oppgave for å vurdere effekten av treslagsandeler for de ulike modellbasene har jeg lyst å omtale artikkelen til *Næsset & Gobakken (2008)* og nærmere forklare prosedyren i steg 3, siden den berører løsningen på problemet hvordan behandle treslagsandelen i regresjonsmodellene for ulike områder.

Artikkelen inneholdt datamateriale fra ti forskjellige områder i sørlige delen av Norge.

Målet med denne artikkelen var å analysere effekten av skogtyper, forskjellige områder og laserinstrumentene som blei brukt i disse områdene på ALS biomasse forhold.

Behandlingen av variabler som skal vise effekten av forskjellige områder og alder (hogstklasse) var her behandlet som dummy variabler mens treslagsandeler var behandlet som kontinuerlige variabler i modeller.

Her var det først utvikla regresjonsmodeller ut i fra laserdataene ved å bruke multipell regresjonsanalyse. Det er foretatt tre steg for analysen. En kort beskrivelse er som følgende:

- Steg 1: Har som mål å finne enkel modell som inkluderer de laservariablene som bidrar mest med forklaring av biomassen. To variabler ($\ln h_{90}$ og $\ln d_0$) blei valgt ut her, både for biomassen over bakken og biomassen under bakken.
- Steg 2: Vurdering av stabiliteten på utvalgte modeller fra steg 1 blir gjort.
- Steg 3: Effekten av skogtype, forskjellige områder og laserinstrumentene var vurdert med å utvide de utvalgte modeller fra steg 1 og 2 med variabler som representerte disse faktorene.

I følge den målsetningen de har satt i artikkelen utvidet de videre de utvalgte modellene i fra steg 2 med variablene som representerte disse faktorene, dette er i gjort i steg 3. Effekten av forskjellige ALS instrumenter var kvantifisert og evaluert ved å modellere dem som kombinasjoner av effekter på områder. Skogtypene var uttrykt som forskjellige grupper av variabler knyttet til bonitet, treslagsandel og alder (*hogstklasse*). Bonitet og treslagsandelen var behandlet som kontinuerlige variabler. Dummy variablene var brukt til å vurdere effekten av alder og områdene.

Etter dette var det igjen beregnet Pearson korrelasjonskoeffisient for å se om det er antydning for potensielle kolinearitetsproblemer iblant uavhengige variablene som er lagt til i regresjonsmodellen fått i fra seg 1 og steg 2, og for å finne variablene som har høyest korrelasjon med biomassen. Det var oppdaget her at treslagsandelen av gran og furu var høyt korrelerte med biomassen over- og under- bakken. Siden dette var påvist har de bestemt seg for å utelatte furuandelen fra endelige modellen. Andel av gran og lauv var brukt videre i analysen, bonitet var også utelatt i fra endelige modellen.

Det var oppdaget i denne artikkel at inkludering av variablene som representerte forskjellige områder, aldersklasser (*hkl*) og treslagsandelen i først utvalgte modeller forbedret modellene signifikant.

Næsset (2004d) fant at skogtyper ikke hadde noen signifikant effekt på estimerte biomassemodeller. Men det er viktig her å understreke at han har brukt bare data fra Våler til denne undersøkelsen. *Næsset & Gobakken (2008)* undersøkte ti forskjellige områder og fant signifikante forskjeller mellom skogtyper og treslagsandeler. Parameterestimater for granandeler indikerte at ved å øke granandelen med 1 % mens alle andre variablene var konstante, så økte predikert biomasse over og under bakken med henholdsvis 0,2 % og 0,3 %.

Dette er forklart med at gran har ofte smal og konisk krone der relativ større andel av laserpulser treffer lenger nede i kronen sammenlignet med furu og lauvtreslag med lignende størrelser der mer runde kroner resulterer i laserskyer lokaliserte høyre oppe i kronene. Derfor, like ALS-verdier utledet fra laserhøydeverdier av gran, furu og lauvtrær vil gi høyere predikerte biomasseverdier for gran. Lik tendens er bitt reportert for tømmervolum (Næsset, 2004 a).

