



Norges miljø- og  
biovitenskapelige  
universitet

**Masteroppgave 2018 30 stp.**

Fakultet for miljøvitenskap og naturforvaltning

## **Klassifisering av endringer ved hjelp av flybåren laserdata**

**Ola Doksrød Strande**

Skogfag



## Forord

Denne masteroppgaven marker avslutningen på min skogfagsutdannelse ved Fakultetet for miljøvitenskap og naturforvaltning (MINA), ved Norges miljø- og biovitenskapelige universitet (NMBU).

Jeg vil rette en stor takk til mine veiledere ved MINA, overingeniør Roar Økseter og Dr. Ole Martin Bollandås. De har hjulpet meg gjennom hele prosessen med å behandle og sortere data, gjøre statistiske beregninger- og modelleringer, og med tekstbehandling. Stipendiat Lennart Noordermeer må også takkes for god hjelp og gode innspill. Takk til medarbeidere under feltarbeidet i Krødsherad, og takk til medstudenter for en fin studietid på Ås.

Norges miljø- og biovitenskapelige universitet

Ås, 14. mai 2018

Ola Doksrød Strande



## Sammendrag

Skogregistreringer og taksering av skog gjøres i dag ved hjelp av flybåren laserskanning (FLS), en teknikk som gir informative data om skogbestands trehøyder og vegetasjonstetthet. De senere årene har enkelte skogområder i Norge blitt taksert med FLS for andre gang, noe som gjør at det er mulig å benytte FLS- data til å identifisere endringer i disse skogområdene. I et miljø- perspektiv, med tanke på avskoging av tropisk regnskog, er det derfor tenkelig at FLS kan brukes til å identifisere avskogede og degraderte områder. Identifisering av endringer er aktuelt også for andre formål, blant annet for en ny type boniteringsmetode som er under utvikling ved Norges miljø- og biovitenskapelige universitet (NMBU). Denne metoden avhenger av FLS- data fra to tidspunkt, som stammer fra skogområder som har hatt en naturlig og upåvirket utvikling. Det var dermed ønskelig å finne en metode som kunne brukes som klassifikator i boniteringsmodeller, og/eller for å identifisere avskogede og degraderte områder i tropisk regnskog.

Datagrunnlaget stammet fra et 50.000 dekar stort område som lå i Krødsherad kommune i Buskerud fylke. I 2001 (T1) og 2016 (T2) ble det samlet inn data ved hjelp av FLS. Det ble også gjort feltmålinger, der data fra 148 prøveflater (233m<sup>2</sup>) ble benyttet i de statistiske beregningene. Det ble laget logistiske regresjonsmodeller som predikerte endringsklasser på prøveflatene, basert på forklaringsvariabler ekstrahert fra FLS- dataene. Disse endringsklassene var *uberørt*, *tynnet* og *avvirket*. De predikerte klassene ble sammenlignet med feltregistrerte endringsklasser på prøveflatene og de beste modellene var de som klassifiserte flest prøveflater korrekt.

Det ble undersøkt hvilke forklaringsvariabler som gav de beste modellene. De beste modellene inneholdt informasjon om skogens tetthet, ca. midt i vegetasjonens høydesjikt. Den beste modellen klassifiserte totalt 85.8% av prøveflatene korrekt. Modellene klassifiserte avvirkede prøveflater med høy presisjon. Det var imidlertid en utfordring å klassifisere tynnede prøveflater korrekt, disse ble ofte klassifisert som uberørte. I tillegg ble det undersøkt to ulike metoder å benytte forklaringsvariablene i modelleringen. I Metode 1 ble det benyttet differanseverdier mellom variabler fra T1 og T2, som forklaringsvariabler. I Metode 2 ble forklaringsvariabler fra både T1 og T2 benyttet i modellene. Det ble funnet at Metode 2 ikke ga bedre resultater enn Metode 1, men i denne sammenheng var det tenkelig at Metode 1 var best egnet til å klassifisere endringer. Modellene ble validert på 60 kontrollflater (ca. 3700 m<sup>2</sup>), og ble funnet å være gyldige.



## Abstract

Forest inventory and forest- registration are done by means of airborne laser scanning (ALS), a technique that gives informative data about tree heights and the density of the forest`s vegetation. Some forest areas in Norway have had a second time of ALS- measurement, which makes it possible to compare ALS- data from two different points of time, and identify changes in these areas. For environmental purposes, considering deforestation and degradation in tropical forests, these kinds of changes may be detected by means of ALS. Identification of changes is also applicable for other purposes. At the Norwegian University of Life Sciences (NMBU), scientists are investigating a new method that can be used for site index (SI) measuring in forests, by means of ALS. This method depends on ALS- data from two points of time, and the data have to come from undisturbed and untouched forest areas. Therefore, we wanted to develop a method that could exclude areas that was not suitable for the new SI- measuring method. We also needed a method that could identify deforestation and degradation in tropical forests.

The data was collected from a 50.000 decaar forest area in Krødsherad municipality, in the county of Buskerud, Norway. In 2001 (T1) and 2016 (T2) data was collected by means of ALS. Field- data was also collected, and field data from 148 plots (233m<sup>2</sup>) was used in the statistical calculations. Different variables were created on basis of ALS- data and was used as explanatory variables in logistic regression models, for estimation of changes. Three classes of changes were used; *untouched*, *thinned* and *harvested*. The models predicted these classes of changes on the 148 plots, and these classes was compared with observed classes based on field- data. The best models were those that predicted a high amount of the plots in the right classes.

We wanted to investigate which explanatory variables that gave the best models. The best models consisted of explanatory variables with information based on the forest`s vegetation - density, from the middle of the vegetation height. The best model classified in a total 85.8 % of the plots in the correct classes. The models classified *harvested* plots with a high precision. However, *thinned* plots were often confused with *untouched* plots We also compared two different methods of using explanatory variables in the models. For Method nr.1 we used the difference between variables from T1 and T2 as explanatory variables. For method nr. 2 we used explanatory variables from both T1 an T2 in our models. We found that Method nr. 2 was no better than Method nr.1, but we considered Method nr.1 as the best when it comes to predicting class changes. The models were validated on 60 control plots and were found to be valid.





# Innholdsfortegnelse

Forord .....	III
Sammendrag .....	V
Abstract .....	VII
1. Innledning.....	1
1.1 Bakgrunn .....	1
1.2 Flybåren laserskanning (FLS) .....	4
1.3 Flybåren laserskanning for prediksjon av klasseendringer .....	5
1.4 Problemstilling.....	6
2. Materiale og metode.....	7
2.1 Materiale .....	7
2.1.1 Studieområdet.....	7
2.1.2 Datagrunnlag .....	8
Feltmålinger 2001.....	8
Feltmålinger 2016/ 2017 .....	9
2.1.3 Laserdata .....	10
2.2 Metode .....	11
2.2.1 Behandling av datamaterialet .....	11
2.2.2 Klassifisering av feltmålte prøveflater .....	12
2.2.3 Tilpassing av logistiske regresjonsmodeller til å klassifisere endringer .....	14
2.2.4 Kryssvalidering og validering av modellene .....	16
2.2.5 Kriterier for rangering av modellene.....	16
2.2.6 Klassifisering av endringer ved bruk av deltvariabler og parede variabler .....	18
3. Resultater .....	19
4. Diskusjon .....	22
4.1 Sammenligning av modellene med hensyn på forklaringsvariablene.....	22
4.2 Sammenligning av Metode 1 og Metode 2 .....	23
4.3 Validering av modellene på kontrollflatene .....	24
4.4 Materiale og metode.....	25
4.5 Klassifisering av endringer ved hjelp av laserdata til bruk i boniteringsmodellering og for overvåking av tropisk regnskog.....	27
5. Konklusjon .....	28
6. Litteratur.....	29
7. Vedlegg .....	31



# 1. Innledning

## 1.1 Bakgrunn

Overvåking av endringer i skog er et dagsaktuelt tema. For land der tropisk regnskog er utbredt, er det viktig med nøyaktige estimater for mengde og endringer i biomasse og utbredelse av skog. Slik kan man påvise eventuell avskoging og overutnyttelse av skogressurser samt karbonbinding. Skogområders evne til å binde CO<sub>2</sub> er aktuelt med tanke på global oppvarming og klimaendringer. Skogen er en av de viktigste bidragsyterne til å binde karbon og dempe klimaendringene (e.g. Dymond et al., 2016). Det er derfor av stor betydning for det globale klimaet og miljøet å hindre avskoging og degradering, slik at ikke mer CO<sub>2</sub> frigjøres til atmosfæren.

Internasjonalt spiller FN en viktig rolle i overvåking og rapportering av klimagassutslipp. FNs klimakonvensjon (UNFCCC) er en rammekonvensjon ratifisert av 195 medlemsland, for å regulere konsentrasjonen av klimagasser i atmosfæren «*slik at det stabiliseres på et nivå, som forhindrer en farlig og negativ menneskeskapt påvirkning på klimasystemet*» (Miljødirektoratet, 2017). I 1997 holdt UNFCCC partsmøte i Kyoto, Japan, der *Kyotoavtalen* ble signert. Til sammen har 191 land og EU ratifisert denne avtalen, som setter tallfestede utslippsforpliktelser for industrilandene. Utslippsforpliktelsene er forskjellige fra land til land, avhengig av hvor vanskelig det er for landene å oppnå forpliktelsene. *Kyotoavtalen* stiller krav til medlemslandene at de holder kontroll på, og rapporterer om utslippsmengdene sine (Miljødirektoratet, 2017).

I henhold til *Parisavtalen* av 2015, ble medlemslandene i UNFCCC enige om at den globale oppvarmingen ikke skal overstige en økning på 2 grader celsius, sammenlignet med førindustriell tid. Dette målet vil være umulig å nå hvis ikke utslippene knyttet til avskoging og uthogging av klimamessig viktige skogområder reduseres. I følge FNs klimapanel utgjør utslipp fra hogst av regnskog om lag en sjettedel av verdens totale klimagassutslipp (Barker et al., 2017). For å begrense utslippene fra avskoging har UNFCCC opprettet REDD- programmet (Reducing emissions from deforestation and forest degradation). Programmet knytter karbonbinding i skog og biomasse til en finansiell verdi, og er ment som en økonomisk insentiv og kompensasjon til utviklingsland, fra FNs industriland, for å redusere klimagassutslipp fra avskoging og uthogging av skogområder (UNREDD, 2017). Utviklingslandene mottar økonomiske insentiver dersom skogbruket drives på en bærekraftig og klimavennlig måte, og dersom landene kan vise at skogbrukstiltakene fører til økt CO<sub>2</sub>- fangst og reduserte utslipp.

