

DETEKSJON AV STORE FURUTRÆR BASERT PÅ FLYBÅREN LASER

DETECTION OF BIG SCOTS PINE TREES USING AIRBORNE LASER

TORGRIM ØSTGÅRD

UNIVERSITETET FOR MILJØ- OG BIOVITENSKAP
INSTITUTT FOR NATURFORVALTNING
MASTEROPPGAVE 30 STP. 2010



Forord

I forbindelse med oppgavevalg våren 2009 kom jeg i kontakt med en lokal bedrift i Tolga kommune, Materialbanken AS med kontorer på Røros og i Vingelen. De har sin spisskompetanse på materialkunnskap og er leverandør av forskjellige produkter i malma furu og såkalte bokser i grove dimensjoner. I en samtale med daglig leder, Jacob Trøan, ble behovet for en database med informasjon om trær med store dimensjoner understreket, noe som per dags dato ikke finnes. Som leverandør av blant annet produktspesifikke laftestokker ville en slik database være et viktig hjelpemiddel for effektiv ressurskartlegging og planlegging for uttak av ressursene.

Flybåren laserscanning og dets muligheter innenfor skogbruket synes å være mangfoldige. Å koble laserdata opp mot funn av store dimensjoner er tidligere ikke blitt gjort etter det jeg kjenner til.

Veileder Vegard Lien fortjener en stor takk for hjelp og for gode og pedagogiske forklaringer underveis i prosessen. Det samme fortjener min hovedveileder, Terje Gobakken, som har vært behjelpelig med raske og konstruktive tilbakemeldinger. I forbindelse med den feltmessige delen vil jeg også få takke min far og mor på Tolga for transport og feltmessig hjelp som effektiviserte takseringen av laserkorridorene.

Torgrim Østgård

Ås 18. mai 2010

Sammendrag

En etablering av en database med opplysninger om kvalitetsvirke, spesielle dimensjoner og lokaliseringen av slike ressurser er ønsket fra flere av skognæringens aktører og interessenter i det offentlige forvaltningsapparatet. Denne oppgaven fokuserer på beregning av et områdes sannsynlighet for forekomst av furutrær med store dimensjoner. Beregningene er basert på bruk av flybåren laser i Tolga kommune i Hedmark fylke. Laserdataene ble samlet inn i 2006 i forbindelse med et forskningsprosjekt på UMB.

Feltdatamaterialet ble samlet inn oktober 2009 og består av i alt 801 furutrær fordelt på 1112 prøveflater á 1000 m². Ulike laservariable er beregnet for de enkelte prøveflater. Sannsynligheten for forekomst av minst ett stort tre innenfor en enkelt prøveflate er estimert basert på, variablene avledet fra laserregistreringer samt eksterne variabler avledet fra kartdata.

Ved hjelp av stegvis regresjonsprosedyre ble det laget tre ulike modeller for forekomst av furuer i diameterklassene >350 mm, >400 mm og >500 mm. Variabelutvelgelsen ble gjort på grunnlag av lav AIC (*Akikes Information Criterion*) og signifikansnivå på $p < 0.05$. Modellene ble brukt til prediksjon av sannsynligheten for forekomst av store dimensjoner i alle besøkte flater. I tillegg ble det laget modeller basert på utvalgte variabler for de samme diameterklasser og for gjennomgang av de nevnte prediksjonsprosedyrene. Felles for disse modellene er lav multikollinearitet.

Valideringen av regresjonsmodellene ble gjort gjennom en Monte Carlo-simulering med 75 % estimeringsandel og 25 % til validering. Totalt 300 gjentak ble gjennomført for hver diameterklasse. Andelen prøveflater hvor prediksjonen stemmer overens med feltobservasjoner av store trær er 63 % 45 % og 14 % for henholdsvis diameterklassene >350, >400 og >500. Simuleringen viste seg i etterkant å ha konverget mot korrekte verdier allerede etter omtrent 100 gjentak. Variasjonen mellom resultatene for de ulike iterasjoner var liten.

Monte Carlo-simulering med ulik estimeringsandel fra 75 % og nedover til 5 % av flatene ble utført for å undersøke betydningen av størrelsen på feltdatamaterialet. Totalt 500 gjentak ble gjennomført for hver femte prosentdel og 25 % av datamaterialet ble benyttet til validering. Resultatet viste et jevnt standardavvik inntil 30 % estimeringsandel hvor økningen ble markant. Dette resultatet innebærer at predikeringen for funn av store dimensjoner kunne vært estimert med lik nøyaktighet for 30 % av totalt antall flater. Mine metoder og resultater kan brukes som bakgrunnsmateriale ved lignende undersøkelser. Høydemåling og tretetthet er dog andre viktige faktorer som bør undersøkes nærmere.

Abstract

An establishment of a database containing information about quality timber, specific dimensions and location of such resources is required. This thesis focuses on the calculation of an area's likelihood of occurrence of Scots pine trees with large dimensions. Calculations are based on airborne laser data in the municipality Tolga in Hedmark County, Norway. The airborne laser data was collected in 2006 in connection with a research project at UMB (Norwegian University of Life Sciences).

Field data was collected in October 2009 and consists of 801 Scots pine trees over a total of 1112 sample plots, each one 1000 m². Different laser variables are calculated for individual sample plots. The probability of occurrence of at least one large tree in a single sample plot is estimated based on variables derived from laser data as well as external variables derived from map data.

Using stepwise regression procedure it was produced three different models for the occurrence of pine trees in diameter classes >350mm, >400mm and >500 mm. Variable selection was done on the basis of low AIC (*Akaike's Information Criterion*) and a significance level of $p < 0.05$. These models were used for predicting the probability of the occurrence of large dimensions in all visited sample plots. In addition, models based on selected variables were created for the same diameter classes and used in the aforementioned prediction procedures. Common to these models is low multicollinearity.

Cross-validation of the regression model was made through a Monte Carlo simulation randomly selecting 75 % of the sample plots. A total of 300 replicates were conducted for each diameter class. The proportion of sample plots where the field observations of large trees matches the prediction of the models are 63% 45% and 14% in diameter classes >350mm, >400mm and >500mm respectively. The simulation proved subsequently to have converged to correct values already after about 100 replicates. The variation between the results for the various iterations was small.

Monte Carlo simulations randomly selecting 75% down to 5% of the plots were carried out to examine the significance of the size of the field data. A total of 500 replicates were conducted for every fifth percentage and constantly 25% of the data was used for validation. The results showed a consistent standard deviation down to a random selection of 30%, where the increase became significant. This result implies that the prediction of occurrence of big trees could be estimated with the same sample error using 30% of the total number of sample plots. My methods and results can be used as background material by similar surveys. Height measurement and stem number is, however, other important factors that should be examined more closely.

Innholdsfortegnelse

1. INNLEDNING	1
1.1 INTRODUKSJON FLYBÅREN LASERSCANNING	1
1.2 FLYBÅREN LASERSCANNING I SKOGBRUKET.....	3
1.2.1 Flybåren laserscanning – Diameterundersøkelser	4
2. MATERIALE OG METODE	5
2.1 STUDIEOMRÅDET	5
2.1.1 Skoghistorikk i Tolga.....	5
2.2 LASERDATA	6
2.3 FELTARBEIDET OG REGISTRERINGSUTSTYR.....	7
2.4 BEARBEIDING AV DATAMATERIALE	8
2.5 LASERVARIABLER.....	9
2.5.1 Høydepersentiler - PF-verdier	10
2.5.2 Tetthetspersentiler - TF-verdier	10
2.6 EKSTERNE VARIABLER.....	10
2.7 STATISTISK ANALYSE	12
2.7.1 Korrelasjon mellom variabler og multikollinearitet i modeller	12
2.8 MONTE CARLO-SIMULERING OG KRYSSVALIDERING	13
2.9 FELTDATAMATERIALETS STØRRELSE	13
3. RESULTATER.....	15
3.1 SLUTTMODELLER.....	16
3.2 PREDIKERING AV SANNSYNLIGHETEN FOR FOREKOMST AV STORE TRÆR	17
3.3 FELTDATAMATERIALETS STØRRELSE	19
4. DISKUSJON.....	21
4.1 MATERIALE OG METODE.....	21
4.1.1 Feltdatamaterialet	21
4.1.2 Feltarbeidet	21
4.2 LASERKORRIDOREN – MONTE CARLO-SIMULERING OG MULTIKOLLEARITET	22
4.3 LASER- OG EKSTERNE VARIABLER I SLUTTMODELLENE.....	23
4.2.1 FELTDATAMATERIALETS STØRRELSE	26
4.3 OPPGAVEN SOM VERKTØY FOR VIDERE UNDERSØKELSER	26
4.3.1 Grafisk sluttresultat – et tenkt eksempel.....	27
5. KONKLUSJON	29
6. LITTERATUR	30
VEDLEGG	

1. Innledning

Riksantikvaren har opp mot 10 000 bygninger som skal opp på et normalt vedlikeholds nivå innen 2020 (pers. medd. Andresen 2010) Dette omfatter så vel private som mottar tilskudd fra det offentlige. Dette vil kreve blant annet materialer av kvalitetsvirke til restaureringsformål og felles for disse materialene er at lengder, dimensjoner og kvalitet er utover dagens standard og dermed ikke tilgjengelig hos trelastutsalgene. Om de politiske målsettingene skal nås, er det forespeilet en bruk av oppimot 50.000 m³ årlig innen 2020 (Venn 2010a). Det arbeides i skrivende stund med å lage materialister over hva hvert bygg vil trenge av materialer. En arbeidsgruppe pekte på under konferansen "Skog for vern" i 2008 etablering av en nettbasert søkebase for anskaffelse av kvalitetsvirke hvor det er mulighet for å legge ut forespørsler og gjøre bestilling etter særskilte behov. En viktig faktor i bakgrunnen for disse planene er anledning for plukkhogst i skogteiger med rett virkeskvalitet, slik at spesialvirke vil kunne tas ut når behovet melder seg.

Trevekst Oppland er navnet på et treårig prosjekt som startet opp i januar 2009. Målet med prosjektet er å fremskaffe, foredle og omsette kvalitetsmaterialer til tradisjonshåndverk og bygningsvern, kort sagt til bruk der det trengs noe spesielt (Venn 2010b). I Hedmark har firmaet Materialbanken AS vært i drift i flere år og firmaet har sin spisskompetanse på materialkunnskap og er leverandør av forskjellige produkter i malma furu og trær i grove dimensjoner. Både prosjektet i Oppland og Materialbanken AS har det samme ønsket om en database for spesielle trær i skogen. Prosjektet i Oppland har allerede etablert en database for 5000 trær i Etnedal som kan brukes til ulike formål. Datafangsten har vært basert på feltundersøkelser av skogtaksatorer.

1.1 Introduksjon flybåren laserscanning

Flybåren laserscanning, også kalt ALS (*Airborne Laser Scanning*) er en form for laseravstandsmåling, LiDAR (*Light Detection and Ranging*). Etter de første studiene av ALS og LiDAR ble gjort på midten av 1990-tallet i Norge er metoden blitt et viktig verktøy i forbindelse med registrering av blant annet skoglige parametre. Lasersystemenes teknologiske utvikling siden den gang har skutt fart både med tanke på kapasitet og nøyaktighet. Måleresultater kan fås med en nøyaktighet på et par desimeter (Næsset *et.al* 2004). Fjernmålingsmetoden er ikke særegen kun for skogsektoren, men et kraftig verktøy med mange applikasjonsmuligheter innenfor mange disipliner.

LiDAR er en form for aktiv fjernmåling. I det ligger det at georefereringen foregår kontinuerlig som laserpulsene bombarderer landskapet og reflekteres slik at man kan vite hvor hver enkelt puls har truffet. Dette skiller den ifra den mest sammenlignbare fjernmålingsmetoden, fotogrammetri, hvor georefereringen foregår i ettetid av registreringen.

Kort forklart baserer teknikken seg på et bombardement av korte høyenergiske lyspulser mot bakken hvorpå de reflekteres av objektet som bryter lysbanen. De såkalte ekkopulsene fanges opp av en sensor på laserinstrumentets mottakerdel om bord i flyet.

Lasermåling opererer på et prinsipp som sier at hvis geografisk lokalisering og orientering av laserinstrumentet er kjent, samt laserpulsens retning og den tidsmessige avstanden fra emittering til registrering, kan man kalkulere treffpunktet på bakken eller trekronene for hver laserpuls. Den tidsmessige avstanden tar utgangspunkt i lysets hastighet og regner ut avstanden mellom objektet som fremkalte refleksjonen av pulsen, og flyet. Resultatet er en punktsverm med utgangspunkt i en x-, y- og z-akse som vil danne et 3D bilde hvor man, avhengig av blant annet skogtyper og laserscanningsmetoder, kan definere hvert enkelt tre. Er punkttettheten i 3D-bildet tilstrekkelig høy, er det utviklet forskjellige algoritmer som kan identifisere enkeltrær eller grupper av trær for deretter å kunne estimere skoglige parametre som hvert tre representerer (Popescu *et.al* 2004). Nøyaktigheten til hvert punkt i 3D-bildet avhenger hovedsakelig av avstanden laserstrålen tilbakelegger (avstanden fra flyet til bakken), posisjonen til laserstrålen (hvor flyet nøyaktig befinner seg da pulsen sendes) og i hvilken vinkel den sendes i forhold til flyet (Baltsivas 1999).

Hvert punkt som laserpulsene produserer må georefereres. En 3D fremstilling av punktene forutsetter at posisjoneringen og orienteringen av lasersystemet om bord i flyet georefereres i forhold til et koordinatsystem, f.eks WGS84. Geokodingen krever full synkronisering mellom scanneren og den integrerte posisjonsdelen bestående av IMU (treghetsnavigasjon) og GPS. IMU registrerer flyets vinkler til enhver tid i luften slik at f. eks et vindkast som forårsaker at flyet krenker litt til den ene siden, vil kunne fanges opp av IMU-instrumentene om bord i flyet. Måleinstrumentene består av akselerometere og gyroer for registrering av akselerasjoner og vinkelhastigheter.