I min oppgave er utvelgelsen av variablene for å finne en enkel modell (steg 1), for forklaring av volumet, nesten lik som i Næsset & Gobakken (2008). Forskjellen er at jeg har brukt statistisk programmet R og variabelutvelgelsen er basert på ”*stepwise both directions*” og *AIC* kriteriet. Jeg har ikke behandlet mine treslagsandeler (0,5 og 0,8 for gran, furu og blanding) som kontinuerlige variabler. Dette medfører at det er vanskelig å tolke resultatene fra de ulike modeller fra treslag. Det er en uheldig kombinasjon av ulik prøveflateantall for utvikling av modeller samt ulik kontrollflate antall som gjekk inn for de ulike treslagsandeler som gjør tolkningen av resultater vanskelig. Dette kommer mer til uttrykk ettersom denne modellbasen er laget bare på områdenivå ($M.O_{0,5-0,8}$) og bare på stratum ($M.S_{0,5-0,8}$) nivå der vi har forskjellige prøveflate antall til modellutvikling og ikke minst kontrollflateantall til validering av de utvikla modeller.

Ved å tilføre treslagsandelen som kontinuerlige variabler vil vi sannsynligvis oppnå forbedringseffekt på de utvalgte regresjonsmodellene som i Næsset & Gobakken (2008), der regresjons modell med to ALS-forklaringsvariabler for biomasse over bakke gjekk fra 0,82 til 0,88 ved å legge til flere variabler som representerte forskjellige områder, hogstklasser og treslagsandeler.

9 KONKLUSJON

Det er kjent at sjølve måten vi stratifiserer på avhenger i stor grad av økonomi. Dette gjelder både i forskning, i planfirmaene og hos skogeieren. Inndeling i strata er ofte også styrt etter ønsket om god nøyaktighet for volumestimatet i den mest verdifulle skogen.

I lasertaksering ser en den positive effekten av stratifisering. Dette gjelder først og fremst skillen mellom lauv og bartrær der det er vesentlige forskjeller imellom.

En forklaring på dette er forskjellige kroneformer, andre forklaringer er refleksjonen av laserlyset fra nåler og løv som igjen er sterkt påvirket av penetreringsgraden til laserpulsene. Bestand som har stor lauvandel blir overestimert og bestand som har stor andel bartrær blir underestimert. På grunn av dette er det nødvendig å avskille treslag der det er høyt lauvinnslag. Dette vil også være et av problemene vi vil få med å slå sammen flere områder.

For å skille bestandsparametrer mellom barbestand og lauvbestand er det også nødvendig med videre stratifisering. Andre studier har også vist at forskjellige treslag forårsaker forskjeller i fordeling av laserdata (laserpulser) og derfor skal en betrakte de adskilt (*Næsset, 1997a*).

Ulike skogbehandlinger påvirker også forholdet mellom bestandsparametrer og fordelingen av laserpulser. Eksempel på dette er at lauv trea reagerer kraftigere på tynninger, der kronestørrelsen øker med større lystilgang, mens bar trær reagerer saktere til tynninger. Akkurat det forholdet mellom kronestørrelsen og stammevolum er relativ ufleksibel gjennom livsløpet til trea og vi får nøyaktigere estimering av barbestand enn lauvbestand. Dette er et argument til på viktigheten av stratifisering.

Det skal også være klart at når man ser på ulike treslag i fra flere områder så skal dette behandles med kontinuerlige variabler for det er bare slik vi kan forbedre effekten av ulike treslag for forskjellige områder og modellbaser. Med to modellbaser $M.O_{0,5-0,8}$ og $M.S_{0,5-0,8}$, slik det er gjort i oppgaven, vil vi likevel ikke oppnå noe bedre effekt i nøyaktigheten siden felles modeller på område og stratum nivå blir då påvirket i stor grad igjen av selve prøveflateantallet, stratum forskjeller i områdene ($M.O_{0,5-0,8}$), område forskjeller mellom stratum ($M.S_{0,5-0,8}$) og ikke minst blir valideringen av disse modeller et dårlig grunnlag siden det også er stor forskjell i mellom antall tilgjengelige kontrollflater.

Inndeling i strata er et viktig tema når teknologien nå har kommet så langt at man begynner å scanne hele land (Sverige og Finland).

Det som kan være en løsning på å sammenslå forskjellige områder er å ta i bruk *full-waveform* laser skannere fordi disse skannere vil fange opp mer informasjon om skogstruktur enn de skannere som er brukt i dag.