Det er utviklingsland med store områder med tropisk skog som har vært i fokus hos REDD - programmet. Økonomiske incentiver for å hindre utslipp fra skogsektoren er aktuelt også for annen type skog, og i andre deler av verden. I følge Olerud (2014) lagrer nordlige boreale skoger mer karbon enn tropiske. Dette skyldes at karbon fra gamle, råtne trær i boreal skog i mindre grad enn i tropisk skog slippes tilbake i atmosfæren, og at karbonet lagres i mer stabile karbonforbindelser i bakken. Derfor bør det vurderes om det også i Norge skal gis økonomisk kompensasjon til skogeiere for å drive et mer klimavennlig skogbruk med økt fokus på karbonfangst og vern av skogområder. I henhold til *Kyotoavtalen* er Norge, på lik linje med alle *Kyoto-land*, blitt tildelt et utslippstak, en kvote, på hvor mye CO<sub>2</sub> som årlig kan slippes ut (Olerud, 2014). Innad i landet kan myndighetene dele ut eller selge en viss mengde kvoter til bedrifter og private aktører. En måte skogeiere skal kunne få økonomisk kompensasjon/ utbytte av karbonet som bindes i skogen sin kan være denne; at skogeiere kan selge en kvote til private bedrifter tilsvarende den mengden karbon som skogen deres binder, for så å sørge for at skogen deres binder mest mulig CO<sub>2</sub> (ved å skjøtte den optimalt, eller ved å verne den).

For at industrilandene skal bli forsikret om at deres økonomiske innsats i REDD- prosjektet faktisk har ført til resultater, er det nødvendig at det finnes et pålitelig system for å detektere uønskede endringer slik som avskoging og degradering av skog. Dette gjelder også for scenarioet om salg av klimakvoter fra private skogeiere i Norge. Et overvåkingssystem må være på plass slik at myndighetene, eller private kjøpere av klimakvoter, kan etterprøve skogeiere som selger kvoter. Et slikt system vil da kunne ta utgangspunkt i skogbruksplaner og skogtakster, som inneholder informasjon om skogens stående kubikkmasse og biomasse på et gitt tidspunkt. I Norge gjennomføres takster i forbindelse med skogbruksplanlegging med ca. 10- 15 års mellomrom. Ved at man har skogtilstanden på to ulike tidspunkt, vil man dermed også ha informasjon om hvilke endringer som har skjedd i skogen i løpet av denne tiden. Disse endringene vil vise om biomassen har økt eller redusert, og om skogen har bidratt til å binde CO<sub>2</sub> eller frigjøre CO<sub>2</sub> til atmosfæren.

Informasjon om endringer i skog er nyttig også i andre sammenhenger, blant annet til estimering av produktivitet (bonitet). Bonitetsverdien er et mål på hvor god markas evne er til å produsere trevirke. I Norge måles bonitet som den akkumulerte dominerende høyden i et bestand ved en indeks- alder på 40 år i brysthøyde (Tveite, 1977). Faktorer som påvirker et bestands bonitet er klimaet; temperatur, vanntilgang, lystilgang og luftforhold, og jord- og grunnforhold (Børset, 1985). Disse faktorene påvirker bestandet under hele livsløpet. I denne sammenheng defineres bestandets livsløp som tiden fra overhøyden passerer 1,3 meter (brysthøyde), til det aktuelle året for registrering. Boniteten som blir målt ut i fra dette kriteriet kalles «livsløpsbonitet». Dersom de ovenfor nevnte faktorene endres over tid, for eksempel som resultat av klimaendringer, vil også boniteten endres. En

dramatisk endring i vekstforholdene vil imidlertid ikke ha stor effekt for livsløpsboniteten, siden økningen i vekstrate i en relativt kort periode av bestandets livsløp blir fordelt over hele livsløpet. Bonitet kan imidlertid også vurderes over en kortere periode. Interceptmetoden (e.g. Bull, 1931) er en slik metode der lengden på et gitt antall skuddstrekninger vurderes, med utgangspunkt i en gitt høyde (for eksempel brysthøyde). På tilsvarende vis kan man også vurdere siste års skuddstrekninger. Det er dog en utfordring knyttet til å bruke siste års skuddstrekninger kontra skuddstrekninger med utgangspunkt i en fast høyde over bakken. De siste års skuddstrekninger kan ikke brukes som et direkte mål på produktiviteten, fordi høydeveksten avtar etter hvert som trærnes alder øker. Ved varierende høydeutgangspunkt er en gitt vekst vanskelig å tolke. Har man imidlertid utgangshøyden, høydeveksten og hvor mange år denne veksten fordeler seg over, kan disse parameterne jmføres med bonitetskurvene til Tveite og Braastad (Tveite, 1977; Tveite& Braastad, 1981). Som omtalt over, er disse kurvene avhengig av alder og høyde, men en gitt kombinasjon av utgangshøyde, høydevekst og vekstperiode har kun en løsning i settet med bonitetskurver. Sammen med intercept- metoden, kan dette defineres som en *høydedifferensiell* bonitering.

For noen år tilbake var takst av skog og registrering av skogparametere basert på feltmålinger fra prøveflater eller linjetakster, i tillegg til visuell informasjon fra flybilder. Senere har data fra fjernmåling blitt en sentral del av slike takster (e.g. Næsset, 2004). Fjernmålingsteknikken som i dag benyttes er såkalt flybåren laserskanning (FLS), som gir gode data om hvert bestands trehøyder. Registrering av bestandvis bonitet i operasjonelle takster gjøres i dag på bakgrunn av gammel takst. Ut fra flyfoto har man oversikt over bestandsgrensene i skogområdet. Bestandets alderen fra gammel takst framskrives, og jmføres med trehøyder beregnet fra flybåren laserdata (Korsvold, 2018). Dersom de nye trehøydene ikke stemmer overens med bonitet fra gammel takst, bli det gjerne gjort nye feltmålinger av alder og dominerende høyde i disse bestandene.

Ved Norges miljø- og biovitenskapelige universitet (NMBU) forskes det på nye, mer effektive metoder å bonitere bestand på (Noordermeer et al., vurderes for publisering), ved hjelp av FLS- data. På bakgrunn av den høydedifferensielle boniteringsmetoden, er det mulig å bonitere med FLS- data, hvis man har data over samme område fra to forskjellige tidspunkt. Metoden har imidlertid noen begrensninger. Området/ bestanden må ikke ha blitt påvirket av eksterne faktorer, fordi boniteten beregnes ut fra en naturlig utvikling i bestandets dominerende høyde. Dersom bestanden har vært utsatt for menneskelig påvirkning, kan det være en fare for at dominerende høyde har hatt en unaturlig utvikling, for eksempel som følge av en høytytning eller sluttavvirkning. Slike uønskede endringer (for boniteringsformål) kan også komme som følge av naturlige prosesser, slik som vindfall, toppbrekk eller unormal mortalitet som følge av insektskader. Ved bruk av FLS- data er det viktig å ha oversikt over hvilke områder i skogen som er uegnet til å bruke for denne nye typen

boniteringsmodeller. I studien til Noordermeer et al (vurderes for publisering) ble en variabel som representerte gjennomsnittlig dominerende høyde basert på laserdata, og laservariabler fra vegetasjonens øverste sjikt, benyttet som klassifikator. Dersom områder hadde hatt en tilbakegang i en av disse laservariablene, mellom første og andre gangs laserskanning, ble områdene fjernet fra datasettet.

## 1.2 Flybåren laserskanning (FLS)

FLS er en fjernmålingsteknikk basert på såkalt Light detection and ranging (LiDAR), der pulser av nærinfrarødt laserlys sendes ut fra et instrument (sensor), som så returneres fra objektene som laserpulsene treffer (TerratecAS, u.å.). Sensoren registrerer tidspunktet for når laserpulsen blir sendt ut, og når den blir returnert. Ved å multiplisere halve denne tidsdifferansen med lysets hastighet, kan avstanden mellom sensoren og objektet beregnes.

Dette prinsippet ligger til grunn for laserskanning av landområder, og terrengmodellering. Et fly flyr over et område med en påmontert LiDAR- sensor på undersiden, som sender laserpulser og mottar retursignaler. Pulsene fra sensoren fordeles også fra side til side. Teknikken benyttes blant annet for digitale terrengmodelleringer (DTM- Digital Terrain Model) av bakken, der hus, trær og andre objekter på overflaten blir filtrert vekk. Teknikken benyttes også for overflatemodellering (DSM- Digital Surface Model), som inneholder laserdata fra bakken og objekter som står på bakken (Terratec AS, u.å.). Bruk av FLS for skoginventering er en form for overflatemodellering, der både bakken og skogen modelleres. Dataene fra skanningen gir detaljert informasjon om skogens og trærnes høyde og tetthet, som er to viktige variabler for estimering av parametere som for eksempel kubikkmasse og biomasse.

Generelt finnes det to forskjellige teknikker som brukes for FLS: diskretpulssystem (pulse ranging system/ discrete return system) og kontinuerligbølgesystem (continuous wave system). Diskretpulssystemet er den metoden som brukes mest ved skoginventering (Næsset, 2004). Metoden baserer seg på at sensoren hvert sekund sender ut ca. 800 000 laserpulser (RIEGL Laser Measurement Systems GmbH, 2015), der hvert treffpunkt på et objekt, et såkalt fotavtrykk, har en radius på ca. 20 cm. Hver laserpuls som sendes kan ofte ha flere returer, såkalte ekkosignaler, knyttet til seg. Da flybåren laserskanning ble tatt i bruk på starten av 2000- tallet for skogregistreringsformål, var det kun et første og et siste ekko som ble registrert. Grunnet teknologiutvikling er det i dag vanlig å registrere opptil 4-5 ekko for hver puls. Dette skyldes at fotavtrykket har treff på forskjellige steder i høydesjiktet til vegetasjonen, og fotavtrykket deler seg i flere ekko.