Rent praktisk foregår ALS ved at et speil i en sveipemekanisme sørger for at laserpulsene blir spredt på tvers av flyretningen for å oppnå en scannings-effekt. Ulike sveipemekanismer gir ulike typer punktmønstre, mest brukt er oscillerende speil som gir en Z-mønster på bakken. Hvor stor scanningsbredden blir, det vil si fra ytterkant til ytterkant av scannet, avhenger av flyets høyde og åpningsvinkelen til scanneren. Til operasjonell bruk her i Norge brukes det vanligvis en vinkel på rundt 20 grader. Frekvensen til både sveipemekanismen og laseren bestemmer hvor tett punktene treffer i forhold til hverandre og flyets framdrift sørger dermed for arealdekningen.

Laserstrålens "*footprint*", eller *fotavtrykket*, er et sentralt begrep innen terminologien og er viktig når det gjelder å forstå prinsippet bak datafangsten for skoglige formål. *Fotavtrykket* er laserstrålens utbredelse, og denne vil variere med innfallsvinkel, avstand fra flyet og ikke minst strålens divergens. Innenfor ALS brukes hovedsakelig verdier mellom 0,3-2 mrad (Wehr & Lohr 1999). En puls som faller vinkelrett på bakken vil oppnå en sirkulær form sammenlignet med pulser som er sendt ut i større vinkler som vil danne en elliptisk form. Dette avhenger av laserstrålens diameter og innfallsvinkel på objektet samt terrengets utforming. Alt dette følger samme prinsipp som når man holder en lommelykt vinkelrett på et ark. Lysstrålen blir kraftig og konsentrert om et visst areal, mens en liten vinkelendring vil føre til at en større arealenheter blir opplyst, dog litt svakere. En stor divergens på strålen vil føre til et stort *fotavtrykk* og dette vil i verste fall ikke gi noen refleksjon av signalet fordi energitapet blir for stort. Energitalp er akilleshælen for rekkevidden til ALS, i tillegg til de forskjellige faktorene ved instrumentet. Typisk for *small-footprint* systemer er et fotavtrykk som varierer fra noen få centimeter til et par meter, mens *large-footprint* tilsier ca 10 meter og oppover (Næsset *et al.* 2004).

Refleksjonen av en laserpuls, ved en gitt bølgelengde, avhenger av egenskapen til objektet. Den reflekterende egenskapen til naturlig vegetasjon ligger mellom 30-50 % (Wever 1999). Desto lysere overflate jo bedre kvalitet på ekkopulsen. ALS bruker hovedsakelig bølgelengder innenfor det optiske spekteret fra 800 – 1600 nm (Wehr & Lohr 1999). Dette er infrarød stråling som innehar den egenskapen at den reflekteres og ikke absorberes når den treffer bakken.

I tillegg til pulsbaserte systemer finnes det også kontinuerlige lysbølger (*Continuous Wave-Laser*), men denne oppgavens datamateriale baserer seg på et pulsbasert system. Det er ikke mulig å få en helhetlig vertikal dekning av trekronene ved bruk av pulsbaserte lysbølger, i motsetning til kontinuerlige lysbølger.

1.2 Flybåren laserscanning i skogbruket

Laserens evne til å registrere en eller flere retursignal fra samme lyspuls gjør nettopp denne metoden svært egnet for måling i skog, men denne måleegenskapen avhenger av hvilken type sensor man bruker i datafangsten. Divergensen gjør at det blir registrert flere objekter innenfor laserstrålens bane. Det kan være to eller flere objekter, f. eks toppen av treet, en grein like nedenfor som gjør at flere returpulser blir generert og noe av laserlysets *footprint* kan også treffe bakken som dermed blir den siste returen som fanges opp. Det er blant annet fra siste returene som oftest treffer bakken at man kan utarbeide DTM-kart (*Digital Terrain Modell*). Bakkenivået må være kjent for at man kan subtrahere laserpuls høyden for å kunne beregne trehøyden over bakken.

I ALS-datafangst kan hvert tre sammenlignes med en populasjon, hvor laserpulsene er eneste verktøy som gir oss informasjon om populasjonen, hovedsakelig fra trekrona. Avhengig av blant annet skogbilde, kronestruktur og sesong vil bare deler av trekronene kunne visualiseres av rådataene fra laserscanning. Laserpulsene kan spesielt om sommeren kun bli reflektert ifra blad- og nålverket. Penetreringsraten blir dermed veldig lav og få pulser vil treffe bakken. Kilian *et al.* (1996) har påvist en penetreringsrate på 30 % om sommeren i barskog. Best resultat fås når atmosfæren er kald, tørr og klar og natta er å foretrekke framfor en solrik dag (Baltsavias 1999).

Lasertakstmetoden til skogbruksformål deles inn arealbasert metode og enkelttremetode. I de nordiske landene er bestandsskogbruket den rådende skogbehandlingen hvor fokuset er på en gitt arealenhet, typisk 1-10 ha, hvor gjennomsnittsverdier fra skoglige parametre legger grunnlaget for ressursbeskrivelsen (Næsset *et al.* 2004). De første norske undersøkelser av dette er gjort av Næsset (1997a,b) som tok for seg henholdsvis gjennomsnittlig trehøyde vektet av grunnflatesum og estimering av bestandsvolum. Informasjon om skoglige parametre kan også fås fra arealbaserte takseringer som dessuten krever mindre arbeid når det gjelder datainnsamlingen, og det er påvist lovende resultater ved en punkttetthet på 1 puls per m² (Næsset *et al.* 2004).

Data på enkelttrenivå krever høy grad av nøyaktighet, med andre ord en høy punkttetthet per m². Næsset *et al.* (2004) og Persson *et al.* (2002) har forespeilet en tetthet på 5-10 laserpulser per m² i bartrebestand, noe som fører til en dyr taksering. Nøyaktigheten forbedres også av *small-footprint* systemer som fanger opp treet utseende bedre.

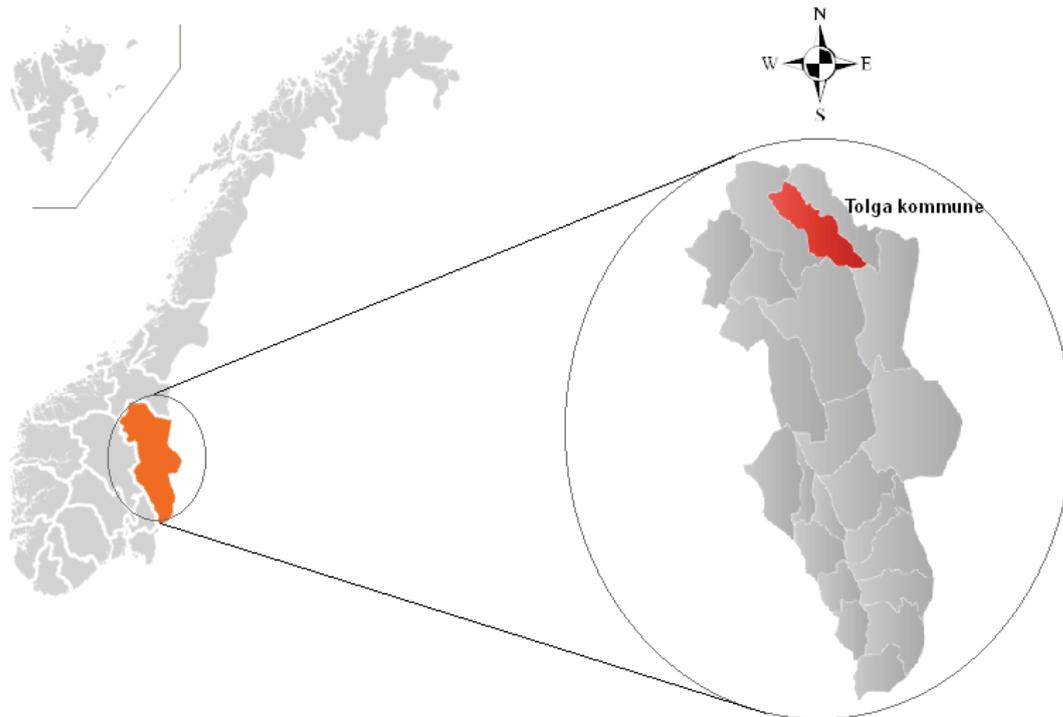
Forskningen som er blitt gjort på dette fagfeltet og den teknologiske utviklingen de seneste år har demonstrert at *small-footprint* systemer er i god stand til å fremskaffe estimater for ulike skoglige parametre. Basert på høydemålinger fra laserdata har man kunnet estimere andre viktige skoglige parametre som f. eks gjennomsnittshøyder, grunnflatesum og bestandsvolum med en nøyaktighet på lik linje eller bedre sammenlignet med konvensjonelle datafangstmetoder som blant annet fotogrammetri (Næsset *et al.* 2004). Disse prosedyrene er nå implementert i konkrete operasjonelle skogtakseringer, den første tilbake til 2002.

1.2.1 Flybåren laserscanning – Diameterundersøkelser

Basert på rådataene fra laserflygningen får man en punktsky hvor hvert punkt har geografiske koordinater samt en høydeverdi. En stor punktetthet vil avbilde trekronas form i nær sammenheng med den naturlige formen, men stammediameteren til treet, i likhet med flere parametre, kan først utvinnes etter prosessering og etterbehandling av laserskya. Å kunne estimere diameterfordelingen er en viktig del av skogtakseringen siden den vil gi skogeier verdifull tilleggsinformasjon om skogen og optimalisere forvaltningen og er et viktig hjelpemiddel for å beskrive variasjonen i uensaldret skog. Korrelasjonen mellom stammediameter og kronediameter ved å måle kronestørrelsen ut fra flyfoto har tidligere blitt undersøkt med tanke på å kunne estimere stammediameteren (Jakobsons 1970). Noen undersøkelser ser på karakteristikk ved laserdata innenfor prøveflater for å estimere diameterfordelinger (Bollandsås & Næsset 2007 & Mehtätalo *et al.* 2007), mens andre bruker enkeltrær som målenivå. Etter det jeg kjenner til er det ingen som direkte har predikert sannsynligheten for å finne spesielt verdifulle store trær.

2. Materiale og metode

2.1 Studieområdet



Figur 1: Oversiktskart over studieområde, Tolga kommune i Hedmark fylke.

Studieområdet i dette arbeidet er utvalgte områder i Tolga kommune i Hedmark (Figur 1). I Tolga er det totalt 64 000 og 124 000 dekar produktiv skog fordelt på henholdsvis bar- og bjørkeskog. Når det gjelder total kubikkmasse er fordelingen omtrent 50/50 mellom bjørk og bartrærne hvor gran utgjør omtrent 2,5 % av sistnevnte. Alle skogbruksmessige tall er hentet fra Mjøsens planavdeling v/Øystein Neby. Tall fra 2007 og 2008.

2.1.1 Skoghistorikk i Tolga

Rørsverket ble anlagt i 1644 og har hatt stor betydning for skogbildet og dermed nåtidens forekomst av store dimensjoner i traktene rundt Røros. Verket fikk nærmest en enerett til skogsvirke innenfor sirkumferensen, deriblant Tolga, men det ble også etter hvert hogd utenfor sirkumferensen. Fra ca 1670 hvor Tolga fikk egen smeltehytte ble det hogd hardt i skogene på Tolga. I hele verkets område er forbruket estimert til å være i gjennomsnitt 50 000m³ trevirke årlig (Sandnes 1942). Behovet for trekull førte til denne rovhogsten og det gikk hardt utover alle dimensjoner selv om det utover 1700-tallet kom bestemmelser om at såkalt mindreverdige virke skulle brukes til kullbrenning. Verkets behov for skog ble betydelig redusert på slutten av 1800-tallet da krutt, steinkull og koks ble mer vanlig å bruke.

I denne sammenheng er det klart at flesteparten av de virkelig store dimensjonene må ha blitt hogd i løpet av denne perioden. Hvis man tar utgangspunkt i slutten av 1800-tallet at skogreisningen begynte som smått har flesteparten av trærne i teorien hatt omtrent 150 år å vokse på. Det må også nevnes at noen områder på Tolga er blitt innkronet som såkalte kronskogteiger og dermed spart for herjingene under verkets periode.

2.2 Laserdata

UMB gjennomførte i 2006 flygninger over Tolga kommune med laser i øst-vest retning i forbindelse med et forskningsprosjekt. Totalt ble det innenfor kommunegrensen taksert 8 striper med 6 kilometers mellomrom. Laserstripenes bredde vil i størst mulig grad avhenge av terrengets utforming, laserscannerens åpningsvinkel og flyhøyden, men vil i gjennomsnitt være omtrent 400-700 meter bred. For denne studien ble det i midten av hver stripe definert en 32 meter bred korridor som ble delt opp i 32 meter lange prøveflater tildelt sine respektive ID-numre. På forhånd var det avtalt å strebe etter et datamateriale med innslag fra alle deler av kommunen, men noen striper forble urørt av forskjellige grunner som blant annet: tidsaspektet, dårlige GSM-forhold og lokal kunnskap om områdene.

Landsskogtakseringens striper på Tolga ble laserscannet 6. og 8. august 2006. Laserflygningen ble utført av Blom Geomatics AS med fly av typen PA31 Piper Navajo og om bord et Optech laserinstrument. Gjennomsnittlig verdi av både diameter på *fotavtrykket* og pulser per m² var henholdsvis 0.68 meter og 2.8 pulser. Til sammen er det registrert omtrent 3 millioner punkt innenfor kommunegrensen (pers.medd. Groesz 2010). Det var hovedsakelig ett fly, LN-AEY, som fløy alle stripene bortsett fra stripe 3 sør for Hodalen som ble fløyet med LN-LAS, med andre ord to forskjellige sensorer. LiDAR-system ALTM 3100 ble brukt i begge flyene. ALTM (*Airborne Laser Terrain Mapper*) består av alle de vesentligste komponentene som scanner, GPS & INS, datamaskin for kontroll og tidssynkronisert datalagring. På Tolga ble følgende parametre brukt :

Flyhøyde (AGL)	800 meter
Repetisjonsfrekvens	100 000 Hz
Speil frekvens	55 Hz
Halvvinkel	17 deg
Hastighet	Ca 75 m/s

Tabell 1: Tallene er basert på gjennomsnittsmålinger fra laserprosjektet Hedmark skog 2006 (Blom Geomatics AS 2006)

Punktdata med x, y og z-koordinater ble levert fra Blom Geomatics AS. Terrenghøyden var fratrukket z-koordinatene ved levering.