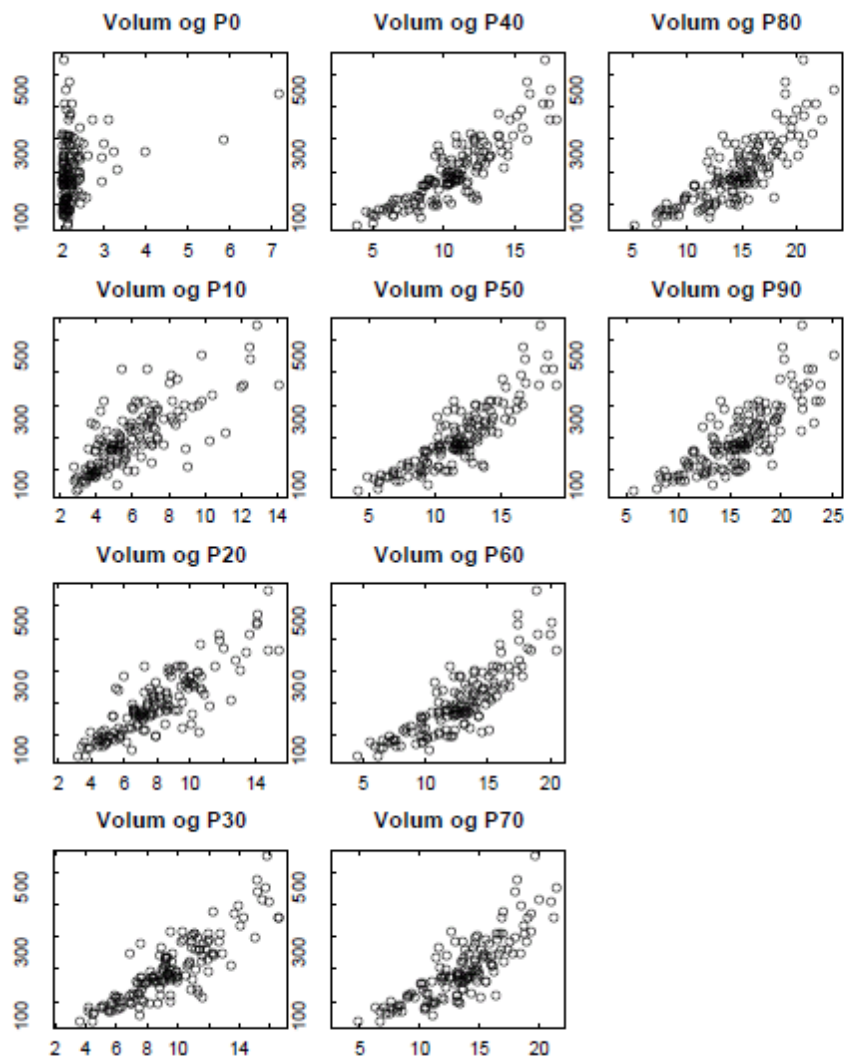
Dersom statistiske forhold mellom bestandsvariabler og laservariabler forblir forholdsvis konstante fra et område til en annen, kan det være mulighet til å gjenbruke tidligere innsamla data og så på denne måten redusere behovet for ytterligere felt arbeid. I denne studie var det funnet at områdene har signifikant effekt og dette bør indikere at store klimatiske, terreng variasjonen og hvordan trær utvikler seg i formen, for områdene påvirker i stor grad estimeringen av volum og andre utleda variabler. Dette bør derfor tas i betraktning dersom man vurderer å slå sammen takstområder.