For skoginventering skiller det mellom to ulike metoder med hensyn på geografisk oppløsning; arealmetoden (Næsset, 2002) og enkelttremetoden (Maltamo et al., 2004). Arealmetoden er den

vanligste metoden for estimering av kubikkmasse og biomasse. Metoden tar utgangspunkt i at det legges ut prøveflater i området som skal takseres, og disse oppsøkes i felt. Det gjøres målinger av trærne på disse prøveflatene, slik at man har en feltreferanse for området. Det blir også gjort en laserskanning av hele området, som også dekker områder som ikke ble feltregistrert. Fra laserskanningen ekstraheres laserdata fra prøveflatene, og en sammenheng mellom laserdata og feltdata kan bli funnet ved hjelp av regresjonsanalyse. Hele arealet deles inn i celler, og for hver celle predikeres det en verdi; volum, treantall, overhøyde, og/ eller andre relevante parametere. Til slutt summeres cellene i henhold til hvilke bestand de tilhører, og skoglige parametere kan beregnes for hele bestandet.

Enkelttremetoden baserer seg på å beregne en parameter (f.eks. volum) for hvert enkelt tre i bestandet, og legge sammen alle enkelttrærne for å få en samlet verdi for hele bestandet. En generell svakhet ved FLS er at teknikken er upresis når det gjelder å estimere stammeantall (Næsset, 2004; Maltamo et al. 2004), spesielt i flersjiktet skog, da denne metoden er mer avhengig av å identifisere alle trær. Arealmetoden er ikke så avhengig av å identifisere alle trærne, siden denne metoden forholder seg til mer aggregerte verdier. Derfor er arealmetoden den mest anvendte metoden i skoginventeringssammenheng.

### 1.3 Flybåren laserskanning for prediksjon av klasseendringer

Tidligere studier har vist at det er mulig å estimere endringer i skog ved hjelp av flybåren laserdata. Bollandås et al. (2013) viste at det var mulig å modellere og predikere endringer i biomasse mellom to tidspunkt, ved bruk av FLS- data innhentet med diskrepulsmetoden. I denne studien ble det funnet at å predikere endringer ved hjelp av såkalte deltavariabler (differansen mellom laservariabler fra tidspunkt 1 og tidspunkt 2) som forklaringsvariabler i regresjonsmodellen gav de beste resultatene. Studien testet også en indirekte metode for å predikere endring i biomasse. Her ble biomassen modellert og predikert på begge tidspunkt, og deretter ble endringen predikert som differansen mellom de to biomasseprediksjonene. Begge metodene ble også undersøkt og sammenlignet av McRoberts et al. (2015), men denne studien kom til den motsatte konklusjonen, nemlig at den indirekte metoden så ut til å gi bedre resultater. McRoberts undersøkte også om den direkte metoden kunne forbedres ved å bruke laservariabler i par (i stedet for deltavariabler), fra både tidspunkt 1 og tidspunkt 2, som forklaringsvariabler. Et argument for å modellere med parvise variable, var at begge variablene inneholdt informasjon om utgangssituasjonen ved tidspunkt 1, og sluttsituasjonen ved tidspunkt 2. Resultatene fra denne undersøkelsen viste at metoden som benyttet laservariabler i par var best.

I en studie fra Våler i Østfold, viste Næsset et al. (2013) også at det var mulig å predikere endringer ved hjelp av FLS- data. I denne studien klassifiserte Næsset et al. (2013) endringer ved hjelp logistiske regresjonsmodeller, slik at endring i biomasse kunne knyttes til en «årsak». Endringene ble klassifisert i kategoriene *uberørt*, *degradert* og *avskoget*, og sammenlignet med observerte endringer fra feltmålinger. Undersøkelsen viste at modellene basert på variabler fra FLS- data klassifiserte endringer nokså presist, spesielt områder som var avskoget. En utfordring var derimot å skille uberørte områder fra degraderte områder.

Studien til Næsset et al. (2013) var viktig, fordi det generelt er et behov for å klassifisere endringer. En god klassifikator er viktig blant annet med tanke på REDD- rammeverket, og overvåking av CO<sub>2</sub>-fangst i skog. Det er også interessant å vurdere om en slik klassifikator kan benyttes i forbindelse med boniteringsmodellering (e.g. Noordermeer et al. vurderes for publisering). I 2016 ble det gjort en skogplantakst over et område i Krødsherad i Buskerud fylke, som sist gang ble taksert i 2001. Det ble benyttet data fra flybåren laserskanning ved begge tidspunkt. Det er hittil gjort få klassifiseringsstudier lik Næsset et al. (2013) sin studie. Det var derfor interessant å undersøke om deres metode kunne egne seg for et annet studieområde, og om det kunne være eventuelle forskjeller mellom studiene. Til forskjell fra studien til Næsset et al. (2013), var det fra taksten i Krødsherad også tilgang på et valideringsdatasett til å vurdere modellene på. Dermed kunne man i større grad kontrollere om modellene var aktuelle for å klassifisere endringer for andre områder.

#### 1.4 Problemstilling

Hensikten med denne oppgaven var å utvikle en modell som kunne klassifisere endringer i skog ved hjelp av data fra flybåren laserskanning. Disse tre endringsklassene var *A; uberørt*, *B; tynnet* og *C; avvirket*. Det var ønskelig å klassifisere i disse tre klassene, siden disse er sammenlignbare med endringer som man ønsker å ha oversikt over i tropisk regnskog, med tanke det internasjonale rammeverket REDD. Avvirkede og tynnede områder vil da være sammenlignbare med avskogede og degraderte områder.

For klassifisering av endringstype, ble det tilpasset logistiske modeller med to typer forklaringsvariabler ekstrahert fra laserdata – deltavariabler og parvise variabler. Modellene som ble beregnet med de to metodene ble vurdert etter hvor nøyaktig de predikerte klasseendringer, sammenlignet med observerte klasseendringer på korresponderende prøveflater. Modellene ble validert på sitt eget datasett ved kryssvalidering, og et uavhengig datasett fra det samme undersøkelsesområdet.



## 2. Materiale og metode

### 2.1 Materiale

#### 2.1.1 Studieområdet

Området der undersøkelsen ble gjort, var i Krødsherad kommune i Buskerud fylke. Området var på ca. 50.000 dekar. Krødsherad er en skogbruks- kommune, og representerer godt områder i Norge og Skandinavia der det drives aktivt skogbruk. Produktivt skogareal er 234 000 dekar, som utgjør ca. 62% av kommunens totalareal (Dæhlen, 2017), og årlig avvirkes det mellom 45 000 og 55 000 m<sup>3</sup> tømmer. Noresund er administrasjonssenteret i kommunen, (UTM 32, 534765 Ø, 6671639 N). Klimaet er kontinentalt, med stabilt kalde vintre og varme somrer med moderate nedbørsmengder, og høye temperaturforskjeller gjennom døgnet. Geologisk består grunnfjellet i kommunen av bergarter som gneis og gabbro, og kvartsitt. I lavereliggende strøk og langs Krøderen er det overliggende sedimenter av morenemasser og marine avsetninger (Thorsnæs, 2017). Studieområdet lå først og fremst i tilknytning til Krøderen, med utlagte prøveflater og kontrollflater på begge sider av innsjøen.



Figur 1:

Venstre: Krødsherad kommune er avmerket med et kvadrat (Norgeskart, u.å).

Høyre: Oversiktskart over Buskerud fylke, og Krødsherad kommune (mørkt skravert område)(KF/StoreNorskeleksikon, u.å.)

### 2.1.2 Datagrunnlag

Datainnsamlingen ble gjort i forbindelse med skogplantakst i Krødsherad, i regi av Viken Skog SA. I tillegg ble dataene samlet inn for å brukes til forskningsprosjekter ved NMBU. Det ble gjort feltmålinger og flybåren laserskanning i 2001 (T1) og 2016 (T2). Noen feltmålinger ble også gjort i 2017. Næsset (2004) brukte data fra T1 i en studie der det ble undersøkt om data fra FLS egnet seg til å brukes i en skogplantakst.

Det ble lagt ut et systematisk nett av totalt 160 prøveflater som ble oppsøkt i felt. I tillegg ble det målt 80 kontrollflater, i den hensikt å ha et uavhengig datasett å validere modellene mot.

Kontrollflatene ble lagt ut subjektivt i T1, for å representere områdets variasjon av hogstklasser, bonitetsverdier og treslag.

Feltemålinger 2001

#### Prøveflater

Ved T1 ble prøveflatene navigert til med håndholdt Garmin GPS. For å få et geografisk samsvar mellom laserdata og prøveflatens posisjon, ble den endelige posisjonen til hver prøveflate bestemt ved bruk av differensiell GPS og GLONASS. Prøveflatenes sentrum ble derfor logget med en Javad Legacy GPS mottaker, i ca. 30 minutter. Samtidig ble en kjent koordinat i nærheten, et fastmerke, logget med samme type GPS, som fungerte som såkalt basestasjon. Koordinaten til prøveflatens sentrum ble senere korrigert for feilen funnet ved basestasjonen. Basestasjon lå i sentrum til studieområdet, maks 10 km fra prøveflatene. Prosedyren med å logge prøveflatenes sentrum, og korrigere for feil ved hjelp av basestasjon, ble også gjort for alle hjørnepunktene til kontrollflatene.

Fra feltemålingene var man i denne klassifiseringsstudien interessert i informasjon om hvilken hogstklasse prøveflatene tilhørte; hogstklasse II (HKL II) eller hogstklasse III-V (HKL III-V) Dette fordi man ønsket å vite om prøveflatene hadde endret seg fra en HKL til en annen mellom T1 og T2. Fra prøveflatene i HKL II var det derfor kun denne informasjonen, altså at prøveflaten tilhørte HKL II, som ble benyttet. Prøveflatene i HKL II ble oppsøkt i felt, og fire sirkulære småflater på 40m<sup>2</sup> (radius på 3.57 m) ble etablert 5.04 m ut fra prøveflatens sentrum, en i hver himmelretning. På disse fire småflatene, som hver ble delt opp i fire kvadranter, ble det gjort registreringer av blant treantall og plantenes gjennomsnittshøyde, men slik informasjonen var ikke relevant for denne studien.