2.3 Feltarbeidet og registreringsutstyr

Feltarbeidet startet 4. oktober 2009 med gjennomgang av noen prøveflater med veileder Vegard Lien fra Institutt for naturforvaltning for sjekk av utstyrets funksjonalitet og gjennomgang av rutiner for registreringen. Datainnsamlingen foregikk fra 5. oktober til og med 21. oktober.

Jeg hadde en mobil basemottaker montert i koffert som var tilgjengelig til bruk med radiokommunikasjon for bruk på fastmerker i områder hvor det var dårlig teledkning, men den kom ikke til anvendelse. GNSS-målingen ble derimot utført med en håndholdt målebok *Topcon FC200* og programmet *TopSURV* versjon 7.1. Måleboka var festet til en brakett på en 2 meter lang stav som ble brukt som oppmålingsreferanse 2 meter fra midten av stammen. Dårlige signalforhold under trekrona førte til en forlengning av avstanden til 4 meter. På toppen av staven satt GPS/GLONASS-antenne med kommunikasjonsmuligheter med roveren. Referansemottakeren, det vil si basestasjonen, befant seg på Kløftryggen på Tynset. Oppkoblet til Internet kunne den sende sine GPS data til en såkalt *caster* som igjen distribuerer data via GSM til den mobile roveren når jeg beveget meg i terrenget slik at roveren kontinuerlig kunne korrigere den målte posisjonen med disse verdiene. De kontinuerlige korreksjonene kan også lagres for senere bruk. Den mobile roveren var av typen *Topcon LegacyE+*, montert i sekk. De samme feilene og avvik for basestasjonen og roveren blir dermed eliminert når avstanden fra disse to mottakerne til de satellittene det måles mot er kjent. Dermed kan basestasjonen kommunisere med roveren og sistnevnte sammenligner fasemålinger, det vil si avstandsforandringer til satellitten, slik at man kan oppnå nøyaktigheter på centimaternivå i sanntid og millimaternivå ved etterprosessering.

Posisjonsbestemmelsen av hvert enkelt tre måtte gjøres med høy grad av nøyaktighet for å kunne koble hvert enkelt funn til gjeldende prøveflate. Dette ble opprettholdt av muligheten til å bruke både GPS og det russiske satellittsystemet GLONASS, man bruker i dag betegnelsen GNSS (*Global Navigation Satellite System*) som en felles betegnelse på global satellittbasert posisjonsbestemmelse. Dette er en form for differensiell GPS-posisjonering (DGPS) og bruker altså et nett av landbaserte referansestasjoner for å kunne måle/beregne de faktiske feilene når de oppstår. Sannsynligheten for riktig posisjonsbestemmelse øker med observasjonstiden, men skoglige forhold som blant annet grunnflatesum og stammeantall kan gi dårligere resultater (Næsset og Gjevestad 2008). Spesielt er tettheten uttrykt ved grunnflatesummen en god indikator på hvorvidt sannsynligheten er stor for å oppnå en *fixed* løsning (Næsset *et al* 2000). I tillegg spiller også antall tilgjengelige satellitter og satellittgeometrien inn. Sistnevnte uttrykk menes hvordan satellittene befinner seg på himmelen og en større spredning over basestasjonen gir bedre nøyaktighet.

Signalkvaliteten som roveren fanget opp ble oppdelt i tre trinn, "*fixed*", "*float*" og "*auto*" hvorpå i hovedsak det mellomste trinnet (*float*) ble brukt. Trinnfordelingen varierer med nøyaktigheten som kan påvises og dette har blant annet rot i hvor mange satellitter man har kontakt med. *Fixed* indikerer den beste signalkvaliteten. *Float* har dermed en dårligere signalkvalitet og dette har som oftest sin sammenheng med at trær og topografiske forhold ofte kan forvrengte eller hindre signaler å nå antennen. I skog er det vanskelig å få til kontinuerlige signaler fra satellittene over en tilstrekkelig tidsperiode for å kvalitetssikre den sanne posisjonen. Grunnen er temporære signalblokkeringer

som opptrer regelmessig forårsaket av greiner og topper. I glissen skog i høyereliggende strøk er sjansen betraktelig større for bedre signalkvalitet (*fixed*).

Måleteknikken heter RTK (*real-time kinematic*), en posisjonsberegning av GPS-data i mottakerens beregningsprogram i sanntid (*real-time*). *Kinematic*, eller kinematisk betyr at en mottaker er under bevegelse under måling og vil med dette oppnå en dårligere nøyaktighet sammenlignet med statisk måling, men til gjengjeld en høy produktivitet.

Kryssvis klaving i brysthøyde med utgangspunkt i himmelretningene nord-sør og øst-vest ble gjort i felt, mens gjennomsnittet av disse to målingene ble benyttet i dataanalysen. Nøyaktigheten er på millimeternivå. Alle trær over 500 mm ble fastsatt diametermål etter sunn fornuft på grunn av klavens begrensning, men dette hadde heller ingen innvirkning på datamaterialet siden alle trær over 500 med mer havnet i samme klasse. Høydemåling foretatt med en *Vertex* høydemåler ble gjort ved hvert femte registrerte tre. På grunn av problemer med utstyret ble ikke dette oppfylt for hele datamaterialet.

	Min	Median	Mean	Max	Totalt
Diameterregistreringer	350	407	421	700	801
Høydemålinger	7.0	16.0	15.5	24.0	120
Antall flater besøkt					1112
Antall flater med forekomst					358

Tabell 2: Kjernetall for feltregistreringer.

2.4 Bearbeiding av datamateriale

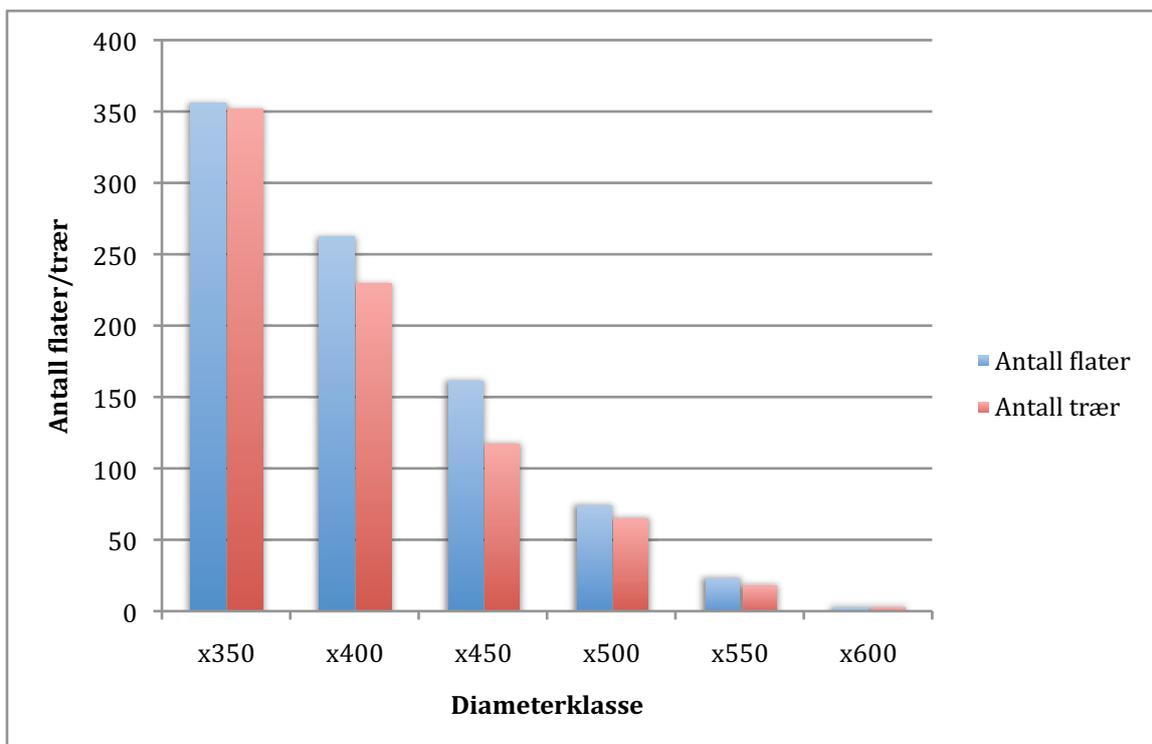
Feltdatamaterialets kjernetall er i hovedsak vist i tabell 2 og figur 2.

Fremgangsmåten for denne oppgaven er i tråd med hva Næsset (2002) og Gobakken & Næsset (2008) kaller en to- stegs prosedyre. Første steg er å bruke georefererte prøveflater med laserdata til å utvikle empiriske forhold mellom laservariabler og skoglige parametre, i mitt tilfelle forekomsten av store trær. I steg to vil forholdet dem imellom resultere i predikerte verdier for hver prøveflate. Predikeringen baserer seg altså på hvilken egenskap ved laserdataene som påvirker sannsynligheten for forekomsten av store trær.

Bearbeidelsen av datamaterialet omfattet blant annet å gi hver prøveflate-ID en kode for henholdsvis besøkt og ikke-besøkt. Dette har sin sammenheng i min fysiske tilstedeværelse i de respektive prøveflatene selv om det ikke var forekomst av store dimensjoner der. Forekomster av store dimensjoner som kunne ansees som tvilstilfeller med tanke på geografisk tilhørighet i prøveflata ble registrert under feltarbeidet. Etterprosesseringen la det endelige grunnlaget for hvilke registreringer som med sikkerhet kunne ansees som innenfor. Registreringer utenfor korridoren fikk en distanseverdi >0 og ble dermed ekskludert i ArcGIS.

En annen feilkilde det er tatt høyde for er nylig tynnede eller sluttavvirkede prøveflater. Siden det er gått 3 år fra lasertakseringen til min feltregistrering er det en viss sjanse for at et hogstinggrep i prøveflatene er blitt foretatt og dette vil få innvirkning på mine resultater. Konsekvensen for slike prøveflater blir en omkoding til "ikke-besøkt".

Under feltarbeidet ble det til sammen registrert 956 trær, men etter ekskludering på grunnlag av feil treslag og dGPS-koordinater utenfor prøveflatesystemet, var det endelige antallet 801 trær (Tabell 2). I tillegg ble det laget såkalte x-kolonner for hver diameterklasse. X-kolonnene er sentrale i den videre dataanalysen og tar utgangspunkt i hver prøveflate ved prediksjonen av store trær. X-kolonnene ble gitt verdier, bionomiske koder (0 og 1), som representerer om det er forekomst av store trær innenfor gjeldende diameterklasse. Diameterklasseinndelingen har følgende gruppering: >350 mm, >400 mm og >500 mm. Alle funn fra 500 mm til og med den største registrerte dimensjonen (Tabell 2), ble plassert i samme klasse. Diameterklassene som refereres til i oppgaven omtales heretter som x350, x400 og x500.



Figur 2: Kjernetall for hver diameterklasse. Antall flater med forekomst og totalt antall trær per diameterklasse.

2.5 Laservariabler

For denne oppgaven er første og siste retur for laserpulsene brukt i forbindelse med estimeringen av laservariablene. Segmenteringen av pulsreturene foregår i praksis ved at lasersystemets mottakerdel programmeres til å måle og avvete alle returpulser etter første pulsen, men kun registrere den siste pulsen. Alle punkter under 1,3 meter fra bakkenivå ble ikke definert som vegetasjonstreff. Dette eliminerer sjansen for at steiner

eller andre terrengobjekter gir innvirkning på datamaterialet. Laservariablene ble klargjort for videre dataanalyse av hovedveileder Terje Gobakken, Institutt for naturforvaltning.

2.5.1 Høydepersentiler - PF-verdier

Såkalte høydepersentiler viser fordelingen av punktene spredt i den vertikale retningen. De har i mange undersøkelser blitt brukt som uavhengig variabel i regresjonsmodeller for estimering av forskjellige skoglige parametre på prøveflatenivå (Gobakken & Næsset 2008; Næsset & Gobakken 2008; Næsset 2002; Lim *et al.* 2003)

Med bakgrunn i høydemåling for hver enkelt puls ble høydepersentilene, referert til PF-verdier i denne oppgaven, beregnet innenfor kategoriene [PF 0, PF 10.....PF 90]. Resultatet ble et sett med høydeverdier innenfor hver respektiv prøveflate. PF-kategoriene representerer en prosentvis sats som gjenspeiler alle laserpulsenes høydeverdi i stigende rekkefølge. Persentilene er grupperingsenheter av statistiske tall som er ordnet i 10 like store intervaller fra 1,3 meter over bakken til den maksimale høyden innenfor hver flate. Den prosentvise satsen er definert som prosent av observasjoner nedenfor. F.eks vil PF 50 ta utgangspunkt i gjennomsnittsverdien i midten av rekkefølgen, mens PF 0 tar utgangspunkt i den første verdien. Sistnevnte vil som regel ha verdier i nærheten av 1,3 siden alle verdier under dette nivået ble utelatt fra det endelige datasettet. I resten av oppgaven omtales PF-verdiene som fPF eller lPF for henholdsvis første og siste retur (*first/last*)

2.5.2 Tetthetspersentiler - TF-verdier

TF verdiene er såkalte tetthetspersentiler og har samme utgangspunkt som PF-verdiene i forhold til bakgrunnen for målingene. I tillegg følger også kategoriinndelingen samme mønster [TF 0, TF 1.....TF 9], men bakgrunnen for inndelingen er annerledes. Punktetettheten kalkuleres innenfor 10 like lange vertikale avstander hvor hver enkelt avstand defineres som 1/10 av avstanden mellom 95 % persentilet og 1,3 meter over bakken (Gobakken & Næsset 2005; 2008 og Næsset (2009). Hver høydemessige sats gjenspeiler hvor mange prosent av den samlede punktetettheten innenfor hver prøveflate som befinner seg over denne satsen. F. eks omfavner TF 0 alle laserpulsmålinger over 1,3 meter i ei prøveflate hvor 90 % persentilet er 11,3 meter. I resten av oppgaven omtales TF-verdiene som fTF eller lTF for henholdsvis første og siste retur (*first/last*)

2.6 Eksterne variabler

Dette er såkalte naturgitte og kulturelterte variabler som ble implementert i dataanalysen som et supplement til oppgaven. Opplysningene ble hentet ut fra Norge Digitalt i ArcGIS med hovedfokus på Tolga kommune. Den kulturbetingede variabelen veiavstand, det vil si avstanden fra prøveflatene til nærmeste vei, ble beregnet ut fra forskjellige veikategorier; fra riksveier og fylkesveier til skogsveier. Høydedataene dannes fra et triangulært, irregulært nettverk (TIN). Dette betyr at punktene langs høydekurvelinjene får samme verdi og organiseres som hjørnene i en trekant.