10 REFERANSER

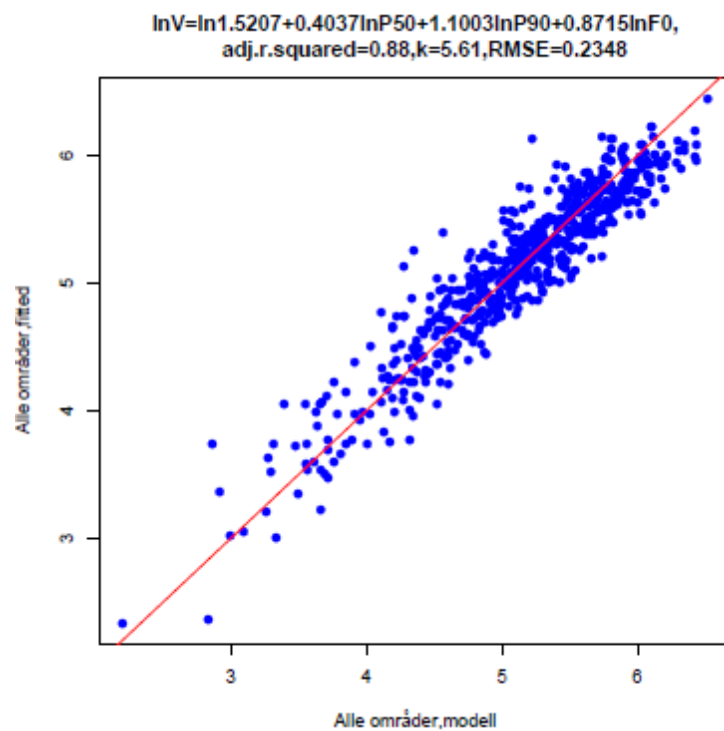
1. Aldred, A.H. and Bonnor, G.M., 1985. Application of airborne lasers to forest surveys. Information Report PI-X-51, Canadian Forestry Service, Petawawa national Forestry Institute, 62 p.
2. Bollandsås, O.M. & Næsset, E. 2007. Estimating percentile-based diameter distributions in uneven-sized Norway spruce stands using airborne laser scanner data. *Scandinavian Journal of Forest Research*, 22(1), 33-47.
3. Brantseg, A. 1959. *Stand volume functions for pine*. Norsk skoghåndbok. (In Norwegian).
4. Brantseg, A. 1967. *Volume functions and tables for Scots pine*. South Norway. Communications of Norwegian Forest Research Institute, 22, 695-739. (In Norwegian with English summary).
5. Braastad, H. 1966. *Volume tables for birch*. Communications of Norwegian Forest Research Institute, 21, 265-365. (In Norwegian with English summary).
6. Flewelling, J. & Pienaar, L. 1981. Multiplicative regression with lognormal errors. *For.Sci.*27:281-289
7. Holte, A. 1993. *Diameter distribution functions for even-aged (Picea abies) stands*. Communications of Norwegian Forest Research Institute 46(1), 1-46.
8. Hyypä, J., Yu, X., Hyypä, H. & Maltamo, M. 2006. Methods of Airborne Laser Scanning for Forest Information Extraction. *Workshop on 3D Remote Sensing in Forestry*, 14-15 Feb 2006, Vienna.
9. Heldal, J. 2006. Logisk regresjon-kurskompendium i byråskolens kurs SM 507, s.17, 2006/54
10. Kraus, K. & Pfeifer, N. 1998. Determination of terrain models in wooded areas with airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing* 53,193-203.
11. Maclean, G., Krabill, W., 1986. Gross-merchantable timber volume estimation using an airborne lidar system. *Canadian Journal of Remote Sensing* 12, pp. 7-18.
12. Mønness, E. 1982. *Diameter distributions and height curves in even-aged stands of Pinus sylvestris L.* Communications of Norwegian Forest Research Institute 36 (15): 1-43.
13. Nelson, R., Krabill, W. and Maclean, G., 1984. Determining forest canopy characteristics using airborne laser data, *Remote Sensing of Environment* 15, pp. 201-212.
14. Næsset, E. 1995. Stand volume functions for Picea abies in Western Norway. *Scandinavian Journal of Forest Research* 10, 42-50.
15. Næsset, E. 1997a. Determination of mean tree height of forest stands using airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing* 52, 49-56.
16. Næsset, E. 1997b. Estimating timber volume of forest stands using airborne laser scanner data. *Remote Sensing of Environment* 61, 246-253.
17. Næsset E. & Tveite, B. 1999. Stand volume functions for Picea abies in Eastern, Central and Northern Norway. *Scandinavian Journal of Forest Research* 14, 164-174.
18. 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, 328-340.