I HKL III- V ble det gjort registreringer med tanke på nøyaktig estimering av middelhøyde, overhøyde, middeldiameter, grunnflate, treantall og volum. Informasjon som var relevant for denne studien var at prøveflaten tilhørte nettopp HKL III-V. I tillegg ble prøveflatenes treantall ved T1 og T2 brukt som informasjon for å identifisere eventuelle endringer. Prøveflatene var sirkulære og hadde et areal på 233 m<sup>2</sup> (radius på 8.61 meter). Hvert tre i prøveflatene ble klavet i brysthøyde og notert i stikkliste

med 2 cm diameterklasser. I hogstklasse 3 ble trær under 4 cm i brysthøydiameter utelatt, og i hogstklasse 4 og 5 ble trær under 10 cm i brysthøydiameter utelatt. For hver prøveflate ble det ved relaskopmåling i flatens sentrum plukket ut ca. 10 prøvetrær (dette antallet varierte fra flate til flate). Prøvetrærnes brysthøydiameter ble målt på millimeternivå, mens høyde ble målt på desimeternivå, med Vertex høydemåler.

Trær som stod på prøveflatens grense ble ikke målt dersom over halvparten av treets diameter ved brysthøyde var utenfor prøveflaten. Døde trær på prøveflaten ble heller ikke målt. For hver prøveflate i HKL III-V, og hver kvadrant i HKL II ble det registrert hovedtreslag (det dominerende treslaget på prøveflaten) og nøyaktig hogstklasse.

### Kontrollflater

I tillegg til feltmålingen av prøveflatene, ble det også gjort feltmåling av totalt 80 kontrollflater i 2001. Hensikten med å taksere kontrollflatene var for å få et uavhengig datagrunnlag, som modellene som var utviklet på data fra prøveflatene kunne valideres mot. Kontrollflatene var kvadratiske, med areal på ca. 3700 m<sup>2</sup> (61× 61m). Det var en utfordring å legge ut så store kvadrater i til dels krevende terreng og tett skog, ved hjelp av målebånd og kompass. Derfor varierte størrelsen og formen på kvadratene mye. For HKL II ble kontrollflaten systematisk inndelt i 64 småflater, der hver småflate hadde et areal på 40 m<sup>2</sup>. Her ble registreringene gjort tilsvarende som på prøveflatene. For kontrollflater i HKL III- V ble registreringene gjort på samme måte som på prøveflatene, men prøvetrærne ble plukket ut annerledes. Hvert  $n$ -te tre som ble klavet var et prøvetre, ut fra en forhåndsbestemt nummerering innad i diameterklassene. For hver kontrollflate ble  $n$  bestemt ut fra relaskopmålinger på kontrollflaten.

### Feltemålinger 2016/ 2017

Prøveflatene og kontrollflatene ble målt om igjen ved T2. Prøveflatene og hjørnepunktene fra T1 ble gjenfunnet og etablert på nytt sommeren 2016 og 2017 ved bruk av CPOS, kartverkets posisjonstjeneste (Kartverket, 2017). Det ble benyttet en TopCon Hiper Sr mottaker som mottok korreksjonsdata fra kartverkets basestasjon i Noresund, ved hjelp av kontakt med kartverket over GSM- nettet. Grunnet dårlig satellittdekning eller dårlig GSM- forbindelse, var det ved noen tilfeller ikke mulig å være sikker på at det nyetablerte punktet lå på den samme koordinaten som i T1. Derfor ble disse punktene logget på nytt med TopCon Legacy E+ mottaker. Punktene nye posisjon ble senere bestemt ved at de ble korrigeret for feil registrert av kartverkets basestasjon i Noresund.

På prøveflatene ble feltemålingene gjort på samme måte som ved T1, for både HKL II og HKL III-V. For kontrollflater i HKL II ble målingene gjort på samme måte som ved T1. For HKL III-V ble kontrollflatene delt inn i 16 celler, der hver celle var på 233 m<sup>2</sup> (15.25 × 15.25 m). Hensikten med

denne oppdelingen var å ha celler fra kontrollflatene som var like store som prøveflatene. Feltnålingene på hver celle ble utført på samme måte som for hele kontrollflaten i 2001. For hver rute ble det bestemt hovedtreslag og hogstklasse.

### 2.1.3 Laserdata

Laserdataene ble samlet inn i 2001 (T1) og 2016 (T2) for å korrespondere i tid med feltdatainnsamling. Ved T1 foregikk lasertaksten fra 23. juli til 1. august. Ved T2 ble taksten utført mellom 7. juni til 31 juli (Tabell 1), av TerraTec AS. Flyet holdt en forhåndsbestemt rute og fulgte flystriper som var ca. 1000 m brede. En overlapp på flystriper på mellom 50% og 20% (ved hhv T1 og T2) medførte en effektiv flystripebredde på 500m og 800m. Flyet holdt en hastighet på ca. 70 m/s, og en høyde på ca. 1000 m over bakken. Ved T1 ble det skannet med en pulsrepetisjonsrate på 10 kHz, mens ved T2 var pulsrepetisjonsraten 534 kHz. En maksimal skanningsvinkel på 15° ved T1 og 20° ved T2 gav en gjennomsnittlig punkttetthet på bakken på 1 laserpuls pr m<sup>2</sup> ved T1, og 11 laserpuls pr m<sup>2</sup> ved T2.

Tabell 1: Oversikt over sentrale fly- parametere fra T1 (2001) og T2 (2016).

Parametere	T1	T2
Lasersensor	<i>ALTM 1210 (Optech)</i>	<i>LMS Q – 1560 (Riegl)</i>
Dato for innsamling	23/6 – 1/8	7/6 – 31/7
Gjennomsnittlig flyhøyde (m)	650	1280
Gjennomsnittlig flyhastighet (m/s)	75	69
Overlapp mellom striper (%)	50	20
Pulsrepetisjonsrate (kHz)	10	534
Skanningsfrekvens (Hz)	30	115
Skanningsvinkel (+/-°)	15	20
Pulstetthet på bakken (m <sup>-2</sup> )	1	11

Fra laserdataene ble laserekkoene klassifisert i to grupper; bakkeekko og vegetasjonsekko.

Bakkeekko var laserekko som traff bakken, og ble brukt til å modellere terrenget. Vegetasjonsekko var laserekko som traff vegetasjonen. For hvert vegetasjonsekko kunne dermed høyden fra bakken beregnes, ved å trekke ifra differansen fra terrengmodellen. Prinsippskisse for FLS er lagt ved som Vedlegg 1. Til videre beregninger ble det kun benyttet første ekko, som generelt er mindre sensitiv overfor variasjoner i trekronen enn det siste ekko er (Næsset, 2005a). Data fra første ekko er også mer robust overfor variasjoner i flyets (og LiDAR- sensorens) høyde over bakken (Næsset, 2004b).

Siden Krødsherad er et område med topografiske variasjoner, var det sannsynlig at flyets høyde over bakken varierte.

For hver prøveflate og kontrollflate ble det beregnet såkalte høydevariabler, som skulle representere høydefordelingen til ekkosignalene. Avstanden mellom en grense på 2 meter over bakken og opp til øverste vegetasjonsekko (Hmax) ble delt inn i ti persentiler, H10, H20,... , H90, Hmax. Disse persentilene representerte høyden til andelen av ekkosignaler returnert fra nedenfor denne høyden. For eksempel var H20-variabelen høyden over bakken, der 20% av ekkosignalene hadde truffet nedenfor denne høyden. Det ble også beregnet tetthetsvariabler, som forklarte hvordan den horisontale fordelingen av vegetasjonen var. Høydesjiktet mellom 2 meter over bakken og opp til høyden H95- høydevariabel ble delt inn i 10 like store deler, som ble tilegnet hver sin tetthetsvariabel, fra D0 til D9. Verdien til for eksempel D2- tetthetsvariabel tilsvarte andelen av antall ekkosignaler som var returnert fra høyere enn 2/10 av høyden til H95 (e.g. Bollandsås et al., 2008)

## 2.2 Metode

### 2.2.1 Behandling av datamaterialet

Av de 160 prøveflatene ble fire fjernet fra datasettet fordi de var delt i to på grunn av forskjellige hogstklasser ved T2. I tillegg manglet det laserdata for fem prøveflater målt i T2. En prøveflate manglet laserdata fra T1. I tillegg ble en prøveflate kun feltmålt i T1, og ikke i T2. Det var også en prøveflate som hadde blitt registrert med feil posisjon. Dermed var det totalt 12 prøveflater som ble fjernet fra datasettet, slik at datasettet som ble brukt til videre beregninger bestod av 148 prøveflater.

Av de 80 kontrollflatene ble 14 flater fjernet fordi de bestod av både HKL II og HKL III- V ved T2. I tillegg manglet det laserdata fra seks kontrollflater. Dermed ble det fjernet totalt 20 kontrollflater, slik at datasettet som ble brukt i valideringen bestod av 60 kontrollflater.

### 2.2.2 Klassifisering av feltmålte prøveflater

Under felttaksten ble prøveflatene registrert som enten HKL II eller HKL III-V. 13% av prøveflatene tilhørte HKL II, mens 87% av prøveflatene tilhørte HKL III-V, ved T2. Ved T2 ble det ikke registrert om det var skjedd noen endringer på prøveflatene; om det var gjort *tykning* eller *avvirkning*, eller om prøveflatene hadde stått *uberørt* siden T1. Endringer på prøveflatene ble derfor klassifisert etter visse kriterier i etterkant av taksten. Dersom en prøveflate tilhørte HKL II ved T1 og hadde utviklet seg til HKL II eller HKL III-V ved T2, ble den klassifisert som uberørt (A) (Tabell 2). Dersom en prøveflate tilhørte HKL III-V ved T1, og ble registrert som HKL II ved T2, ble den klassifisert som avvirket (C). Dersom en prøveflate tilhørte HKL III-V ved T1 og T2 ble den klassifisert som tynnet (B) dersom prøveflaten hadde hatt en tilbakegang i treantallet som oversteg det man kan forvente for naturlig avgang. For gran (*Picea abies*) og furu (*Pinus sylvestris*) tilsvarer det en avgang på henholdsvis 0.72% og 0.54% pr. år (Eid & Tuhus, 2000). Dette er gjennomsnittlig mortalitet basert på landsskogstakseringens materiale. Dersom avgangen var lik eller mindre enn dette, ble prøveflaten klassifisert som uberørt.