Trekantene kan ha ulik form og størrelse og utgjør dermed et irregulært nettverk. Interpoleringsteknikk sørger for at en høydeverdi kan leses ut fra nær sagt hvor som helst. Et annet kartlag tar utgangspunkt i himmelretning oppgitt i grader for hver pixel på kartet. Eksponeringen kan graderes etter en nord til sør gradient. Dette innebærer blant annet at himmelretningene øst og vest får samme verdi (90 grader). En såkalt *slope-funksjon* angir stigningen målt i prosent. Et arealressurskart med muligheter for å kategorisere områder etter bonitet (m³/ha/år) ble tilslutt utelatt fra dataanalysen på grunn av en mer forenklet beregning uten hensyn til denne.

VARIABEL	MEAN	MIN	MAX	SD(F)	VARIABEL	MEAN	MIN	MAX	SD(L)
fMean	6.456	1.39	17.30	2.94	lMean	6.57	1.39	17.99	3.22
fcv	36.86	0	39.80	10.04	lcv	39.80	0	132.9	12.06
fPF0	1.50	1.30	8.19	0.58	lPF0	1.51	1.30	9.71	0.74
fPF10	3.52	1.30	13.35	1.99	lPF10	3.31	1.30	14.10	2.32
fPF20	4.39	1.30	15.30	2.40	lPF20	4.31	1.30	16.37	2.81
fPF30	5.10	1.30	16.31	2.68	lPF30	5.20	1.30	17.58	3.12
fPF40	5.73	1.32	17.30	2.92	lPF40	5.84	1.32	18.34	3.33
fPF50	6.35	1.32	17.90	3.13	lPF50	6.53	1.32	18.92	3.51
fPF60	6.98	1.32	18.55	3.32	lPF60	7.20	1.32	19.46	3.69
fPF70	7.65	1.39	19.21	3.50	lPF70	7.92	1.39	20.06	3.84
fPF80	8.45	1.39	19.95	3.70	lPF80	8.75	1.39	20.72	3.99
fPF90	9.55	1.39	21.03	3.97	lPF90	9.85	1.39	21.65	4.20
fTF0	0.39	0	0.95	0.23	lTF0	0.24	0	0.70	0.14
fTF1	0.38	0	0.94	0.23	lTF1	0.23	0	0.66	0.14
fTF2	0.35	0	0.93	0.22	lTF2	0.21	0	0.64	0.13
fTF3	0.32	0	0.92	0.21	lTF3	0.19	0	0.63	0.13
fTF4	0.28	0	0.88	0.20	lTF4	0.16	0	0.61	0.12
fTF5	0.23	0	0.80	0.18	lTF5	0.14	0	0.57	0.10
fTF6	0.19	0	0.66	0.15	lTF6	0.11	0	0.48	0.09
fTF7	0.13	0	0.53	0.11	lTF7	0.09	0	0.33	0.07
fTF8	0.09	0	0.35	0.07	lTF8	0.06	0	0.22	0.05
fTF9	0.05	0	0.16	0.03	lTF9	0.03	0	0.11	0.02

Tabell 3: Deskriptiv statistikk for laservariablene i alle besøkte flater på Tolga, dvs 1112.

VARIABEL	MEAN	MIN	MAX	SD
Stigning	10.5 (11.2)	0	50 (44)	7.6 (6.6)
Eksponering	176.5 (193.8)	0	357	104.4 (101.2)
Eksponering2	87.4 (88.2)	0	180	48.9 (45.5)
H.O.H	724.6 (695.4)	493	989 (855)	95.4 (76.6)
Veiavstand	420 (440)	1	1621 (1116)	335 (311)

Tabell 4: Deskriptiv statistikk for eksterne variabler i alle besøkte flater på Tolga, 1112. Tall i parentes tar utgangspunkt i flater med forekomst, 358.

2.7 Statistisk analyse

Utvelgelsen av laservariabler og analyser er foretatt i *R* (R Development Core Team 2007). Sannsynligheten for forekomst er tilpasset en generalisert lineær modell (GLM) basert på en bionomisk fordeling. I begynnelsen var det interessant å se hvilke av de uavhengige laservariablene som best kunne beskrive forekomsten av store dimensjoner. Regresjonsprosedyren "*forward selection*" inkluderer statistisk signifikante variabler (p -verdi $<0,05$) for å bygge opp modellene til de ulike diameterklassene. Det vil fortløpende velges den uavhengige variabelen som forbedrer modellen best, til modellen ikke kan forbedres lengre. Kriteriet for inkludering av variabler er *Akaike Information Criterion* (AIC) som ligger som en egen funksjon i *R*, og sluttmodellen baserer seg på høy modelltilpasning og lav kompleksitet. Dette er teoretisk to motstridende mål, da gjerne høy modelltilpasning fås ved inkludering av mange variabler for bedre tilpasning til variasjonen og lav kompleksitet innebærer få variabler for blant annet å unngå korrelasjon. AIC bygger dermed på en inkludering av få variabler, men som passer godt til dataene. AIC er et verktøy for å sammenligne modeller, ikke et mål for å teste modeller. Lav AIC var dermed bakgrunnen for utvalget av variabler i en stegvis regresjonsprosedyre, mens signifikansnivået ($p < 0.05$) var "taket" for antall variabler i sluttmodellen.

Modellene ble etter hvert brukt til å predikere sannsynligheten for funn av store dimensjoner og dermed sammenligne predikeringen opp mot om det faktisk var forekomst av store dimensjoner der, med andre ord mine registreringer. Sannsynlighetsnivået er satt til 50 %.

2.7.1 Korrelasjon mellom variabler og multikollinearitet i modeller

Multikollinearitet defineres som når det er høy korrelasjon mellom to eller flere av forklaringsvariablene i en regresjonsmodell. Høyde- og tetthetspersentilene (PF- og TF-verdier) i tillegg til eksterne variabler gir veldig mange potensielle forklaringsvariabler for modellene. Dermed øker sjansen for multikollinearitet betraktelig. Variablene som inngår i de signifikante sluttmodellene ble undersøkt for korrelasjoner i *R*. Resultatet ble en korrelasjonsmatrise hvor korrelasjonskoeffisienten gir opplysninger om hvor mye to variabler samvarierer, f. eks om en økning hos den ene fører til en like stor økning hos den andre og motsatt (Figur 3). Denne informasjonen gis av tallverdier mellom eller lik -1 og +1 hvor det i denne oppgavens sammenheng strebes etter verdier så nærme 0 som mulig. Verdien 0 betyr ikke at det ikke er noen relasjon mellom variablene, men *Pearssons korrelasjonskoeffisient* måler bare styrken av lineær sammenheng. Ut fra verdiene i korrelasjonsmatrisen ble egenverdier (*eigenvalues*) beregnet. *Condition number* var beregnet som kvadratroten av største egenverdi delt på minste egenverdi og brukt som mål for påvisning av multikollinearitet i denne oppgaven.

Formel for condition number (κ -value):

$$\kappa = \sqrt{\frac{\lambda_{max}}{\lambda_{min}}}$$

2.8 Monte Carlo-simulering og kryssvalidering

Iterasjon er et matematisk begrep som brukes for en tilnærming av et utfall. Å iterere vil si å bruke en beregningsprosess som gjentas et visst antall ganger slik at man kontinuerlig får et forbedret resultat. Prosessen kjennetegnes ved at det etter hvert vil utgjøre en liten forskjell på resultatene mellom to påfølgende gjentak, den vil konvergere mot en verdi som synes å være det beste estimatet for modellen. En kryssvalidering etter dette prinsippet ble utført ved bruk av Monte Carlo-simulering. Zhang (1996) peker på fire forskjellige metoder for validering av regresjonsmodeller hvor datasplitting og kryssvalidering er én av disse metodene som ble anvendt for denne oppgaven.

Monte Carlo-simuleringen startet med at det ble foretatt et tilfeldig utvalg på 834 flater, det vil si 75% av de besøkte flatene. Disse flatene ble brukt til å beregne parameterestimater for hver parameter som inngår i de signifikante sluttmodellene for de respektive diameterklassene. Den tilfeldige utvelgelsen ble gjentatt 300 ganger, hvert utvalg unikt med tanke på flateutvalget. Verifiseringen av disse modellene ble gjort på de resterende 278 flatene som ikke ble trukket ut i simuleringen. De fungerte dermed som et uavhengig datamateriale for kontroll av sluttmodellene. Disse flatene fikk predikert en sannsynlighetsverdi basert på parameterestimater fra sluttmodellen til simuleringsflatene. Til grunn for sannsynlighetsverdiene ligger altså de modellerte parameterestimater som kontrolleres i forhold til parameterestimater for hver av de 278 kontrollflatene. Man får dermed en sannsynlighetsverdi oppgitt i prosent som beskriver hvor godt hver flate er tilpasset sluttmodellen og dermed sannsynligheten for å finne store trær der. Sannsynlighetsverdiene ble delt i over/under 50 % og satt i en forekomstmatrise som viser hvor mange flater hvor predikerte verdier sammenfaller eller ikke sammenfaller med henholdsvis forekomst eller ikke forekomst (Tabell 8 og 9). Siden simuleringen ble gjentatt 300 ganger med forskjellig utvalg av flater som beregningsgrunnlag fikk man forskjellig resultat fra gang til gang. Ettersom flere simuleringer blir kjørt vil gjennomsnittstallet bli mindre influert av variasjonene mellom utvalgene og etter 300 simuleringer har gjennomsnittstallene i forekomstmatrisen stabilisert seg på et nivå som synes å gjenspeile sluttmodellen i best mulig grad.

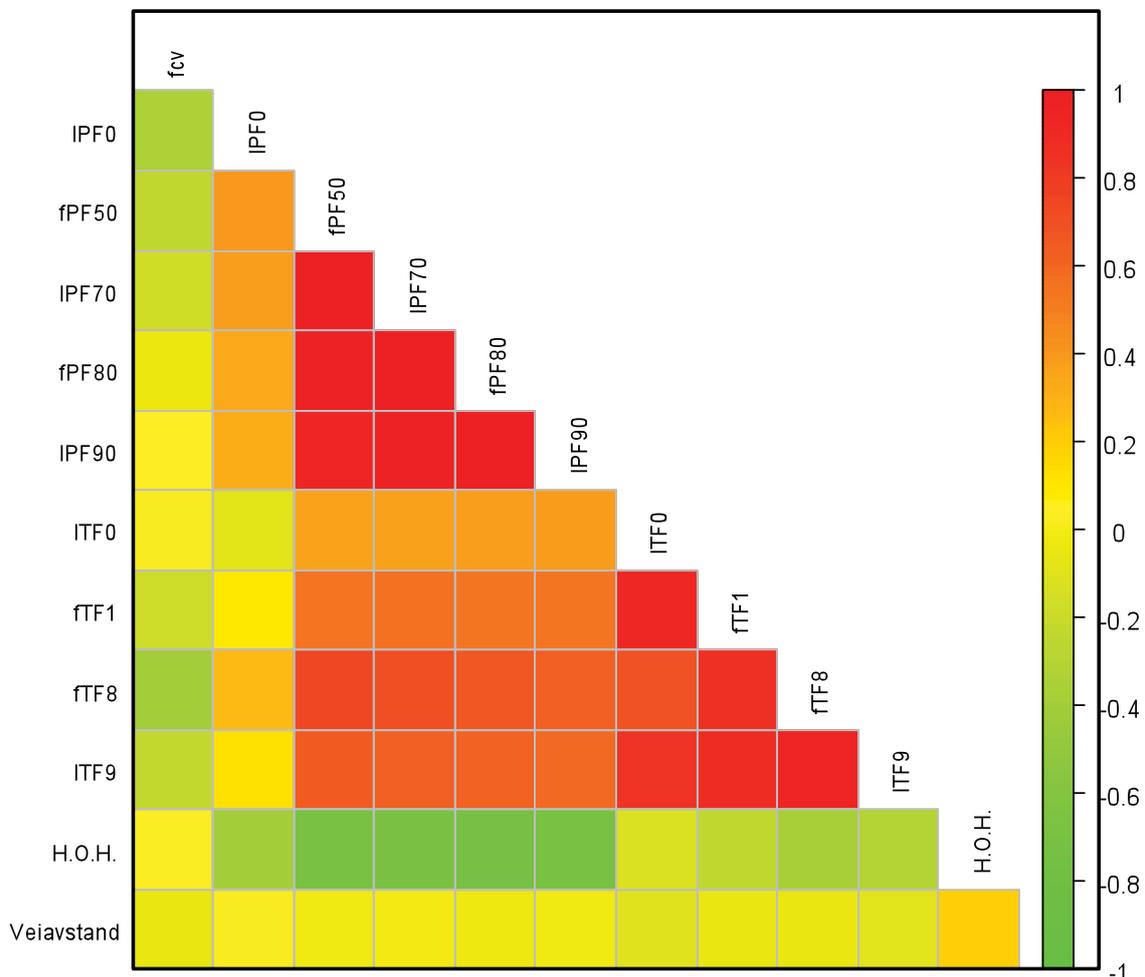
2.9 Feltdatamaterialets størrelse

En interessant vinkling på denne oppgaven er å se på hvordan feltdatainnsamlingen kan effektiviseres. Tidsforbruket er vanskelig å måle, men en annen metode som kan brukes er å måle mine modellers nøyaktighet i forhold til den kvantitative datamengden. Hvor stor innsats er nødvendig i felt for å kunne predikere sannsynligheten for forekomst med tilnærmet like god nøyaktighet? Dette ble gjort ved å sammenligne standardavviket opp mot forskjellige prosentandeler av mitt totale datamateriale. Rent praktisk ble dette løst ved en Monte Carlo-simulering med uttak fra 5% 10% 15%.....75% av det totale datamaterialet. Gjentakene ble gjennomført 500 ganger og verifiseringen ble foretatt på 25% av flatene som ikke ble trukket ut i simuleringen.