19. Næsset, E. 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sensing of Environment*, 80, 88–99.
20. Næsset, E. 2004a. Practical large-scale forest stand inventory using a small-footprint airborne scanning laser. *Scandinavian Journal of Forest Research*, 19, 164–179.
21. Næsset, E. 2004b. Accuracy of forest inventory using airborne laser-scanning: Evaluating the first Nordic full-scale operational project. *Scandinavian Journal of Forest Research* 19, 554–557.
22. Næsset, E. 2004c. Effects of different flying altitudes on biophysical stand properties estimated from canopy height and density measured with a small-footprint airborne scanning laser. *Remote Sensing of Environment*, 91, 243–255.
23. Næsset, E. 2004d. Estimation of Above- and Below-Ground Biomass in Boreal Forest Ecosystems. In M. Thies, B. Kock, H. Spiecker & H. Weinacker (Eds.), *Laser-scanners for forest and landscape assessment*. (pp. 145–148). Freiburg, Germany: International society of photogrammetry and remote sensing. *International archives of photogrammetry, Remote sensing and spatial information sciences*
24. Næsset, E., Bollandsås, O. M., & Gobakken, T. 2005. Comparing regression methods in estimation of biophysical properties of forest stands from two different inventories using laser scanner data. *Remote Sensing of Environment*, 94, 541–553.
25. Næsset, E. 2007. Airborne laser scanning as a method in operational forest inventory: Status of accuracy assessments accomplished in Scandinavia. *Scandinavian Journal of Forest Research*, 22, 433–442.
26. Næsset, E. & Gobakken, T. 2008. Estimation of above- and below- ground biomass across regions of the boreal forest zone using airborne laser. *Remote Sensing of Environment*, 112, 3079–3090
27. Næsset, E., Gobakken, T., Holmgren, J., Hyypä, H., Hyypä, J., Maltamo, M., Nilsson, M., Olsson, H., Persson, Å. & Söderman, U. 2004. Laser scanning of forest resources: The Nordic experience. *Scandinavian Journal of Forest Research*, 19, 482–499.
28. Solodukhin, V., Zukov, A., Mazugin, I., 1977. Possibilities of laser aerial photography for forest profiling, *Lesnoe Khozyaisto* (Forest Management) 10, pp. 53–58 (in Russian).
29. Tveite, B. 1977. *Site index curves for Norway spruce (Picea abies (L.) Karst.)*. Report of Norwegian Forest Research Institute, 33, 1–84. (In Norwegian with English summary).
30. Vestjordet, E. 1959. *Stand volume functions for spruce and birch*. Norsk skoghåndbok. (In Norwegian).
31. Vestjordet, E. 1967. *Functions and tables for volume of standing trees. Norway spruce*. Communications of Norwegian Forest Research Institute 22, 543–574. (In Norwegian with English summary).
32. Vestjordet, E. 1972. *Diameter distribution and height curves for even-aged stands of Norway spruce*. Communications of Norwegian Forest Research Institute 29(8), 469–557.
33. Wehr, A., Lohr, U., 1999. *Airborne laser scanning – an introduction and overview*. *ISPRS Journal of Photogrammetric and Remote Sensing* v. 54, pp. 68–82.