Tabell 2: klassifisering av endringer ut ifra feltmålingene

HKL T1	HKL T2	Avgang i treantall	Observert endringsklasse
II	II		A
II	III – V		A
III – V	III – V	Mindre eller lik grensen for naturlig avgang	A
III – V	III – V	Mer enn grensen for naturlig avgang	B
III – V	II		C

HKL: hogstklasse, inndelt i to strata; hogstklasse II og hogstklasse III-V. T1: tidspunkt 1 (2001) T2: tidspunkt 2 (2016). A: uberørt. B: tynnet. C: avvirket

I Tabell 3 presenteres minimums-, maksimums- og gjennomsnittsverdier for stående kubikkmasse for endringsklassene A, B og C, basert på feltmålingene ved T1 og T2, på prøveflatene og kontrollflatene. De tilsvarende verdiene for treantall er også oppgitt. I Tabell 3 er det også oppgitt treslagsfordeling ved T2, basert på data fra prøveflatene.

Tabell 3: Oversiktsdata for endringsklasse A, B og C basert på feltdata fra prøveflatene (størrelse ca. 233 m<sup>2</sup>) og kontrollflatene (størrelse ca. 3700 m<sup>2</sup>).

	Prøveflater		Kontrollflater	
	min- maks	Gjennomsnitt	min- maks	Gjennomsnitt
<i>A. Uberørte flater</i>				
Antall flater		101		42
m <sup>3</sup> / daa T <sub>1</sub>	2 – 68.3	18.3	4 – 28.5	16.3
m <sup>3</sup> / daa T <sub>2</sub>	1.8 – 81.8	24.5	6.8 – 43.2	22.4
Endring m <sup>3</sup> /ha		6.2		6.1
N/ daa T <sub>1</sub>	17 - 137	69	24 - 148	74
N/ daa T <sub>2</sub>	17 - 206	9	27 - 163	96
Endring N/ daa		25		22
<i>B. Tynnede flater</i>				
Antall flater		28		14
m <sup>3</sup> / daa T <sub>1</sub>	6.5 – 64.4	24.1	10.5 – 36.0	25.5
m <sup>3</sup> / daa T <sub>2</sub>	3.0 – 90.9	27.2	8.1 – 53.8	29.9
Endring m <sup>3</sup> / daa		3.1		4.4
N/ daa T <sub>1</sub>	22 - 193	82	48 - 144	84
N/ daa T <sub>2</sub>	13 - 142	63	20 - 122	67
Endring N/ ha		-19		-19
<i>C. Avvirkede flater</i>				
Antall flater		19		4
m <sup>3</sup> / daa T <sub>1</sub>	8.1 – 47.6	30.6	21.1 – 38.4	28.4
m <sup>3</sup> / daa T <sub>2</sub> *	x	x		
N/ daa T <sub>1</sub>	34 - 180	93	87 - 130	106
N/ ha T <sub>2</sub>	x	x		
<i>Treslagsfordeling</i>				
Gran T <sub>2</sub>		50 %		
Furu T <sub>2</sub>		38 %		
Lauv T <sub>2</sub>		12 %		

m<sup>3</sup>/daa: stående tømmervolum (m<sup>3</sup>) pr. dekar. N/daa: treantall pr dekar. T1: tidspunkt 1 (2001). T2: tidspunkt 2 (2016)

\*for klasse C ble det ikke registrert m<sup>3</sup> ved T2, siden alle prøveflater og kontrollflater på det tidspunktet tilhørte HKL II. Det ble ikke registrert m<sup>3</sup> for prøveflater i HKL II.

### 2.2.3 Tilpassing av logistiske regresjonsmodeller til å klassifisere endringer

For å tilpasse modeller til prediksjon av endringer ved hjelp av variabler ekstrahert fra data fra flybåren laserskanning ble det brukt logistisk regresjon. Den observerte endringsklassen ble brukt som responsvariabel i modellene, høyde- og tetthetsvariablene fra laserdataene ble brukt som forklaringsvariabler. Til å beregne modellene ble dataprogrammet Rstudio (Rstudio inc., 2017) benyttet. Funksjonen «*multinom*» ble brukt for å lage de statistiske modellene. «*Multinom*» kommer av «multinomial logistic regression», som betyr at modellenes responsvariabel bestod av flere enn to kategorier (Kwak & Clayton- Matthews, 2002), nemlig A, B og C. Funksjonen «*predict*» ble brukt for å predikere nye klasser på prøveflatene. Den generelle formelen for logistisk regresjon var (Mendenhall & Sincich, 2011):

$$y = \frac{e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_i x_i}}{1 + e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_i x_i}} \quad (1)$$

*y*: observert responsvariabel; A, B eller C.

$\beta_i x_i$ : forklaringsvariabel fra flybåren laserdata

Et aspekt som ble tatt hensyn til i modelleringene, var multikollinearitet. Dette er et problem som oppstår dersom to eller flere forklaringsvariabler i samme modell korrelerer med hverandre. Variablene inneholder da samme type informasjon, noe som kan føre til at modellene klassifiserer upresist. For å unngå multikollinearitet ble det ikke benyttet forklaringsvariabler som hadde en VIF-verdi (variansinflasjonsfaktor) høyere enn 10 i samme modell (Mendenhall & Sincich, 2011).

$$VIF_i = \frac{1}{1 - R_i^2} \quad (2)$$

$R_i$ : korrelasjonskoeffisienten mellom to variabler



For at prosessen med å modellere ikke skulle bli for omfattende og tidkrevende ble det gjort et utvalg av de ti høydevariablene ekstrahert fra laserdataene. Det ble valgt ut tre høydevariabler, en fra nedre sjikt, en fra midtre og en fra øvre, for å bevare informasjon fra hele høydesjiktet. Sannsynligheten for at man forkastet en høydevariabel som potensielt modellerte bra var dermed liten. Dette ble også gjort for de ti tetthetsvariablene; tre tetthetsvariabler fra ulike høydesjikt ble plukket ut. Dermed var det høydevariablene H20, H40 og H80 og tetthetsvariablene D1, D3 og D7 som ble brukt i modelleringene. I Tabell 4 presenteres alle modellformene som ble testet, og hvilke forklaringsvariabler som inngikk i hver modellform.

*Tabell 4: Ulike alternative modeller med de respektive selekterte forklaringsvariabler.*

<b>Modell nr.</b>	<b>Forklaringsvariabler</b>
1	H20 + D1
2	H20 + D3
3	H20 + D7
4	H20 + D1 + D7
5	H20 + D3 + D7
6	H20 + H80 + D1
7	H20 + H80 + D3
8	H20 + H80 + D7
9	H40 + D1
10	H40 + D3
11	H40 + D7
12	H40 + D1 + D7
13	H40 + D3 + D7
14	H80 + D1
15	H80 + D3
16	H80 + D7
17	H80 + D1 + D7
18	H80 + D3 + D7

*H20, H40, H80: Høydevariabler. D1, D3, D7: Tetthetsvariabler*

#### 2.2.4 Kryssvalidering og validering av modellene

Et viktig moment for utvelgelse av de beste modellene var at resultatene (de predikerte klassene) ikke skulle være avhengig av sitt eget datasett, som baserte seg på de 148 prøveflatene. Det var av interesse å forsikre seg om at modellene var uavhengige. Med uavhengighet menes det at datasettet som modellene var utviklet på grunnlag av, ikke hadde en betydelig innvirkning på modellenes resultater. Det var ønskelig å vite at de samme modellene også kunne benyttes for å klassifisere endringer for andre typer datasett, i andre skogområder. Modelleringene ble derfor gjort med kryssvalidering (Mendenhall & Sincich, 2011). En og en prøveflate ble utelatt, og de resterende 147 prøveflatene ble brukt til å predikere klassen for den utelatte prøveflaten. Dette ble gjentatt for alle de 148 prøveflatene.

I tillegg til at modellene ble kryssvalidert, ble de også validert på kontrollflatene. Dette ble gjort for å undersøke om modellene hadde generell gyldighet. Validering av modeller mot et annet datasett, slik som kontrollflatene, var en av de mest effektive måtene å kontrollere at modellene kunne brukes for andre uavhengige datasett (Mendenhall & Sincich, 2011). Modellene ble først laget på bakgrunn av de 148 prøveflatene, før de ble validert på kontrollflatene. Ved T2 ble kontrollflatene delt inn i 16 celler. Her ble da hver celle i kontrollflatene predikert i en endringsklasse, og den klassen som oftest ble predikert på de 16 cellene ble den predikerte klassen for hele kontrollflaten.

#### 2.2.5 Kriterier for rangering av modellene

Modellene ble rangert etter hvor presist de klassifiserte prøveflatene. For hver modell ble det beregnet *overall accuracy* ( $p_o$ ) (Næsset et al., 2013). Dette er en prosent over hvor mange prøveflater modellen klassifiserte korrekt, i forhold til totalt antall prøveflater.

$$p_o = \frac{n}{N} \tag{3}$$

$p_o$ : *overall accuracy*

$n$ : *totalt antall prøveflater som ble predikert korrekt.*

$N$ : *totalt antall prøveflater.*

Modellene ble også vurdert etter *Cohens Kappa Index (k)* (Cohen, 1960), som er et mål på hvor god en modell er, sammenlignet med om klassifiseringen var gjort helt tilfeldig.

$$k = \frac{p_o - p_e}{1 - p_e} \quad (4)$$

*K: kappa*

*p<sub>o</sub>: overall accuracy*

*p<sub>e</sub>: forventet sannsynlighet for at observerte prøveflater ble klassifisert som korrekt klasse A, B eller C, ut ifra en tilfeldig prediksjon.*

Sammenlignet med *p<sub>o</sub>* var *kappa* en verdi som tok hensyn til skjevfordeling i datasettet. Det var en overvekt av prøveflater som var uberørte, mens avvikede prøveflater var i mindretall. *p<sub>o</sub>* kunne bli høy dersom mange av prøveflatene i uberørt- klassen ble predikert korrekt, selv om modellen var unøyaktig for de to andre klassene, som bestod av færre prøveflater. Dette elementet tok *kappa* hensyn til, slik at i denne oppgaven, med dette skjevfordelte datasettet, var det mest hensiktsmessig å først vurdere etter *kappa*. *Kappa* gir en verdi mellom 0 og 1, der *kappa* = 0 indikerer at modellen ville klassifisert like presist som en ren tilfeldig klassifisering. Dersom *kappa* = 1, vil modellen klassifisere alle prøveflater i korrekt endringsklasse.