Forklaringsvariabler	
LiDAR-data, første og siste puls	Eksterne
PF0, PF10.....,PF90	Høyde over havet
TF0, TF1.....,TF9	Stigning (%)
Gjennomsnittlig høyde (mean)	Eksponering
Kovarians (cv)	Eksponering ^{nord-sør}
Sensor (AEY og LAS)	Veiavstand
Forklaringsvariabler utelatt fra dataanalysen	
Maksimal høyde (max)	Bonitet
Gj.snittlig ant. pulser per prøveflate (n)	
Totalt antall pulser per prøveflate (ntot)	

Tabell 5: Oversikt over totalt antall variabler som både inngår i analysene og er utelatt. Eksponering^{nord-sør} = modellert med tanke på en nord-sør gradient.

3. Resultater



Figur 3: Korrelasjonsmatrise for variablene som inngår i sluttmodellene (tabell 6 & 7)

Figur 3 bruker *Pearssons korrelasjonskoeffisient* som bakgrunn for en grafisk framstilling av hvordan variablene i sluttmodellene korrelerer med hverandre. F. eks er IPF90 og høyde over havet negativt korrelert, som er markert med grønt, i den forstand at jo høyere over havet man kommer desto lavere blir IPF90-verdien. (lavere trær). I motsatt ende av skalaen finner vi positiv korrelasjon markert med rød farge. Dette forekommer når en økning i f.eks IPF90-verdier svarer til en økning av fPF80 fordi de er nært beslektet med hverandre.

3.1 Sluttmodeller

Som følge av multikollinearitet variablene imellom innskrenkes mulighetene av hvor mange og hvilke variabler som kan brukes i sluttmodellene med tanke på κ -value. Tabell 6 viser sluttmodellene for hver diameterklasse basert på *forward-selection* regresjonsprosedyre med signifikante p-verdier (<0.05). Både x350 og x500 modellene indikerer en høy grad av multikollinearitet. Næsset (2002) bruker en κ -value høyere enn 30 som grunnlag for indikasjon av multikollinearitet.

Som nevnt tidligere er stripe 3 fløyet med en annen sensor, men i og med at det ikke er registrert noen forekomster der tross 55 takserte flater hadde det i utgangspunktet ingen store konsekvenser. En inkludering av variabelen "sensor" i sluttmodellene ga et ikke-signifikant resultat ($p=0.974$) som bekrefter at laserdataene fra denne stripa ikke har noen signifikante innvirkninger på resultatene. Det er med andre ord som om det samme flyet er fløyet over hele Tolga.

Modell		κ -value
X350	$-10,820+1,429_{IPF90}-0,257_{fPF50}+0,007_{H.O.H.}-0,482_{IPF0}-0,875_{fPF80}$	24,90
X400	$-16,674+0,629_{IPF90}-0,313_{fPF50}+0,011_{H.O.H.}+0,063_{fcv}+2,693_{ITF0}$	11,59
X500	$-2,254e+01+1,123e-01_{fcv}+1,091e+01_{fTF1}-5,754e+01_{ITF9}+4,991e-01_{IPF70}+1,460e-02_{H.O.H.}-1,530e-03_{Veia\text{vst.}}-2,133e+01_{fTF8}$	35,51

Tabell 6: Variablene og multikollinearitet for modellene basert på stegvis regresjonsprosedyre.

Variabelutvelgelsen i Tabell 7 er gjort med bakgrunn i Tabell 6 og hvor mange modeller de er representert i, det vil si modeller med utgangspunkt i kun bruk av materiale fra laser og eksterne variabler hver for seg. Disse modellene vises ikke her. Samtidig er utvelgelsen av variabler begrunnet ut fra ønsket om liten multikollinearitet (κ -value <30)

Modell		κ -value
X350	$-15,753+0,096_{fcv}+0,454_{IPF90}+0,008_{H.O.H.}+1,636_{fTF1}$	3,18
X400	$-18,346+0,105_{fcv}+0,395_{IPF90}+0,011_{H.O.H.}+2,129_{fTF1}$	3,18
X500	$-19,905+0,100_{fcv}+0,257_{IPF90}+0,013_{H.O.H.}+3,013_{fTF1}$	3,18

Tabell 7: Variablene og multikollinearitet for modellene basert på utvalgte variabler.

På forhånd ble en inkludering av følgende modell forsøkt for diameterklassene: fcv+IPF90+H.O.H+Veiavstand. Alle variablene ga signifikante resultater, bortsett fra veiavstand (p-verdi = >0,05). I tillegg ble estimatverdien til denne variabelen negativ for x400 og x500. Modellen predikerer altså at sannsynligheten for å finne store dimensjoner øker desto nærmere vei man kommer, noe som virker ulogisk. Jeg byttet dermed ut denne variabelen med en annen gjenganger for ulike modeller; fTF1.

Modellen ble dermed: fcv+IPF90+H.O.H.+fTF1 (Tabell 7). Dette var den eneste modellen som ga signifikante verdier (p-verdi = <0.001) for alle diameterklassene. En beregning av multikollineariteten ga en κ -value på 3.18, noe som indikerer at dette ikke er noe problem for sluttmodellene.

3.2 Predikering av sannsynligheten for forekomst av store trær

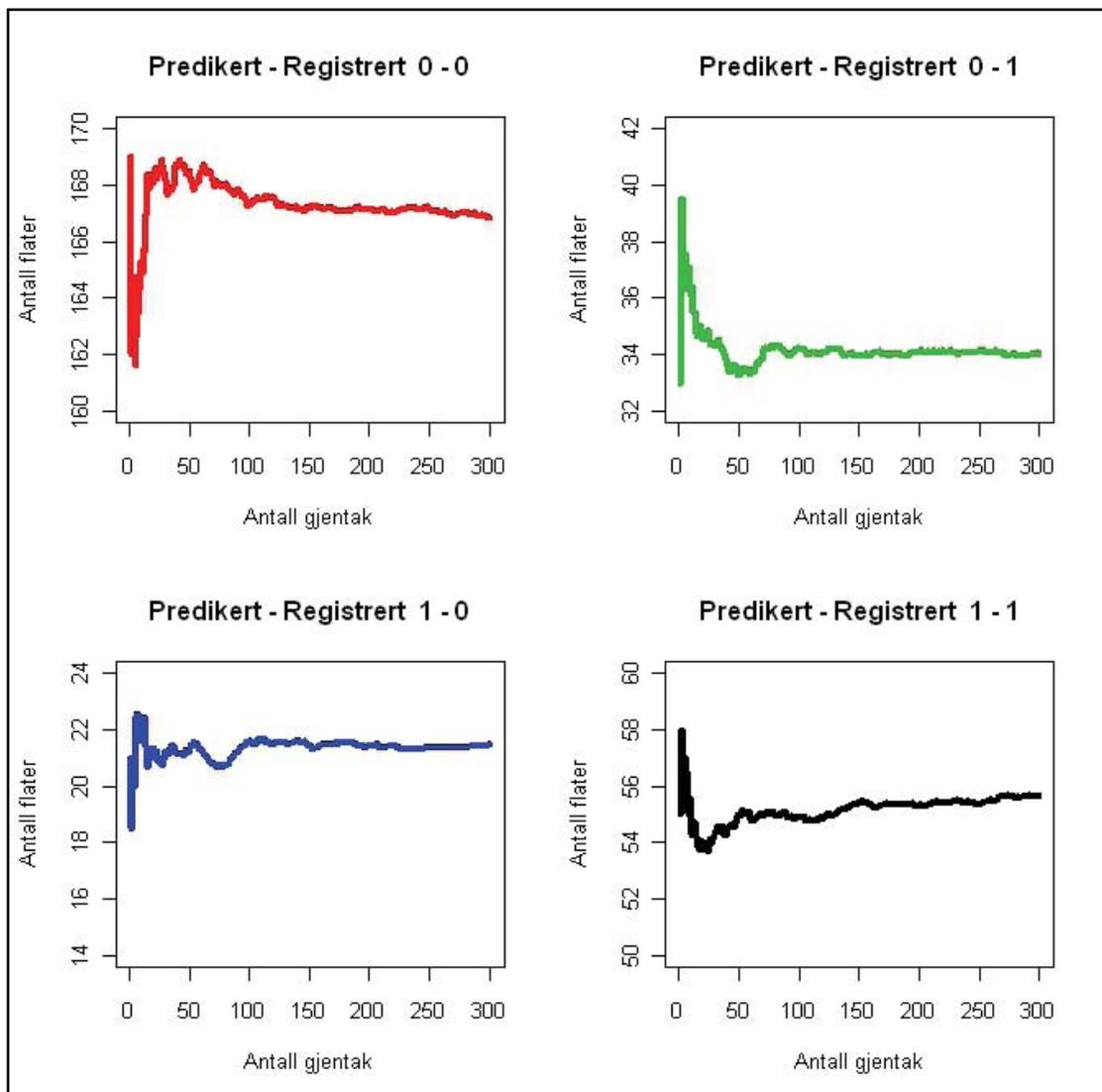
Monte Carlo-simulering			Regresjonsmodell		
x350					
Predikert 0	666.4	136.1	Predikert 0	668	134
Predikert 1	87.0	222.5	Predikert 1	86	224
	Registrert 0	Registrert 1		Registrert 0	Registrert 1
x400					
Predikert 0	795.9	147.6	Predikert 0	800	147
Predikert 1	51.1	117.4	Predikert 1	47	118
	Registrert 0	Registrert 1		Registrert 0	Registrert 1
x500					
Predikert 0	1024.7	66.0	Predikert 0	1026	66
Predikert 1	9.6	11.6	Predikert 1	9	11
	Registrert 0	Registrert 1		Registrert 0	Registrert 1

Tabell 8: Forekomstmatrise 2x2, prediksjon av forekomst basert på Monte Carlo-simulering og stegvise regresjonsmodeller.

Resultatene av den multiple regresjonsanalysen er satt inn i en forekomstmatrise (Tabell 8) som setter modellenes predikeringer og registreringer opp mot hverandre. Grunnlaget for inndelingen er en sannsynlighetsberegning for hvor godt hver prøveflate er tilpasset sluttmodellen. Hvis sannsynligheten er større eller lik 50 % predikeres det forekomst av store dimensjoner. Er sannsynligheten mindre enn 50 % predikeres ingen forekomst. En optimal fordeling er selvfølgelig at hvor modellen predikerer forekomst har jeg registrert forekomst og omvendt. Med utgangspunkt i regresjonsmodellen treffer den på 63 %, 45 % og 14 % for henholdsvis x350, x400 og x500 hvor predikeringen av forekomst stemmer overens med mine resultater (Figur 2 og Tabell 8).

Etttersom diameteren øker vil det bli færre registrerte forekomster og modellene predikerer dermed færre flater med forekomster (Figur 2).

Summen av alle tallene i hver matrise blir 1112 som gjenspeiler totalt antall flater besøkt (Tabell 2).

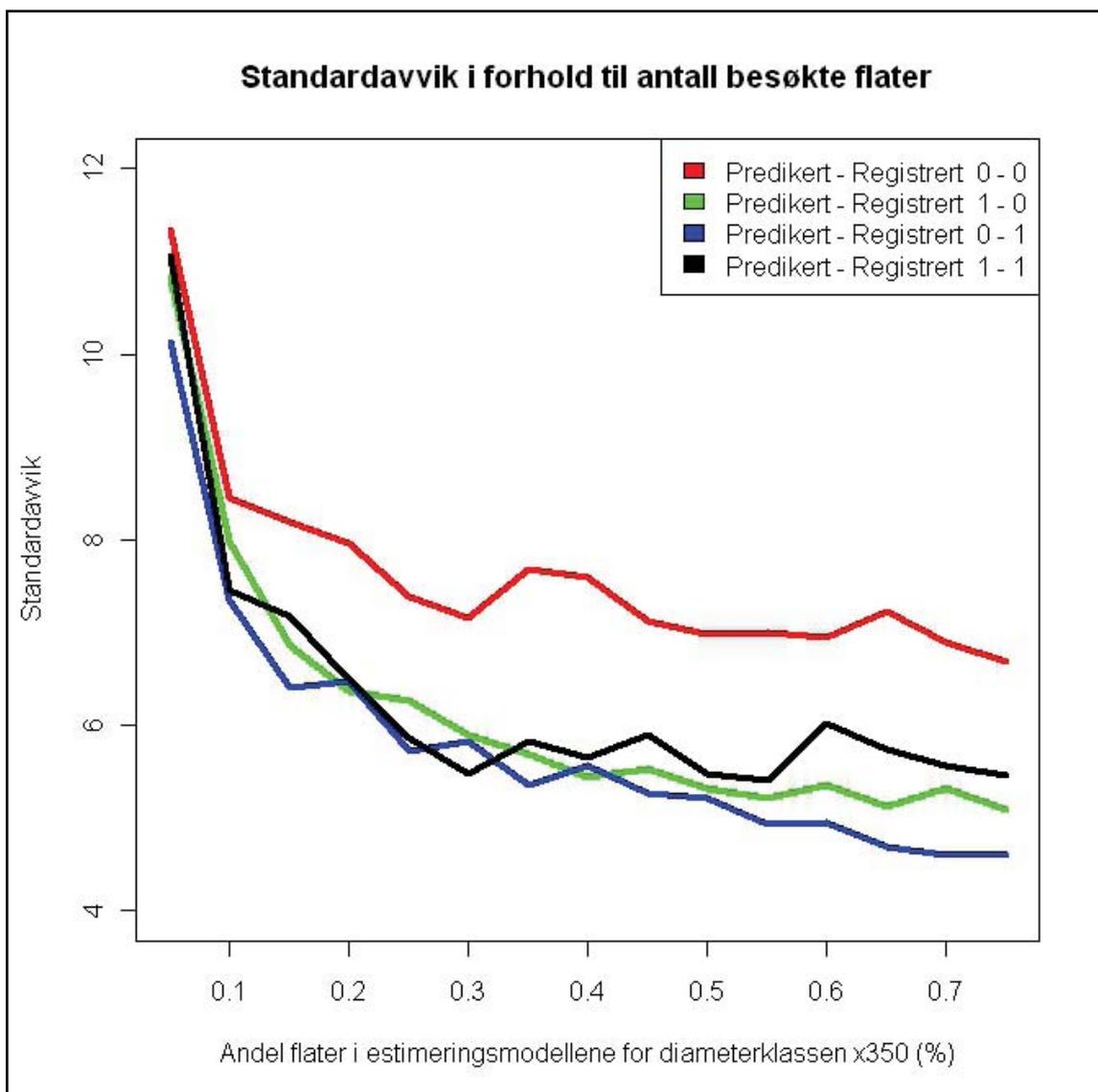


Figur 4: Monte Carlo-simulering for diameterklassen x350. Verifiseringen er foretatt på 25 % av de besøkte flatene, derav de lave verdiene på y-aksen.