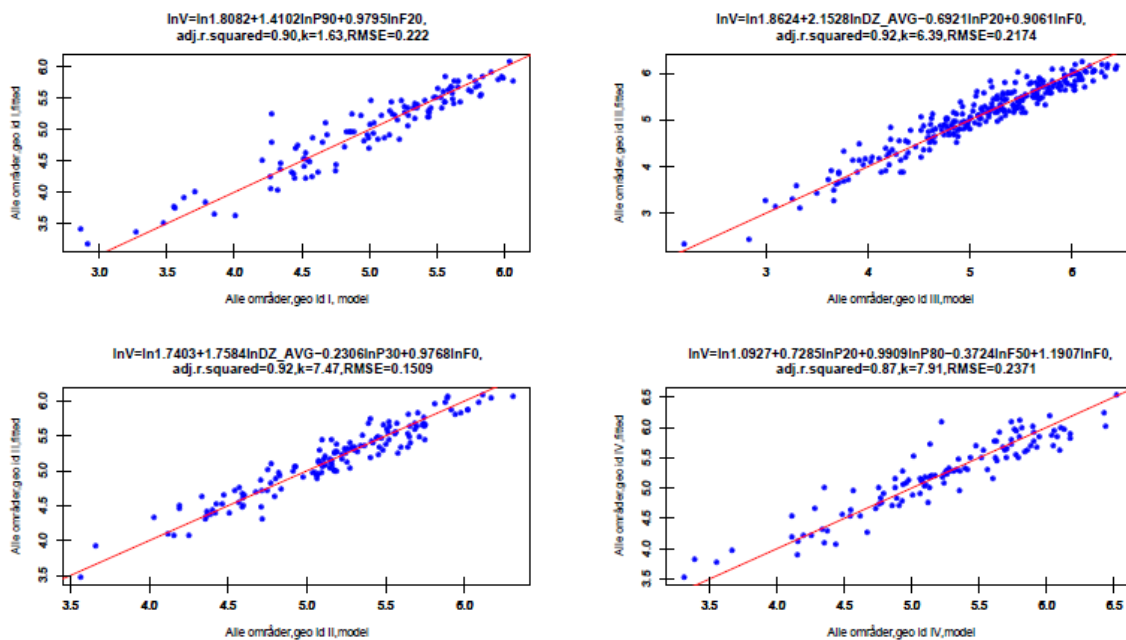
11 VEDLEGG



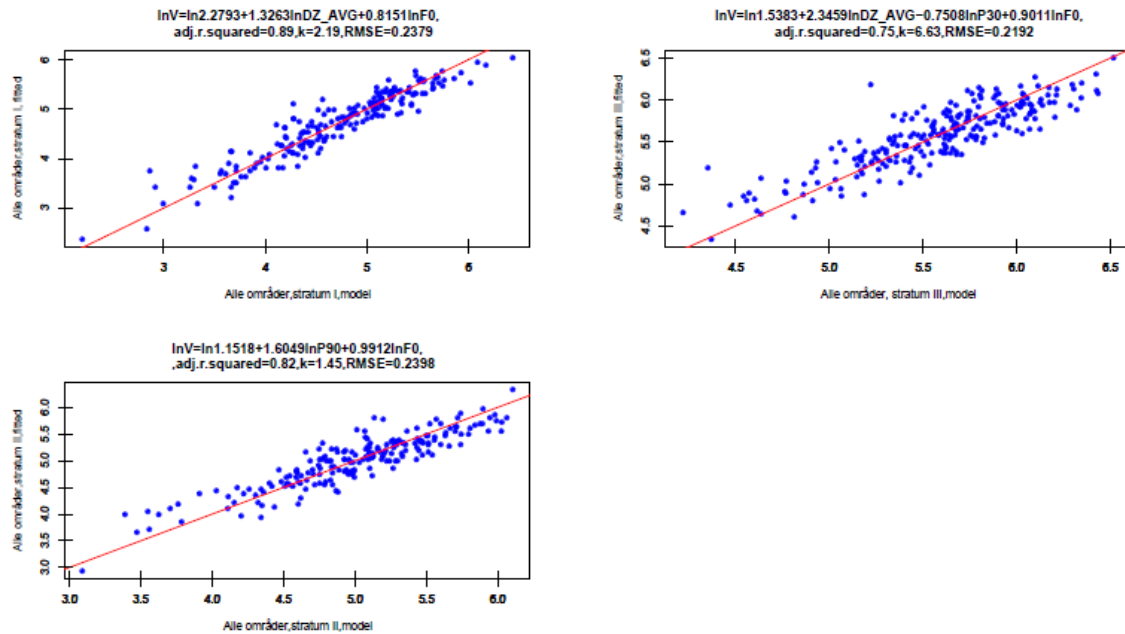
GRAF 1:FORHOLDET MELLOM VOLUM OG PERSENTILER I ORDINÆRSKALA



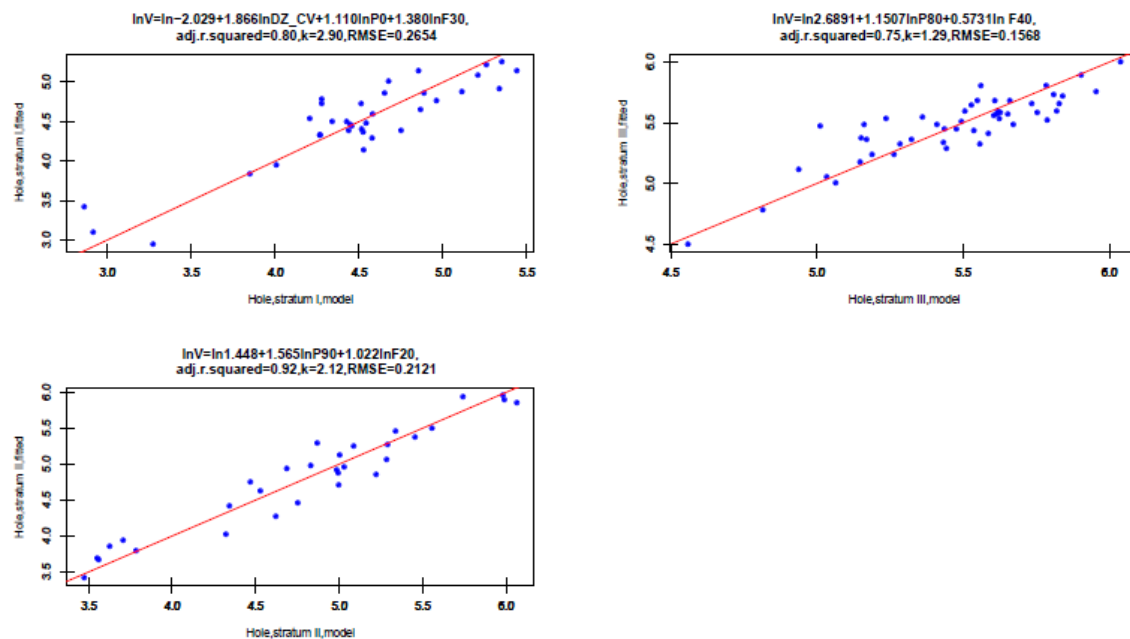
GRAF 2: MODELL FOR ALLE FIRE OMRÅDER



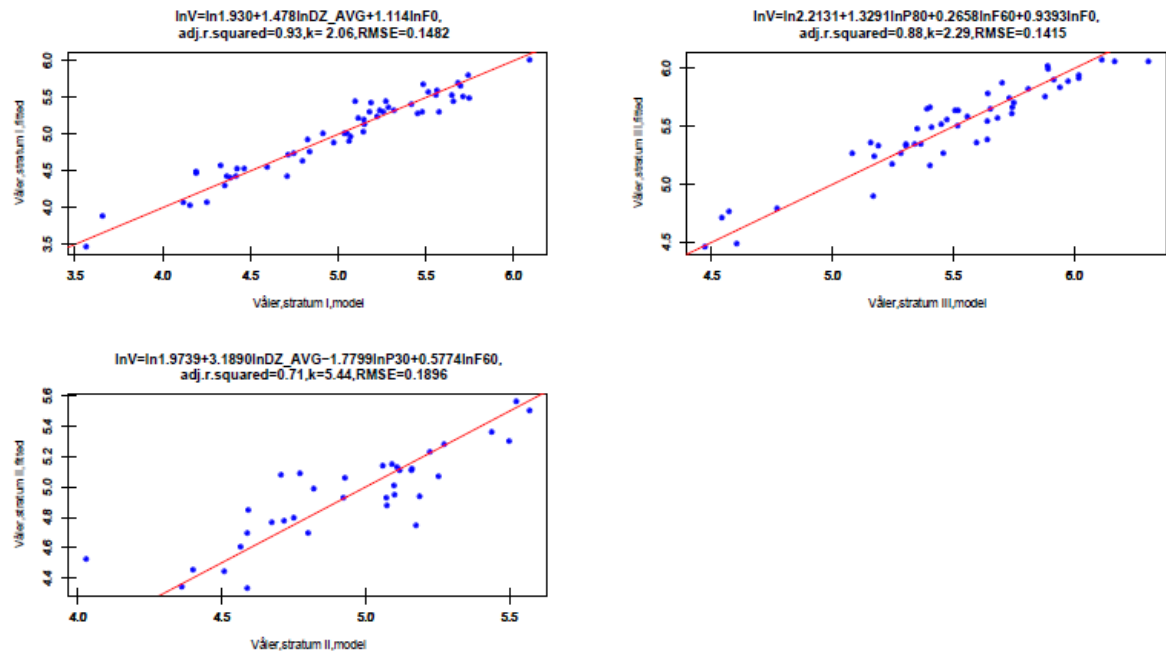