Under modelleringen ble det tydelig at klasse B var vanskeligst å predikere, denne klassen ble ofte forvekslet med klasse A. Derfor ble modellene også vurdert etter *producers accuracy* for klasse B (*p<sub>pb</sub>*), et mål på modellenes evne til å predikere tynnede prøveflater korrekt (Næsset et al., 2013)

$$p_{pb} = \frac{n_b}{N_{ob}} \quad (5)$$

*p<sub>pb</sub>: producers accuracy for klasse B*

*n<sub>b</sub>: antall korrekt predikerte tynnede prøveflater.*

*N<sub>ob</sub>: totalt antall observerte tynnede prøveflater.*

Dersom andelen *p<sub>ob</sub>* var høy, var dette en indikator på at en stor andel av de observerte prøveflatene i klasse B faktisk ble predikert som klasse B. Det ble også beregnet *users accuracy* for klasse B (*p<sub>ub</sub>*), en prosentverdi for hvor mange av de predikerte tynnede prøveflatene som faktisk var tynnet (Næsset et al., 2013). En høy *p<sub>ub</sub>* var en indikasjon på at modellene predikerte tynnede prøveflater presist, i forhold til hvor mange prøveflater som ble predikert som tynnet.

$$p_{ub} = \frac{n_b}{N_{pb}} \quad (6)$$

$p_{ub}$ : users accuracy for klasse B

$n_b$ : antall korrekt predikerte tynnede prøveflater

$N_{pb}$ : totalt antall predikerte tynnede prøveflater

I tillegg til de overfor nevnte kriteriene ble det også laget *confusion- matrix* (Tabell 8) for noen utvalgte modeller (Næsset et al., 2013). En *confusion- matrix* var en tabell over antall observerte prøveflater i hver klasse, og antall predikerte prøveflater i hver klasse. Ut fra en slik tabell kunne man beregne  $p_{pb}$  og  $p_{ub}$  for de ulike klassene, samt  $p_o$ .

#### 2.2.6 Klassifisering av endringer ved bruk av deltavariabler og parede variabler

De fem kriteriene  $p_o$  (3),  $kappa$  (4),  $p_{pb}$  (5),  $p_{ub}$  (6) og *confusion- matrix* (Tabell 10) ble benyttet for å vurdere hvilke forklaringsvariabler som predikerte klasseendringer mest presist. På bakgrunn av de samme kriteriene ble det også undersøkt to metoder for å predikere klasseendringer; modeller basert på delta- variabler (Metode 1), og modeller basert på parede variabler (Metode 2) fra laserdataene. I Metode 1 ble det brukt differanseverdier mellom forklaringsvariabler fra T2 og fra T1. I Metode 2 ble det benyttet forklaringsvariabler fra både T1 og T2 i modellene. Med utgangspunkt i likning (1) ble det brukt følgende forklaringsvariabler for Metode 1:

$$X_i = \Delta x_i = x_{iT2} - x_{iT1}$$

$\Delta x_i$ : forklaringsvariabel som ble brukt i modellene

$x_{iT2}$ : forklaringsvariabelen  $i$  ved tidspunkt T2

$x_{iT1}$ : forklaringsvariabelen  $i$  ved tidspunkt T1

For Metode 2 ble det også tatt utgangspunkt i likning (1). Her ble både  $x_{iT2}$  og  $x_{iT1}$  brukt som forklaringsvariabel i modellen. Dermed bestod modeller basert på metode 2 av dobbelt så mange forklaringsvariabler som modeller basert på metode 1.

### 3. Resultater

I tabell 5 framkommer resultatene for alle modellene, med de ulike kombinasjonene av forklaringsvariabler. Hver variabelkombinasjon ble modellert med Metode 1 og 2, som så både ble kryssvalidert, og validert mot kontrollflatene.

Som Tabell 5 viser, gav Metode 1 en gjennomsnittlig kappa på 0.62 ved en kryssvalidering, og 0.65 ved validering mot kontrollflatene. For Metode 1 var det modell nr. 3 og nr.11 om gav lavest kappa, ved en kryssvalidering, på 0.52 for begge modellene. For Metode 1, kryssvalidering, var det modell nr. 2 og nr. 15 som hadde høyest kappa, på henholdsvis 0.67 og 0.68.

Tabell 5: Kappa og overall accuracy for alle modellene som ble testet. Overall accuracy ( $p_o$ ) er oppgitt i parentes.

Modell nr.	Forklaringsvariabler	Kappa ( $p_o$ , %)			
		Kryssvaliderte modeller		Modeller validert mot kontrollflatene	
		Metode 1	Metode 2	Metode 1	Metode 2
1	H20 + D1	0.62 (83.1)	0.63 (83.8)	0.76 (90.0)	0.69 (86.7)
2	H20 + D3	0.67 (85.4)	0.64 (83.8)	0.76 (90.0)	0.72 (88.3)
3	H20 + D7	0.52 (79.7)	0.53 (79.0)	0.57 (83.3)	0.53 (81.7)
4	H20 + D1 + D7	0.62 (83.1)	0.65 (84.5)	0.72 (88.3)	0.61 (83.3)
5	H20 + D3 + D7	0.64 (83.8)	0.66 (84.5)	0.72 (88.33)	0.59 (83.3)
6	H20 + H80 + D1	0.63 (83.8)	0.64 (83.8)	0.68 (86.7)	0.67 (86.7)
7	H20 + H80 + D3	0.6 (81.8)	0.62 (82.4)	0.65 (85.0)	0.59 (83.3)
8	H20 + H80 + D7	0.55 (80.4)	0.55 (79.7)	0.51 (81.7)	0.57 (83.3)
9	H40 + D1	0.66 (85.1)	0.67 (85.1)	0.64 (85.0)	0.69 (86.7)
10	H40 + D3	0.67 (85.1)	0.64 (83.8)	0.69 (86.7)	0.76 (90.0)
11	H40 + D7	0.52 (79.7)	0.57 (81.0)	0.53 (81.7)	0.54 (81.7)
12	H40 + D1 + D7	0.62 (83.1)	0.63 (83.1)	0.64 (85.0)	0.59 (83.3)
13	H40 + D3 + D7	0.67 (85.1)	0.66 (84.5)	0.65 (85.0)	0.64 (85.0)
14	H80 + D1	0.63 (83.8)	0.62 (83.1)	0.68 (86.7)	0.64 (85.0)
15	H80 + D3	0.68 (85.8)	0.64 (83.8)	0.69 (86.7)	0.67 (86.7)
16	H80 + D7	0.55 (81.1)	0.54 (79.7)	0.54 (81.7)	0.49 (80.0)
17	H80 + D1 + D7	0.63 (83.8)	0.63 (83.1)	0.64 (85.0)	0.72 (88.3)
18	H80 + D3 + D7	0.65 (84.5)	0.61 (82.4)	0.67 (86.7)	0.56 (86.7)
Gjennomsnittet til modellene		0.62 (83.2)	0.62 (82.8)	0.65 (85.8)	0.63 (85)
Standardavviket til modellene		0.05 (1.9)	0.04 (1.77)	0.07 (2.49)	0.07 (2.61)

Med utgangspunkt i Tabell 5 ble det gjort et utvalg av de to modellene som gav høyest kappa, og de to som gav lavest. Dette ble gjort for kryssvaliderte modeller modellert med Metode 1 (Tabell 6). De samme modellene ble valgt ut, som var modellert med Metode 2 (Tabell 7), med den hensikt å sammenligne de to metodene.

Tabell 6: Kryssvaliderte modeller, modellert med Metode 1

Modell nr.	Forklaringsvariabler	kappa	$p_o$	$p_{pb}$	$p_{ub}$
2	H20 + D3	0.67	85.1	46.4	68.4
15	H80 + D3	0.68	85.8	46.4	72.2
3	H20 + D7	0.52	79.7	21.4	46.2
11	H40 + D7	0.52	79.7	21.4	46.2

$p_o$ : overall accuracy.  $p_{pb}$ : producers accuracy for klasse B.  $p_{ub}$ : users accuracy for klasse B.

Tabell 7: Kryssvaliderte modeller, modellert med Metode 2

Modell nr.	Forklaringsvariabler	kappa	$p_o$	$p_{pb}$	$p_{ub}$
2	H20 + D3	0.64	83.8	42.9	63.2
15	H80 + D3	0.64	83.8	42.9	63.2
3	H20 + D7	0.53	79.1	28.6	44.4
11	H40 + D7	0.57	81.0	35.7	52.6

$p_o$ : overall accuracy,  $p_{pb}$ : producers accuracy for klasse B,  $p_{ub}$ : users accuracy for klasse B.

For modell nr. 2 og nr. 15 (de to beste modellene) hadde Metode 1 høyere verdier enn Metode 2, for samtlige kriterier modellene ble vurdert etter. Av de utvalgte modellene hadde den beste modellen, modell nr. 15, en  $p_{pb}$  på 46.4% for Metode 1, og en  $p_{pb}$  på 42,9% for Metode 2 (Tabell 6 og7). Den dårligste modellen, modell nr. 3, hadde en  $p_{pb}$  på 21.4% for Metode 1, og en  $p_{pb}$  på 28.6% for Metode 2.

Det ble laget en confusion- matrix for modell nr. 15 og nr. 11, modellert med Metode 1, kryssvalidert. Disse matrisene fremgår i henholdsvis Tabell 8 og 9.

*Tabell 8: Confusion- matrix av kryssvalidert modell nr. 15, modellert med Metode 1. Antall observerte prøveflater, mot antall predikerte prøveflater i de tre endringsklassene.*

Observert	Predikert				Producers accuracy
	A	B	C	Totalt	
A. Uberørt	97	4	0	101	96.0 %
B. Tynnet	14	13	1	28	46.4 %
C. Avvirket	1	1	17	19	89.5 %
Totalt	112	18	18	148	
Users accuracy	86.6 %	72.2 %	94.4 %		
Overall accuracy				85.8 %	

*Tabell 9: Confusion- matrix av kryssvalidert modell nr. 11, modellert med Metode 1.*

Observert	Predikert				Producers accuracy
	A	B	C	Totalt	
A. Uberørt	96	5	0	101	95.0 %
B. Tynnet	21	6	1	28	21.4 %
C. Avvirket	1	2	16	19	84.2 %
Totalt	118	13	17	148	
Users accuracy	81.4 %	46.2 %	94.1 %		
Overall accuracy				79.7 %	

Det ble også laget en confusion- matrix for modell nr. 15, validert på kontrollflatene, modellert med Metode 1 (Tabell 10). Dersom man sammenlignet Tabell 8 og 10, som bestod av samme forklaringsvariabler i sine modeller, var det en  $p_o$  på henholdsvis 85.8% og 86.7%. Dette tydet på at variabelkombinasjon nr. 15 generelt klassifiserte kontrollflater presist.