Simuleringen i Figur 4 viser variasjoner innenfor hver kvadrant i matrisen. I forhold til Tabell 8 er det her kjørt en kryssvalidering på 25 % av datamaterialet. Etter ca. 100 gjentak har resultatene konvergert mot en verdi som synes å være den beste for modellen. Etter såpass mange uttrekk vil store sprik i beregningene ikke gi nevneverdige utslag for gjennomsnittstallet.

Prøveflater som ikke predikerer forekomst og hvor heller ingen forekomster er registrert viser den største variasjonen i begynnelsen. Denne tar utgangspunkt i 167 flater, det vil si 668 flater totalt hvis man ser bort fra verifiseringsandelen på 25 %. (Tabell 8). Dette er over halvparten av alle besøkte flater (Tabell 2).

3.3 Feltdatamaterialets størrelse



Figur 5: Standardavvik ved Monte Carlo-simulering med økende estimeringsandel av flatene (5 % - 75 %). Til grunn for beregningene ligger regresjonsmodellen for x350.

Simuleringen i Figur 5 viser en utjevning av standardavviket for alle fire kvadrantene med utgangspunkt i diameterklassen x350 for regresjonsmodellen (Tabell 6 og 8). Utjevningen av standardavviket starter allerede ved en estimeringsandel for omtrent 30 % av alle besøkte flater. Deretter holder det et jevnt nivå mot den høyeste estimeringsandelen, 75 % av alle flater.

4. Diskusjon

4.1 Materiale og metode

4.1.1 Feltdatamaterialet

Til sammen 1112 flater ble besøkt, og av disse er det omtrent 1/3 eller helt eksakt 358 flater hvor det er registrert minst en forekomst av store dimensjoner (Tabell 2).

Ekskluderingen av 49 gran fordelt på til sammen 9 prøveflater, vil neppe ha stor innvirkning på datamaterialet siden det berører så få flater. Noen av disse "granflatene" vil fortsatt ha forekomst av store dimensjoner av furu, mens andre flater uten forekomst vil kun være angitt som besøkt i datasettet. Mehtätalo *et al.* (2007) påpekte i deres forsøk, som blant annet gikk ut på å predikere diameterfordelinger basert på ALS-data i furuskog, at systemet av deres likninger vil være ugjenomførlig i enkelte bestand hvor furu ikke er dominerende treslag. Dette kommer av at laserdataene mangler sammenheng med gjennomsnittsverdier fra furudominert skog. Kroneformen vil variere mellom treslagene og vil ha innvirkning på fordelingen av høydepersentiler (Næsset 1997; 2004b). Dette gjelder i hovedsak forskjeller mellom bar- og lauvtrær. Gran vil muligens skille seg litt ut fra furu når det gjelder tetthetspersentilene siden disse vil være en funksjon av tretetthet og kronestørrelse (Næsset 2004b). Granforekomstene på Tolga var enkelte steder spredt fordelt, men i all hovedsak var forekomsten konsentrert der korridoren gikk gjennom et granbestand i omtrent 200 meters lengde.

Lasertaksering med høy punkttetthet har gitt lovende resultater for å kunne skille mellom ulike arter. Holmgren & Persson (2004) har kunnet skille mellom gran og furu med 95 % nøyaktighet i deres forsøk ved enkelttresegmentering relatert til kronestørrelse og -form samt egenskapene til returpulsene. Det er også mulig å bruke optiske instrumenter som et tillegg til ALS for ytterligere hjelp.

4.1.2 Feltarbeidet

Basestasjonens plassering i nabokommunen Tynset gjorde at avstanden til roveren var opptil 40-43 km på det maksimale. Dette gjelder for stripe 1 ved Langsjøen helt øst i Tolga. Ser man bort fra denne stripa, er den maksimale avstanden 20-25 km, noe som må kunne sies å være veldig kort (Deckert og Bolstad 1996). Dette vil eliminere faren for feilposisjoneringer som kan oppstå fordi basestasjonen og roveren har kontakt med forskjellige satellitter. En begrensning av såkalte *cut-off angles* hindrer basestasjonen i å kommunisere med andre satellitter. En svært liten del av mine takserte korridorer befinner seg i daler; de fleste går heller på tvers av hoveddalføret hvor Glomma renner. Taksering på høydedrag gir i barskoger en mye bedre nøyaktighet enn tilfellet er for daler (Deckert & Bolstad 1996).

Under feltarbeidet ble posisjonsbestemmelsen til flesteparten av forekomstene bestemt ved bruk av signalkvalitet i "float-posisjon", det vil si en usikker posisjonsbestemmelse. Allikevel vil ikke feilregistrering ha noen nevneverdig innvirkning for min oppgave siden en feilposisjonering av et tre kan føre til at registreringen gjøres i feil prøveflate sett i forhold til hva som er den geografisk korrekte prøveflaten. Dette kan skje når trekronen deles av grenselinjen mellom to prøveflater. I verste fall risikerer man å

registrere et tre innenfor ei prøveflate som egentlig ikke inneholder noen store dimensjoner og omvendt. Dette er lite sannsynlig ut fra korrelasjonen av flatene, det vil si at skogbildet sjelden går brått over i et annet format når prøveflatene ligger helt inntil hverandre.

Bruken av GSM-basert kommunikasjon innebærer at dGPS-korreksjonssignaler sendes kontinuerlig fra basestasjonen via GSM-mobiltelefonen som igjen er kablet til roveren slik at signalene kan bearbeides i sanntid. Min undersøkelse skiller seg ut fra tradisjonell grid-basert prøveflatesystem hvor prøveflatene spres jevnt utover et geografisk område. Registrering innenfor slike områder etter dGPS-prinsippet kan gjøres med radiobasert kommunikasjon hvor man bytter ut GSM-mobiltelefonen med en mobil basemottaker plassert over et fastmerke. Radiokommunikasjon har dog kort rekkevidde. Med tanke på at laserkorridorene er plassert med 6 km mellomrom ville en radiobasert feltregistrering ført til mye flytting av den mobile basemottakeren.

Sanntidskorrigeringer av dGPS over mobilt internett trenger kun lav båndbredde (bit/s) og det er små datamengder som overføres. GPRS bruker også GSM-mobiltelefonnettet, men har en større maksimal overføringshastighet. Med GPRS betaler man for overført datamengde, ikke for tiden man er oppkoblet. Dermed varte 200 kr på kontantkortet hele datainnsamlingsperioden igjennom. Svakheter med dette systemet er at man gjør seg avhengig av og overlater kontrollen til en teleoperatør, i mitt tilfelle Telenor. Ved manglende eller dårlig GSM-dekning har man dermed få alternativer. Under feltarbeidet sørget dårlig GSM-dekning for at arbeidet måtte avbrytes i to striper, blant annet stripe 6 i Vingelen på vestsiden av Knausvola. Her ble teledekningen borte, men da befant jeg meg i fjellbjørkeskog uten furuforekomster. En mulig løsning ved dårlig teledekning er en rover-til-rover løsning. Dette innebærer i praksis at man må ha to roversekker med radio tilgjengelig i begge sekkene. Dermed vil man i områder med dårlig teledekning sette opp den ene roversekken over et kjentpunkt og måle med den andre roveren med korreksjoner via radio. Nøyaktigheten er den samme ved rover-mot-rover som ved annen tradisjonell RTK-måling (pers. medd. Bøhle 2010).

Under feltarbeidet hadde jeg lite kommunikasjonsproblem med basestasjonen på Tynset. Fordeler ved å bruke egne baser som settes opp selv, er at man har mer kontroll over innstillinger, baseoppsett, korreksjonsformat og kompatibilitet enn når man kobler seg til en tjeneste, f. eks et referansenettverk (pers. medd. Bøhle 2010). Det er dessuten bedre muligheter for feilsøking på basesiden. En re-start av modemmet var løsningen når jeg fikk problemer med basestasjonen på Tynset.

4.2 Laserkorridoren – Monte Carlo-simulering og multikollinearitet

Prøveflater brukes i skogtakseringen som redskap for å hente ut informasjon fra skogen. Tidlige undersøkelser for laserdata gjorde bruk av en systematisk utlegging av prøveflater (grid-basert) som løste problemet med dårlige og usikre resultater når det gjaldt høydemålinger. Derfor ble det systematisk lagt ut prøveflater (f. eks 15 x 15 meter) hvor man tok maksimumsverdien fra hver prøveflate til grunn for utregning av f. eks gjennomsnittshøyden (Næsset *et al.* 2004).

I motsetning til en systematisk utlegging av prøveflater skiller denne oppgavens metode seg ut i og med at prøveflatene ligger helt inntil hverandre og det foretas dermed en

form for totalregistrering i laserkorridorene i midten av stripene. Registreringen oppnår dermed en høy korrelasjon mellom flatene og det vil få innvirkning på de statistikkbaserte sluttmodellene (Salas *et al.* 2009). Dette innebærer at forekomst av et stort tre eller flere store trær i én flate vil gi stor sannsynlighet for forekomst i de to tilgrensende naboflatene. Kissling & Carl (2008) har påpekt at klassiske statistiske analyser som blant annet regresjon har problemer med autokorrelasjon fordi metodene tar høyde for en uavhengig fordeling av feil.

En metode for validering av mine modeller er å dra ut i felt for å registrere forekomster av store dimensjoner i de gjenværende flatene som jeg ikke fikk besøkt. Selv om en slik datainnsamling foretrekkes, er det i skoglig sammenheng ofte vanskelig og dyrt å skaffe til veie et godt og uavhengig datamateriale med tilstrekkelig kvalitet og kvantitet i forhold til dataene som ligger til grunn for modellutformingen (Zhang 1996). Legger man til grunn resultatet i Figur 5 kunne det vært nok med omtrent 30 % av totalt antall flater. Dette forutsetter at flatene i et valideringsdatasett representerer denne oppgavens datamateriale gjennom blant annet skogstrukturene, treslags-sammensetningen og forholdet mellom antall flater med forekomst og antall flater uten forekomst. For denne oppgavens vedkommende ble det ikke tid nok til å kontrollere validiteten på et nytt uavhengig datasett, derfor ble det utført Monte Carlo-simuleringer.

En kryssvalidering med Monte Carlo-simulering sørget for et tilfeldig og spredt utplukk på 75 % av flatene jeg har besøkt (Tabell 8 & Figur 4). Dermed blir sjansen mindre for at naboflata, eller begge naboflatene, er med gjennom alle gjentakene. 300 gjentak lå til grunn for tallene i Tabell 8, men det viste seg at allerede ved omtrent 100 gjentak hadde tallene stabilisert seg rundt et gjennomsnittstall som syntes å passe modellene best (Figur 4). Størst utslag vises for kvadranten hvor det ikke predikeres og heller ikke er registrert funn. I teorien skulle verdien her ha konvergert til 188 flater siden dette er 25 % av totalt antall flater uten forekomst. Modellen opererer derimot med noen feilvurderinger i forhold til enkelte flater og predikerer en sannsynlighet på over 50 % for forekomst, selv om dette ikke stemmer med feltregistreringene. Allikevel utgjør disse flatene med null registreringer omtrent 1/3 av det totale antallet flater. Et større antall av flater fanger opp variasjonen i datasettet bedre og dermed øker sjansen for store utslag mellom gjentakene. Kryssvalidering vil ofte resultere i for optimistiske resultater (Maltamo *et al.* 2009). En validering opp mot et uavhengig datasett vil derimot som regel vise noe dårligere sammenhenger.

4.3 Laser- og eksterne variabler i sluttmodellene

På forhånd vil man ha en generell oppfatning av hvilke variabler man tror inneholder informasjon som oppfattes som sterkt innvirkende på responsvariabelen. Man har ikke fullstendig oversikt over alt som vil være assosiert og i sammenheng med forekomst av store trær. Selv med 44 variabler kan det være andre bakenforliggende ukjente faktorer som ikke ble eller ikke kunne bli målt, komplekse interaksjoner variablene imellom og heterogenitet (Hoeting *et al.* 2006).

Noen av laservariablene som ble laget ut av høydepersentilene og avledet av prøveflatene ble ikke med i den stegvise regresjonsprosedyren (Tabell 5). Blant annet Næsset (2004b) og Næsset & Gobakken (2005) unnlot å bruke maksimumshøyden i sine beregninger siden den varierer mye avhengig av prøveflatestørrelsen og punkttetthet.

Resultatene fra Næsset indikerer at andre laservariablers egenskaper kan erstatte maksimumshøyden ved regresjoner. *Fotavtrykket* kan i noen tilfeller unngå å treffe toppen av trekrona, avhengig av blant annet antall pulser/m² og størrelsen på *fotavtrykket*. Lim *et al.* (2003) presiserer at sjansen for å treffe toppen av krona minsker ved bruk av *small-footprint* systemer, samme system som min oppgave baserer seg på (gjennomsnittlig *fotavtrykk* 0.68 m).