GRAF 3: MODELL FOR HVERT OMRÅDE UAVHENGIG AV STRATUM



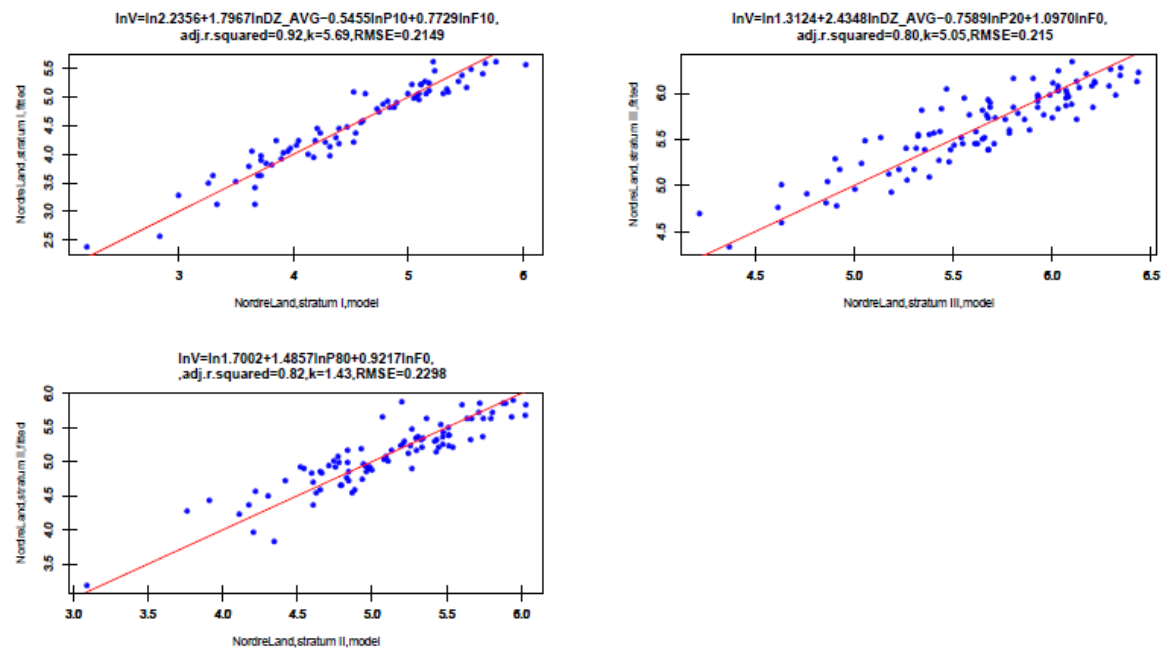
GRAF 4: MODELL FOR HVERT STRATUM UAVHENGIG AV OMRÅDE



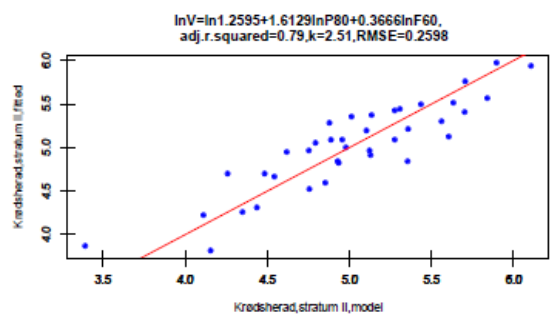
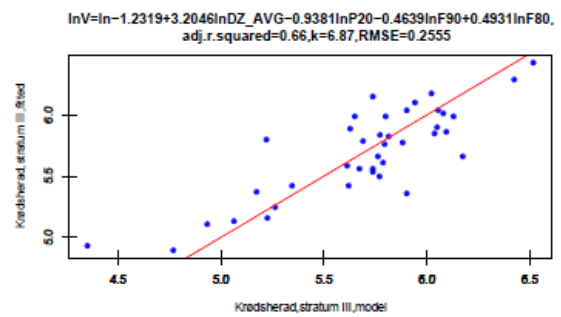
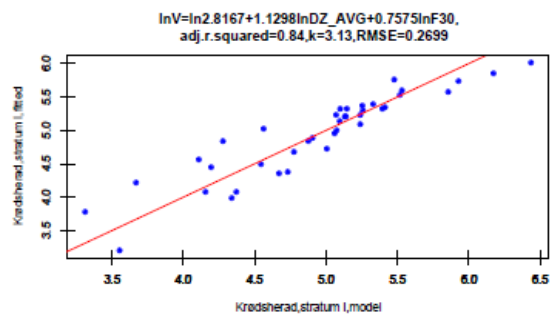
GRAF 5: MODELL FOR HVERT OMRÅDE PER STRATUM- MODELL FOR HOLE



GRAF 6: MODELL FOR HVERT OMRÅDE PER STRATUM- MODELL FOR VÅLER



GRAF 7: MODELL FOR HVERT OMRÅDE PER STRATUM- MODELL FOR NORDRE LAND



GRAF 8: MODELL FOR HVERT OMRÅDE PER STRATUM- MODELL FOR KRØDSHERAD