*Tabell 10: Confusion- matrix av modell nr. 15, validert mot kontrollflatene, modellert med Metode 1.*

Observert	Predikert				Producers accuracy
	A	B	C	Totalt	
A. Uberørt	41	1	0	42	97.6 %
B. Tynnet	5	7	2	14	50.0 %
C. Avvirket	0	0	4	4	100.0 %
Totalt	46	8	6	60	
Users accuracy	89.1 %	87.5 %	66.7 %		
Overall accuracy				86.7 %	

## 4. Diskusjon

### 4.1 Sammenligning av modellene med hensyn på forklaringsvariablene

Det var, med hensyn til oppgavens problemstilling, ønskelig å finne de forklaringsvariablene, ekstrahert fra laserdataene, som gav modeller som klassifiserte endringer mest presist. Generelt klassifiserte modellene en stor andel av de uberørte områdene korrekt. Tynnede prøveflater ble derimot ofte klassifisert feil, og da som uberørte (Tabell 8 og 9). For avvirkede prøveflater var modellene nokså gode, i Tabell 8 kan man for eksempel se at modell nr. 15 feilpredikerte kun 2 av 19 avvirkede prøveflater.

Dersom man sammenlignet de kryssvaliderte modellene som var utviklet med Metode 1, var det ingen markant store forskjeller mellom modellene (Tabell 5). Det var allikevel en trend at modeller som benyttet tetthetsvariabel D7 hadde en litt lavere kappa, sammenlignet med lignende modeller som benyttet andre tetthetsvariabler. Det så også ut til at modeller med tetthetsvariabel D3 hadde litt høyere kappa- verdier, for eksempel bestod den beste modellen av variablene D3 og H80. For Metode 1 gav denne kryssvaliderte modellen en  $p_o$  på 85.8% og en  $p_{pb}$  på 46.4%. I Tabell 5 var det ingen tilsvarende tendenser dersom man studerte høydevariablene. Det så ut til at modeller som bestod av ulike høydevariabler hadde nokså like kappa- verdier, og at det derfor ikke var noen vesentlige forskjeller mellom høydevariablene.

I undersøkelsen til Næsset et al. (2013) ble det gjort en tilsvarende klassifiseringsstudie. Han benyttet også logistisk regresjon i sine modellberegninger, til å klassifisere endringer ved hjelp av laserdata. Klassene han benyttet var *avskoget*, *degradert* og *uberørt*, og er sammenlignbare med klassene *avvirket*, *tynnet* og *uberørt* i denne masteroppgaven. Han fant at den modellen som gav best resultat benyttet deltavariablene H70 og D0 som forklaringsvariabler. Den modellen gav da en  $p_o$  på 93.8% og en  $p_{pb}$  på 56.3%. Sammenlignet med resultatene i denne masteroppgaven kunne det dermed virke som at tetthetsvariabler fra nederste del av vegetasjonen (D0 og D3) var variabler som klassifiserte de tre endringsklassene mest presist. Dette kan muligens forklares med de prinsipielle forskjellene mellom tetthets- og høydevariabler. Høydevariablene gir et bilde på hvordan de reflekterte laserpulsene fordeler seg i vegetasjonens vertikale sjikt. Tetthetsvariablene beskriver derimot andelen av laserpulser, av alle laserpulsene, som blir reflektert i ulike høydelag i vegetasjonen. Etter en tynning eller slutthogst vil det oppstå flere og større åpninger i skogen. Det vil sannsynligvis føre til at en større andel av laserpulsene reflekteres fra bakken, og ikke fra vegetasjonen. Laserpulser som reflekteres fra vegetasjonen vil da også forventes å ha en jevnere fordeling etter slike inngrep, noe som påvirker tetthetsvariablene i større grad enn høydevariablene. Slike inngrep vil også kunne forventes å påvirke lave tetthetsvariabler mest, fordi en større andel av laserpulsene vil bli reflektert



fra den nederste delen av vegetasjonen. Næsset et al. (2013) presenterte kun én modell for klassifisering av endringer, slik at det ikke var mulig å påvise en trend om at lave tetthetsvariabler generelt gav bedre resultater. Det var uansett en lav tetthetsvariabel, D0, som ble benyttet i Næsset et al. sin beste modell. Årsaker til forskjellen i de to studiene vil bli vurdert senere i diskusjonskapitlet.

#### 4.2 Sammenligning av Metode 1 og Metode 2

I tillegg til å undersøke hvilke forklaringsvariabler som klassifiserte endringer mest presist, ble det også vurdert to ulike metoder å benytte forklaringsvariablene i modellene på. Metode 1 benyttet deltavariabler som forklaringsvariabler i modellene. Metode 2 benyttet parvise variabler fra T1 og T2 i modellene. Fra Tabell 5 kunne man se at gjennomsnittlig *kappa* var den samme for de to metodene. *Kappa*-verdiene til Metode 1 hadde et litt høyere standardavvik mellom sine modeller enn det Metode 2 hadde, en indikasjon på at Metode 1 sine modeller hadde større innbyrdes forskjeller. Dersom man sammenlignet Tabell 6 og 7, kunne man også se at Metode 1 gav de to modellene med høyeste verdier, men også de to modellene med de laveste verdiene. I Tabell 5 kan man se at det var modellering ved Metode 1 som gav modellene med høyest *kappa*. Det må nevnes at utvalget av modellene ble gjort med kryssvalidering med Metode 1, og at disse modellene ble modellert med Metode 2 i etterkant. Dermed var det i utgangspunktet mest sannsynlig at Metode 1 ville ha den beste modellen.

McRoberts et al. (2015) gjorde en tilsvarende sammenligning, hvor de sammenlignet bruk av parede variabler og bruk av deltavariabler i modeller, for å estimere endring i mengde biomasse. De kom fram til at modeller som benyttet variabler i par gav best resultater. Siden det var endring i mengde biomasse McRoberts et al. (2015) estimerte var det vesentlig å ha informasjon om mengden biomasse som var tilstede i utgangspunktet, noe laservariabler fra ett bestemt tidspunkt ville gi en indikasjon på. For en klassifiseringsstudie av endring mellom to tidspunkt, var det derimot viktigst å ha informasjon om hva slags type endring som hadde skjedd. Det kan tenkes at denne informasjonen best kom til uttrykk ved bruk av deltavariabler som forklaringsvariabler i modellene. Dette kan være grunnen til at modellering med Metode 2 ikke gav den beste modellen i denne masteroppgaven.

Et aspekt som også talte for å benytte Metode 1, var at man ikke ønsket multikollinearitet i modellene. Dersom det benyttes mange variabler i en modell, som inneholder mye av den samme informasjonen, øker sjansen for multikollinearitet. Det kan ha vært en viss fare for at variablene som ble benyttet i par korrelerte sterkt med hverandre, siden de stammet fra samme lokalitet (prøveflate), men fra forskjellig tidspunkt. Deltavariablene var variabler som hadde blitt «komprimert» til kun én variabel, med utgangspunkt i data ved T1 og T2. Derfor ville antakelig

modeller utviklet med Metode 1 ha en mindre grad av multikollinearitet, sammenlignet med Metode 2.

Multikollinearitet i modellene kan ha vært et generelt problem i denne undersøkelsen, for både Metode 1 og 2. Det ble gjort en korrelasjonsanalyse av variablene, for å unngå at sterkt korrelerende variabler ikke ble benyttet i de samme modellene. Allikevel viste korrelasjonsanalysen at mange variabler korrelerte med hverandre, selv om de ble vurdert til å kunne inngå i de samme modellene. Mange av variablene hadde en VIF- faktor som var lavere enn 10, som var grensen som ble satt for korrelasjon mellom to variabler. Allikevel hadde mange av disse variablene en VIF- faktor nær 10, noe som indikerte at variablene inneholdt en stor del av den samme informasjonen (se Vedlegg 2, Tabell 11, 12 og 13).

#### 4.3 Validering av modellene på kontrollflatene

I tillegg til en kryssvalidering, ble det også gjort en validering av modellene på kontrollflatene. Dette ble gjort for å undersøke om modellene ville gi et annet og eventuelt dårligere resultat dersom modellene skulle benyttes på andre, mer uavhengige datasett. Med utgangspunkt i Tabell 5 virket det som om kappaverdiene i gjennomsnitt ikke var lavere sammenlignet med de kryssvaliderte modellene, heller tvert imot. De ulike endringsklassene ble klassifisert omtrent like godt ved validering på kontrollflatene (Tabell 10) som ved kryssvalidering (Tabell 8). Dette var en indikator på at denne modellen egnet seg godt til å predikere klasseendringer på andre datasett og andre områder. Fra Tabell 5 så dette også ut til å gjelde for samtlige modeller. Resultatene fra valideringen på kontrollflatene viste at modellene generelt var nokså uavhengige av prøveflatene de var utviklet på grunnlag av. Dermed så det ut til at kryssvalidering var en valideringsmetode som ikke gav ugyldige resultater i denne sammenheng.

#### 4.4 Materiale og metode

Det ble vurdert om noen momenter kunne vært gjort annerledes i denne klassifiseringsstudien. Modellenes største svakhet var at områder som var observert som uberørte eller tynnede, ofte ble forvekslet med hverandre, og ble klassifisert i feil endringsklasse (Tabell 9). Generelt klassifiserte modellene en stor andel av de uberørte områdene korrekt. Tynnede prøveflater ble derimot ofte klassifisert feil, og da som uberørte. Av de 148 prøveflatene var 68% observert som uberørte, mens 19% var tynnede. Dersom datasettet hadde bestått av flere tynnede prøveflater, kunne det muligens bidratt til en høyere presisjon når det kom til å klassifisere tynnede prøveflater korrekt. I prosessen med å legge ut prøveflater ved T1 kunne man for eksempel valgt ut områder der man visste at det kom til å bli tynnet mellom T1 og T2. Eventuelt kunne man ved T2 lagt ut noen ekstra prøveflater som man visste var blitt tynnet mellom T1 og T2. Man kunne også benyttet flybilder for å skaffe flere tynnede prøveflater til datasettet. Ved å studere flybilder fra studieområdet kunne man funnet skogområder som var blitt tynnet i tidsrommet mellom T1 og T2. Fra disse områdene kunne man laget prøveflater der man hadde laserdata, og inkludert disse prøveflatene i datasettet.