Min datafangst foregikk imidlertid midt i laserstripa, det vil si at alle laserverdier i korridoren vil være målt omtrent 0 grader *off-nadir*. Dette innebærer i praksis at laserpulsene treffer vinkelrett på trekronene og man unngår dermed uheldige bivirkninger med laserdata registrert med store scanningsvinkler. Store vinkler fører til en større arealdekning både med tanke på *fotavtrykket* og sveipebredden, men datakvaliteten blir fort dårlig. Holmgren (2003) anbefaler ikke større åpningsvinkler enn 10 grader *off-nadir*. Holmgrens simulering av scanningsvinkler viste at de øvre verdiene av både høyde- og tetthetspersentilene ble mest påvirket ved store scanningsvinkler, dog mer merkbart for gran enn for furu hvor sistnevnte kjennetegnes av kortere kronelengde. Høydepersentilene endret seg mer ved store scanningsvinkler i skog med liten stammetetthet, en betegnelse som passer for den glisne furuskogen på Tolga.

Siden maksimumshøyden er utelatt er LPF90 en gjenganger i sluttmodellene av den stegvise regresjonsmodellen. Det faktum at siste returen skulle gjøre seg gjeldende med tanke på PF90-persentilet virker ulogisk siden siste returen ofte vil trenge dypere ned i trekronene sett i forhold til den første. Antall returer som genereres er i all hovedsak avhengig av hvilke fremmedelementer laserstrålen treffer i sin bane. Dermed kan systereturen sammenfalle med første returen eller være alt fra den andre til f. eks den fjerde returen av en puls. Dette er avhengig av hvor mange returer som ble generert av pulsen (Popescu *et al.* 2003). Gobakken & Næsset (2005) undersøkte blant annet forskjeller i høydepersentilene over en toårig vekstperiode og observerte at 90-persentilet og maksimumshøyden endret seg for siste returen. Dette indikerer at en liten del av siste returene også vil bli reflektert av den øverste delen av trekronene. Det må samtidig nevnes at siste returen er mye mer sensitiv til endringer i flyhøyde og *fotavtrykkstørrelse* enn første returen.

Min oppgave inkluderer 44 laservariabler samt 5 eksterne (Tabell 5). Næsset (2004) og Bollandås & Næsset (2007) brukte i alt henholdsvis 44 og 46 forklaringsvariabler i sine studier og dette innebærer en stor sjanse for høy korrelasjon (multikollinearitet) mellom enkelte, det vil si en stor lineær sammenheng. Dette gjelder spesielt for laservariabler som er nært i beslektet i den forstand at de inneholder informasjon fra omtrent samme høydenivå innad i prøveflata (Figur 3). I en skogstruktur med noenlunde samme treslag, alder og høyde vil f.eks fPF80 og fPF90 være høyt korrelert siden de er utledet fra det øvre sjiktet i skogen. Dette gjelder også selvsagt for samme høyde- og tetthetskategorier på tvers av første og siste retur. Spesielt x500 i regresjonsmodellen (Tabell 6) fikk en κ -value >30 , som innebærer en høy grad av multikollinearitet. Dette har nok sammenheng med at ialt 7 forklaringsvariabler inngår i sluttmodellen. 3 av dem kommer fra tetthetspersentilene (fTF1, fTF8 og lTF9). Regresjonsmodellen x350 fikk også forholdsvis høy grad av multikollinearitet (κ -value = 24,90). I denne modellen inngår LPF90 og fPF80 som er nært beslektet (Figur 3).

Både fcv og lcv gir en pekepinn på den romlige utformingen av den vertikale strukturen i prøveflata. Det er en forholdsverdi mellom standardavviket og den gjennomsnittlige høyden av trærne (Zimble et al. 2003). Dens uavhengighet i forhold til prøveflatestørrelse betyr at verdien er sammenlignbar for andre laserdataundersøkelser med samme utstyrsnivå, men ulik prøveflatestørrelse fra min oppgave. F. eks vil man finne en lav cv-verdi i fjellbjørkeskog hvor det er en jevn sjiktning (gir lavt standardavvik) og lav gjennomsnittlig trehøyde. Jeg har observert i felt at et fåtall av mine forekomster befinner seg i fjellbjørkeskog og innslag av store furutrær er med på å dra opp cv-verdiene mot nivåer som en flersjiktet furuskog i hogstklasse 5 ligger på. Fcv er med i regresjonsmodellen for x500, med andre ord de største trærne i datasettet (Tabell 6). Nettopp disse trærne ble vanligvis registrert i fjellbjørkeskog høyt over havet. I tillegg var de vanligvis langt unna vei og variabelen veiavstand er med i regresjonsmodellen for denne diameterklassen.

Det faktum at ulike TF-verdier først kom med i diameterklassene x400 og x500 i regresjonsmodellene kan forklares ut fra blant annet endret penetreringsgrad for laserpulsene. At de laveste TF-verdiene har en signifikant forklaringsgrad på forekomsten av furu >500 mm kan f. eks avhenge av at fjellbjørkeskjermen hadde en negativ innvirkning på penetreringsgraden. Det ble tross alt fløyet i begynnelsen av august da bladmassen til fjellbjørka har hatt tid til å utvikle seg. At de høyeste TF-verdiene har en signifikant forklaringsgrad for forekomsten av furu >500 mm kan f. eks forklares av at et ruvende furutre over bjørkeskjermen vil fange opp et mindretall av det totale antall pulser som reflekteres innenfor hele prøveflata sammenlignet med en forholdsvis jevnt sjiktet furuskog.

Mine prøveflater strekker seg over en forholdsvis stor skala når det gjelder høyde over havet og mange flater havner innenfor fjellskogkategorien hvor fjellbjørk dominerer (Tabell 4). Innimellom befinner det seg enkelte forekomster av store furutrær med en høyde på 10-14 meter. Slike flater vil ha en innvirkning på dataene, spesielt når det er forekomster av store furudimensjoner. Næset (2004b) har pekt på at en høy andel av lauvtrær i barskog påvirker nøyaktigheten delvis fordi kroneformen og variasjonene av dens utforming er forskjellig fra bartrærne. En bonitets- eller aldersstratifisering av prøveflatene har vist seg å kunne separere forskjellige skogtyper som man forventet ville gi forskjeller i fordelingen av høydepersentilene (Gobakken & Næset 2004 og Næset 1997). Ofte vil skogstrukturen endre seg fra kulturskoger til naturskoger og naturskoger befinner seg lengre unna vei enn kulturskoger. Naturskogen bærer preg av kontinuitet og er dermed ofte flersjiktet, uensaldret og med en ujevn trefordeling. Høyde- og tetthetspersentilene vil dermed få verdier som skiller seg fra en en-sjiktet skjøttet kulturskog med jevnaldrede trær.

Noe som ikke vises i resultatkapitlet er modelleringen for kun laservariabler. Her ble det for x350 hele 11 signifikante variabler med i sluttmodellen. Det virker som om en inkludering av variabelen høyde over havet hadde en innvirkning på antall variabler som ble signifikante i den videre variabelseleksjonen når både laser- og eksterne variabler var med i modelloppbyggingen.

4.2.1 Feltdatamaterialets størrelse

Figur 5 viser at samme nøyaktighet ved predikeringen kunne vært oppnådd ved adskillig færre flater, en halvering av datamaterialet og vel så det. Grunnen til at det ble brukt 500 gjentak her, mot 300 gjentak for Monte Carlo-simuleringen i Tabell 8 og Figur 4, er fordi resultatene kommer tydeligere frem. Dette gjelder spesielt for lave estimeringsandeler som i praksis innebærer at få flater inngår i hvert gjentak. Derfor ble simuleringen stoppet ved 5 % uttak, altså 55 flater. Et utvalg av såpass få flater per gjentak gjør det vanskelig å fange opp den totale variasjonen som datasettet mitt gjenspeiler. Hvis f. eks et tenkt gjentak innebærer at alle 55 flatene som tilfeldig blir plukket ut befinner seg innenfor fjellbjørkeskog uten furu er det opplagt at modellen vil mislykkes i å predikere forekomster av store furutrær med dette som grunnlag.

En typisk størrelse på ei prøveflate, basert på hva som ansees som standard prosedyre for skogtaksering, er 200 – 250 m² (Gobbaken & Næset 2008). Denne oppgaven baserer seg på 4-5 ganger så store prøveflater og denne størrelsen vil dermed øke sjansen for at en forekomst befinner seg innenfor prøveflata. Med tanke på at hver tredje flate inneholder minst én forekomst av store dimensjoner (Tabell 2) ville bruken av 250 m² prøveflater i min korridor ført til en mindre effektiv registrering og et større antall besøkte flater. 1000 m² store prøveflater og en drøy halvering av antall prøveflater synes som en god ramme rundt denne oppgavens metodikk (Figur 5).

4.3 Oppgaven som verktøy for videre undersøkelser

Hvorvidt det er mulig å bruke mine resultater for funn av store dimensjoner i et annet geografisk område avhenger ikke bare i graden av likhet med Tolga kommunes skogstruktur. Flaskehalsen vil være utstyret som brukes i datafangsten. Sammenhengen mellom skoglige parametre og laserdata avhenger også av egenskapene til laserdataene. Nilsson (1996) peker på at forskjellig fotavtrykkstørrelse gir forskjellige estimat på trehøyde og andre skogvariabler. Næset (2004a) hevder at denne effekten sannsynligvis kommer bedre til uttrykk for siste returen enn for den første. En annen faktor som innebærer at mine resultater ikke er direkte flyttbar til andre områder, beror på mitt datamateriale som kun er basert på furuforekomster. Dette vil gjøre metoden uegnet i skoger med en annen treslagsfordeling hvor blant annet store grantrær forekommer. Hvis f. eks laserområder i Lom i Gudbrandsdalen skal undersøkes for store dimensjoner av furu, må det på lik linje etter denne oppgavens metodiske ramme, etableres øvingsområder med etterfølgende feltarbeid og registrering av store dimensjoner. Figur 5 viser at man kan klare seg med drøyt halvparten av denne oppgavens datamengde. Man kan altså klare seg, med noenlunde lik nøyaktighet, med 600 prøveflater.

Det er stor variasjon i design og konfigurering av ALS-instrumentene (Lim *et al.* 2003). Uansett teknologisk utvikling i forbindelse med ALS, så vil mine prøveflater med registreringene være koordinatfestet og en ny flyvning med andre utstyrsspesifikasjoner vil kunne gi andre laservariabler som best predikerer forekomsten av store dimensjoner. Lim & Treitz (2004) peker på at mange undersøkelser bærer preg av at forklaringsvariablene er for stedsspesifikke og relevante kun for undersøkelsens formål. De kan dermed ikke bli generalisert til å gjelde andre skogtyper eller voksesteder.

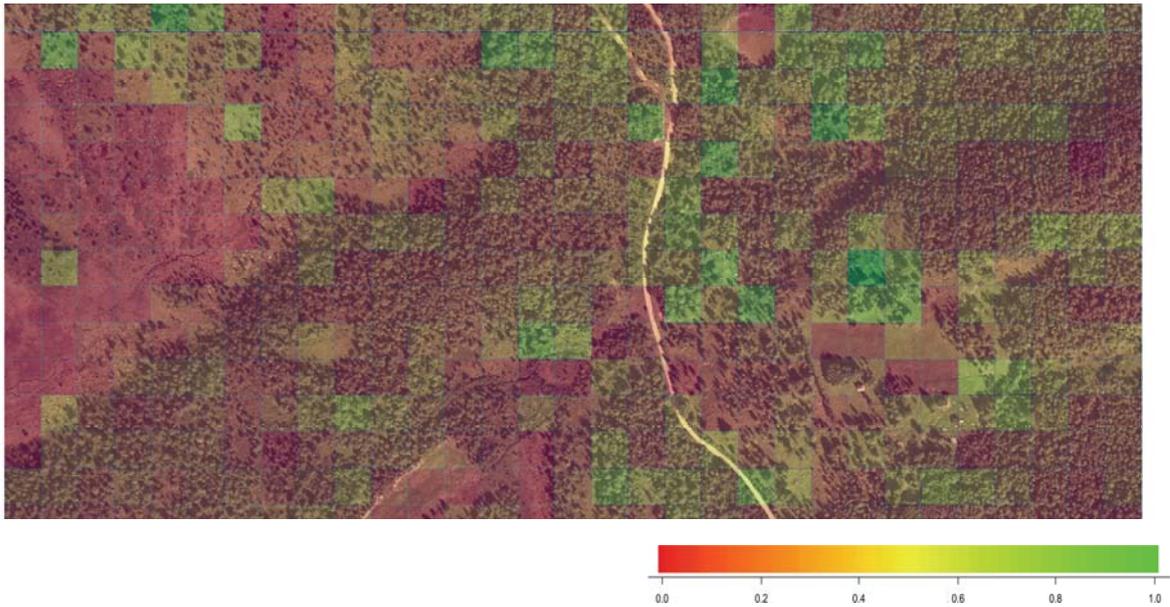
Forskere har diskutert hvilken innvirkning en mer intensiv datainnsamling, altså større punkttetthet, har på høydeinndelingen og dermed den høydemessige fordelingen av høydepersentilene. Lim *et al.* (2008) peker på at hvis punkttettheten har noe å si for inndelingen av høydepersentilene, kan man vanskelig fastslå om det er punkttettheten selv eller de forskjellige faktorene ved selve undersøkelsen som er årsaken. Dette dreier seg om f. eks pulsfrekvens, scanningsvinkel, scannefrekvens, eller flyets hastighet og høyde. Det er vanskelig å oppnå noen synergieffekt mellom de ulike parametrene da f. eks en større scanningsvinkel vil gi en lavere scannefrekvens siden sveipemekanismen vil bevege seg over en lengre avstand for hvert sveip. I deres undersøkelse med forskjellig punkttetthet var høyde- og tetthetspersentilene lite sensitiv for laserdata hentet fra to ulikt konfigurerte flyvninger og forskjellige prøveflater. Ulike resultater mellom de to datasettene ble funnet for første returen i begge ender av persentilene, altså maksimums- og minimumsverdi.

4.3.1 Grafisk sluttresultat – et tenkt eksempel

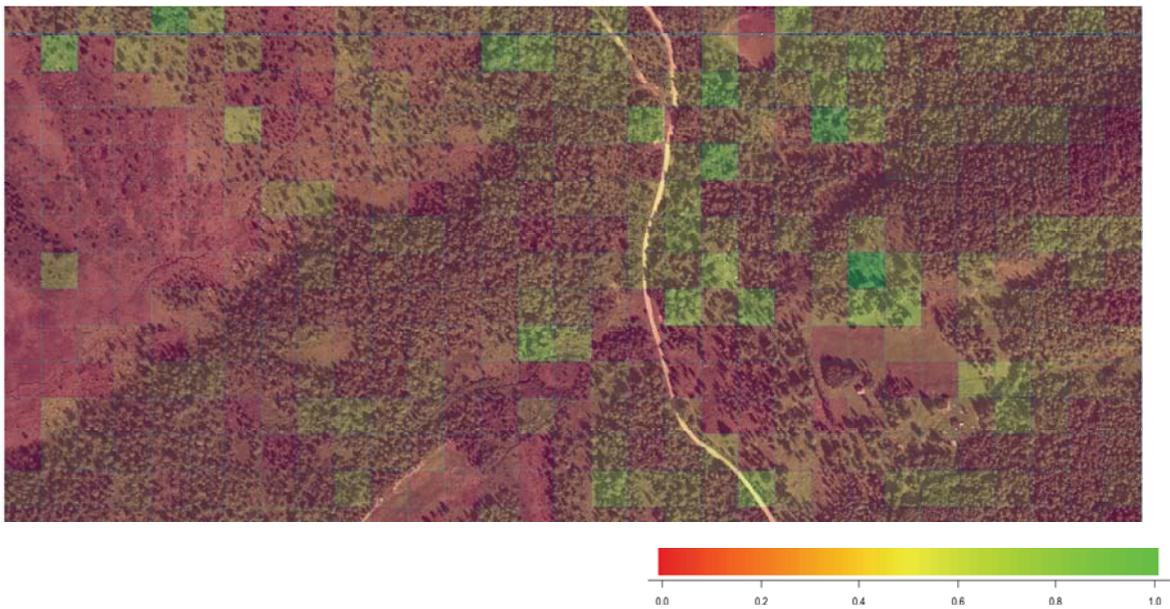
Resultatene fra min oppgave kan framstilles i et grafisk design som Figur 7, 8 og 9 viser. Figurene tar utgangspunkt i et utsnitt fra laserstripe 6 i nærheten av Olaberget. Laserscannets bredde er blitt inndelt i et grid-system av prøveflater med samme størrelse som mine korridorflater, 1000m². Utsnittet representerer forskjellige skogstrukturer med varierende tetthet og andre landskapsmessige trekk som blant annet innmark og tresatte myrområder. Områdets utforming gjenspeiles i flyfotoet i figur 6. Predikeringen av sannsynligheten for funn av store trær baserer seg på modellene for gjeldende diameterklasser fra tabell 7. Et kartlag med fargegradienter er lagt oppå flyfotoene for å gi et grafisk bilde av hvor sannsynligheten er størst for å finne store trær.



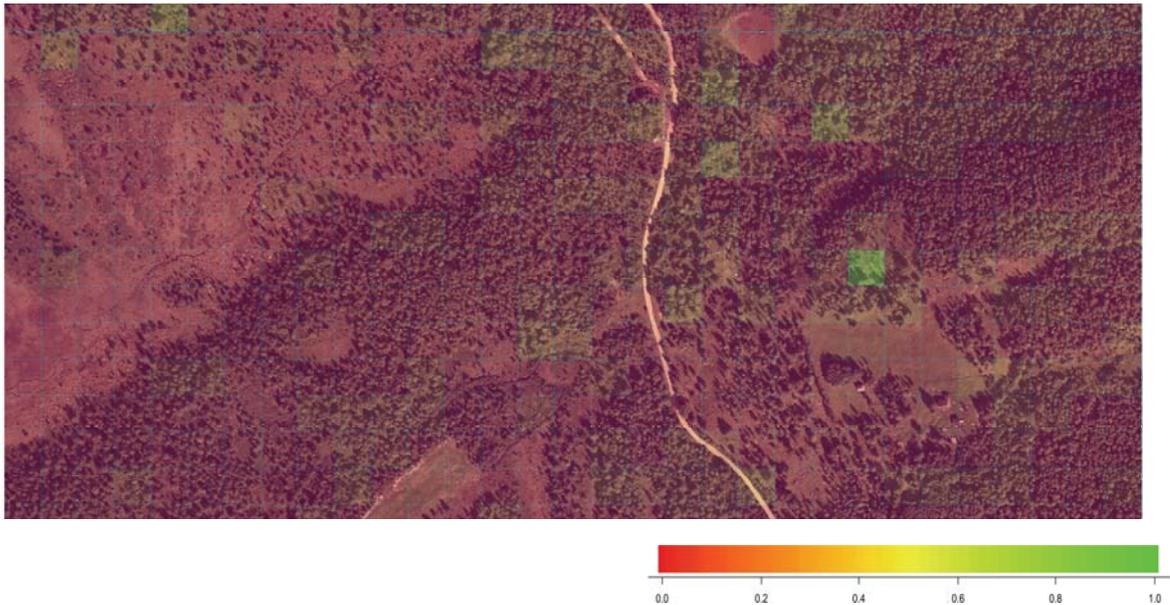
Figur 6: Utsnitt av flyfoto i nærheten av Olaberget



Figur 7: Utsnitt av flyfoto i nærheten av Olaberget. Sannsynlighet for funn av furu >350 cm i brysthøydiameter. Fargeskalaen angir sannsynlighetsnivåer.



Figur 8: Utsnitt av flyfoto i nærheten av Olaberget. Sannsynlighet for funn av furu >400 cm i brysthøydiameter. Fargeskalaen angir sannsynlighetsnivåer.



Figur 9: Utsnitt av flyfoto i nærheten av Olaberget. Sannsynlighet for funn av furu >500 cm i brysthøydediameter. Fargeskalaen angir sannsynlighetsnivåer.

5. Konklusjon

Resultatene viser at flybåren laser kan benyttes til å finne fram til områder med furutrær med store dimensjoner. Regresjonsmodellens prediksjon av forekomst av store furutrær >350 mm stemmer overens med 63 % av registrerte forekomster. Prosentandelen synker ved prediksjon av større furutrær, >400 mm og >500 mm. Feltarbeidet og datainnsamlingen kan med god margin kuttes ned til drøyt halvparten av totalt antall besøkte flater uten at nøyaktigheten til sluttmodellene vil svekkes nevneverdig (Figur 5) og dette vil gjøre datainnsamlingen vesentlig rimeligere. I denne oppgaven er det fokusert på å finne fram til aktuelle områder. I framtidige studier bør det undersøkes hvor mange trær som finnes innenfor de aktuelle områdene og hvor høye disse trærne er.

I sluttmodellene var høyde over havet en forklaringsvariabel som viste seg å gi signifikante verdier for predikeringen av store dimensjoner for alle diameterklassene. Ut ifra feltmessige erfaringer sammenlignet med grafisk fremstilling fra ArcGIS er det enkelte steder man bør studere høyde over havet og bonitetskart nøye. Modellen predikerer 50 % sjans for funn av store dimensjoner, men materialutnyttelsen, sett med Materialbankens øyne, innskrenkes når det f. eks viser seg at treet er 10 meter høyt.

6. Litteratur

- Baltsavias, E. P. (1999). Airborne laser scanning: basic relations and formulas. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54: 199–214.
- Blom Geomatics AS (2006). Rapport BNO06711. Hedmark skog 2006. Intern rapport.
- 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: 33-47.
- Deckert, C., Bolstad, P. V. (1996). Forest canopy, terrain and distance effects on global positioning system point accuracy. *Photogrammetric Engineering & Remote Sensing*, 62 (3): 317-321.
- Gobakken, T. & Næsset, E. (2005). Weibull and percentile models for lidar-based estimation of basal area distribution. *Scandinavian Journal of Forest Research*, 20 (6): 490-503.
- Gobakken, T. & Næsset, E. (2008). Assessing effects of laser point density, ground sampling intensity, and field sample plot size on biophysical stand properties derived from airborne laser scanner data. *Canadian Journal of Forest Research*, 38: 1095-1109.
- Hoeting, J. A., Davis, R. A., Merton, A. A. & Thompson, S. E. (2006). Model selection for geostatistical models. *Ecological Applications*, 16 (1): 88-98.
- Holmgren, J. (2003). Estimation of Forest Variables using Airborne Laser Scanning. *Doctoral dissertation. ISSN 1401-6230, ISBN 91-576-6512-5*
- Holmgren, J. & Persson, A. (2004). Identifying species of individual trees using airborne laser scanner. *Remote Sensing of Environment*, 90: 415-423.
- Jakobsons, A. (1970). The correlation between the diameter of the tree crown and other tree factors – mainly the breast-height diameter. Analysis based on sample trees from the National Forest Survey. *Report 14, Department of Forest Survey, Royal College of Forestry, Stockholm, Sweden. 75 s.*
- Kilian, J., Haala, N. & Englich, M. (1996). Capture and evaluation of airborne laser scanner data. *International Archives of Photogrammetry and Remote Sensing*, Tilgjengelig fra:
http://www.ifp.uni-stuttgart.de/publications/1996/laser_wien.pdf
(lest 03.05.2010).
- Kissling, W. D., & Carl, G. (2008). Spatial autocorrelation and the selection of simultaneous autoregressive models. *Global Ecology and Biogeography*. 17: 59–71.
- Lim, K. S., Treitz, P. M., Wulder, M., St-Onge, B. & Flood, M. (2003). LiDAR remote sensing of forest structure. *Progress in Physical Geography*, 27: 88-106.

- Lim, K. S. & Treitz, P. M. (2004). Estimation of above ground forest biomass from airborne discrete return laser scanner data using canopy-based quantile estimators. *Scandinavian Journal of Forest Research*, 19: 558-570.
- Lim, K. S., Hopkinson, C. & Treitz, P. M. (2008). Examining the effects of sampling point densities on laser canopy height and density metrics. *Forestry Chronicle*, 84: 876-884.
- Maltamo, M., Næsset, E., Bollandsås, O. M., Gobakken, T. & Pacalen, P. (2009). Non-parametric predictions of diameter distributions using airborne laser scanner data. *Scandinavian Journal of Forest Research*, 24: 541-553.
- Mehtätalo, L., Maltamo, M., & Packalén, P. (2007). Recovering plot-specific diameter distribution and height-diameter curve using ALS based stand characteristics. *IAPRS, Vol. XXXVI, Part 3/W52*
- Nilsson, M. (1996). Estimation of tree heights and stand volume using an airborne lidar system. *Remote Sensing of Environment*, 56: 1-7.
- 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.
- Næsset, E. (1997b). estimating timber volume of forest stands using airborne laser scanner data. *Remote sensing of environment*, 61 (2): 246-253.
- Næsset, E., Bjerke, T., Avstedal, O. & Ryan, L. H. (2000). Contributions of differential GPS and GLONASS observations to point accuracy under forest canopies. *Photogrammetric Engineering & Remote Sensing*, 66 (4): 403-407.
- Næsset, E. (2002). Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sensing of Environment*, 80: 88-99.
- Næsset, E., Gobakken, T., Holmgren, J., Hyypää, H., Hyypää, 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 (6): 482-489.
- Næsset, E. (2004a) 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.
- Næsset, E. (2004b). Practical large-scale forest stand inventory using small-footprint airborne laser scanner. *Scandinavian Journal of Forest Research*, 19: 164-179.
- 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.

- Næsset, E. & Gjevestad, J. G. (2008). Performance of GPS precise point positioning under conifer forest canopies. *Photogrammetric Engineering & Remote Sensing*, 74 (5): 661-668.
- Næsset, E. (2009). Effects of different sensors, flying altitudes, and pulse repetition frequencies in forest canopy metrics and biophysical stand properties derived from small-footprint airborne laser data. *Remote Sensing of Environment*, 113: 148-159.
- Persson, Å., Holmgren, J. & Söderman, U. (2002). Detecting and measuring individual trees using an airborne laser scanner. *Photogrammetric Engineering & Remote Sensing*, 68: 925-932.
- Popescu, S. C., Wynne, R. H. & Nelson, R. F. (2003). Measuring individual tree crown diameter with lidar and assessing its influence on estimating forest volume and biomass. *Canadian Journal of Remote Sensing*, 29 (5): 564-577.
- Popescu, S. C. & Wynne, R. H. (2004). Seeing the trees in the forest: Using lidar and multispectral data fusion with local filtering and variable window size for estimating tree height. *Photogrammetric Engineering & Remote Sensing*, 70: 589-605.
- R Development Core Team (2007). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
- Salas, C., Ene, L., Gregoire, T. G., Næsset, E. & Gobakken, T. (2010). Modelling tree diameter from airborne laser scanning derived variables: A comparison of spatial statistical models. *Remote Sensing of Environment*, 114: 1277-1285.
- Sandnes, T. (1942). Skogen og skogbruket i Rørosdistriktet. Rørosboka, bind 1, 147-179.
- Venn, L. (2010a). Krok og andre "feil" til heder og verdighet. *Norsk Skogbruk*, No. 2 – 2010, 20-21.
- Venn, L. (2010b). Materialbank på rot til tradisjonshåndverk. *Norsk Skogbruk*, No. 2 – 2010, 22-23.
- Wehr, A., Lohr, U. (1999). Airborne laser scanning—an introduction and overview. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54: 68-82.
- Wever, C. & Lindenbeger, (1999). Experiences of 10 years laser scanning. *Proc. Photogramm. Week*, 125-132.
- Zhang, L. (1996). Cross-validation of non-linear growth functions for modeling tree height +- diameter relationships. *Annals of Botany*, 79: 251-257.
- Zimble, D. A., Evans, D. L., Carlson, G. C., Parker, R. C., Grado, S. C. & Gerard, P. D. (2003). Characterizing vertical forest structure using small-footprint airborne LiDAR. *Remote Sensing of Environment*, 87: 171-182.

Personlige meddelelser:

Andresen, K. (2010). Riksantikvaren. (e-post til Torgrim Østgård 21.03.2010)

Bøhle, K. L. (2010). Blinken (e-post til Torgrim Østgård 04.05.2010)

Groesz, F. (2010). Project Manager Forestry – Blom Geomatics AS (e-post til Torgrim Østgård 18.03.2010)

Vedlegg



Kartet viser alle 8 laserstripene i Tolga kommune. Grønt markerer besøkte flater.