For alle modellene som ble testet, med både Metode 1 og 2, var det ofte de samme prøveflatene som ble klassifisert som uberørte, og som var observert som tynnet ut fra felttaksten. En sannsynlig årsak til dette var at prøveflatene ble klassifisert feil ut fra felldataene; de skulle i utgangspunktet ha vært klassifisert som uberørte. Under felttaksten kunne det med fordel ha blitt registrert hvilken endringsklasse hver prøveflate tilhørte; *uberørt*, *tynnet* eller *avvirket*. Fordi dette ikke ble gjort, ble prøveflatene klassifisert ut ifra ulike kriterier basert på skogparametere fra feltmålingene. Blant annet ble det satt en mortalitetsgrense pr. år på 0.72% for gran og 0.54% for furu, som var gjennomsnittstall fra landsskogstakseringens materiale (Eid & Tuhus, 2000). Dersom avgangen i treantall var høyere enn dette, på prøveflater som tilhørte HKL 3-5 ved T1 og T2, ble prøveflatene klassifisert som tynnede, ut fra felttaksten. Denne grensen for naturlig avgang kunne vært satt høyere, fordi det normalt i en tynning blir tatt ut langt flere trær enn det som går bort naturlig. Dermed ville flere prøveflater blitt klassifisert som uberørte, og muligens blitt predikert i korrekt endringsklasse.

I undersøkelsen til Næsset et al. (2013) ble feltmålte endringer klassifisert på en annen måte. De tok ikke utgangspunkt i endring fra HKL II til HKL III-V (eller motsatt vei), men klassifiserte ut fra endring i biomasse. Dersom mengden biomassen ved tidspunkt 2 (2010) var mindre enn 10% av biomassen ved tidspunkt 1 (1999) ble endringen klassifisert som *avskoget*. Dersom biomassen i 2010 utgjorde mer enn 10% av biomassen i 1999, ble endringen klassifisert som *degradert*, avhengig av hvor stor reduksjon det hadde vært i treantallet. Dersom treantallet i 2010 var under 70% av treantallet i 1999, for eldre produksjonsskog, ble endringen klassifisert som *degradert*. Dersom treantallet var høyere

enn dette i 2010, ble endringen klassifisert som *uberørt*. Næsset et al. (2013) fant at sin beste modell predikerte klasseendringer litt mer presist enn det som ble funnet i denne masteroppgaven, med en  $p_o$  på 93.8% og en  $p_{pb}$  på 56,3%. Dette skyldes kanskje at klassifisering av feltmålte endringer ble gjort på en mer presis måte i Næsset sin studie.

Det bør også nevnes at Næsset et al. (2013) sin studie ble gjort i et annet område, og i en annen landsdel med ulik topografi. Basert på tidligere studier er det en generell forståelse om at variasjoner i topografien kan ha en innvirkning på laserdataene (Næsset & Gobakken, 2008; Næsset 2004b). Dersom et område er preget av høye fjell og dype daler, slik som i Krødsherad, er det en utfordring å foreta en flygning med jevn høyde over bakken. Ujevnheter i flyets høyde over bakken kan føre til at punktettheten til laserekko fra bakken og vegetasjonen varierer. Størrelsen til laserpunktene fotavtrykk på bakken og vegetasjon vil sannsynligvis også variere over det laser skannede området. Dette kan være en årsak til at Næsset et al. (2013) fant en mer presis modell, siden topografien varierte mindre i undersøkelsesområdet i Våler.

En annen årsak til forskjellen i resultatene, var at det i studien til Næsset et al. (2013) ble taksert med 11 sesongers mellomrom, mens det i undertegnede studie var 15 år mellom T1 og T2. Dermed var det en større sjanse for at det hadde skjedd færre og mindre omfattende endringer i Næsset et al. (2013) sitt studieområde. For eksempel var 78% av alle prøveflatene i studien deres observert som *uberørte*, 9% som *tynnet* og 13% som *avvirket*. I studien fra Krødsherad var disse tallene henholdsvis 68%, 19% og 13%. Det bør nevnes at denne forskjellen også kan skyldes feilklassifisering i henhold til feltdataene, som nevnt tidligere. Fordi modellene generelt klassifiserte tynnede prøveflater upresist, kan det tenkes at denne forskjellen i andel tynnede prøveflater kan ha ført til en større andel av prøveflatene ble predikert korrekt, i Næsset et al. (2013) sin studie. Det kan også tenkes at en kortere periode mellom T1 og T2 ville forbedre modellenes evne til å skille *uberørte* prøveflater fra *tynnede*. Lang tid mellom T1 og T2 kan føre til at kronen i tynnede bestand får tid til å vokse, og tette igjen åpninger. Dermed kan det i studien fra Krødsherad ha vært mer problematisk å skille tynnede prøveflater fra *uberørte*.

Som man kunne se fra Tabell 8 og 9 var det noen prøveflater som ble klassifisert som *avvirket* ut fra felttaksten, men som allikevel ble klassifisert som *uberørte* eller *tynnede* med modellene. Denne feilen var noe uventet, siden det er en stor forskjell mellom *avvirkede* og *tynnede/uberørte* områder. Denne typen feilklassifisering skyldtes sannsynligvis at det ikke ble registrert trær i HKL III-V på de gjeldende prøveflatene under feltmåling ved T2, samtidig som det kan ha blitt registrert vegetasjon på prøveflaten under laser skanningen ved T2. Dette kan ha vært trær som hadde stammen sin utenfor prøveflaten, men med en vesentlig del av trekronen sin hengende inn i

prøveflaten, såkalte grensetrær. Man kunne vurdert å registrere slike grensetrær under felttaksten. En annen måte å redusere feil knyttet til grensetrær, hadde vært å benytte større prøveflater. Som en generell regel vil «grensetrærnes relative innvirkning reduseres ettersom prøveflatenes størrelse øker» (Næsset et al., 2013). Feilklassifiseringen kan også skyldes at det stod døde trær på prøveflaten, som i henhold til takstinstruksen ikke ble registrert under feltmåling. Disse kan allikevel ha blitt registrert under laserskanningen.

#### 4.5 Klassifisering av endringer ved hjelp av laserdata til bruk i boniteringsmodellering og for overvåking av tropisk regnskog

Det er interessant å vurdere om modellene som ble utviklet i denne studien kan benyttes som klassifikator i boniteringsmodellering, og/eller til å klassifisere endringer i tropisk regnskog i henhold til REDD- rammeverket. For boniteringsprosjektet (e.g. Noordermeer et al., vurderes for publisering) er det avgjørende å ha en klassifikator som presist kan skille ut områder som er ugunstige å bruke til boniteringsmodellering. I Noordermeer et al. (vurderes for publisering) sin studie ble det brukt gjennomsnittlig dominerende høyde, og høydevariabel H90 som klassifikator. Prøveflater som hadde hatt en reduksjon i enten dominerende høyde eller H90 mellom T1 og T2, ble fjernet fra datasettet. For boniteringsmodellering er det områder med signifikante endring i vegetasjonens vertikale sjikt som er mest kritisk å identifisere. Det er nemlig den naturlige endringen i høyden til bestandets dominerende trær som det er viktig å ha kontroll over. Endringer i skogens horisontale sjikt har ingen direkte innvirkning på registrering og måling av bonitet. Man bør allikevel tenke seg om dersom man ønsker å inkludere tynnede prøveflater i boniteringsmodeller, da det kan være en fare for at dominerende trær har blitt tatt ut. Den beste modellen som ble funnet i denne masteroppgaven var lite presis når det kom til å skille uberørte områder fra tynnede, i Tabell 6 kunne man se at  $p_{ub}$  og  $p_{pb}$  var nokså lave, for Modell 15. Dermed anbefales det ikke å benytte denne modellen som klassifikator ved boniteringsmodellering.

Klassifiseringsmetoden ble vurdert også med tanke på å benyttes til overvåking av tropisk regnskog, i henhold til REDD- rammeverket til UNFCCC. Til å klassifisere områder som var avvirket hadde vår beste modell høy presisjon (Tabell 8). Med tanke på REDD- rammeverket kan det dermed være aktuelt å benytte liknende modeller til å klassifisere avskogede og snauhogde områder. For tynnede områder var det derimot knyttet større usikkerhet til denne modellens presisjon. Fra Tabell 8 kom det fram at modellen klassifiserte 46.4% av de tynnede prøveflatene korrekt, med hensyn på  $p_{pb}$ . Dermed var modellen lite egnet til å klassifisere degraderte områder i tropisk regnskog.

## 5. Konklusjon

I denne oppgaven har det blitt undersøkt hvor presist modellering ved bruk av logistisk regresjon, ved hjelp av data fra flybåren laserskanning, kan predikere klasseendringer i et skogområde. Det ble funnet at de beste modellene benyttet tetthetsvariabel D3 som en av forklaringsvariablene. Næsset et al. (2013) kom i sin undersøkelse fram til at D0 inngikk som en av forklaringsvariablene i sin beste modell. Basert på disse to studiene er det dermed en klar indikasjon på at tetthetsvariabler fra vegetasjonens nederste sjikt er viktige for modeller som skal klassifisere denne typen endringer (uberørt, tynnet og avvirket). Sammenlignet man modeller bestående av ulike høydevariabler, var det ingen markante forskjeller mellom disse modellene. Det var heller ingen vesentlig forskjell mellom bruk av deltavariabler og parede variabler i modellene, men det kan tenkes at det for denne studien, der man ønsket å kun klassifisere endringer, var mest aktuelt å benytte deltaverdier. Til å klassifisere områder som var avvirket var modellene generelt gode og hadde høy presisjon. Til å klassifisere tynnede områder var det derimot knyttet større usikkerhet til modellenes presisjon. Modellene var, ut fra resultatene fra validering på kontrollflatene, gyldige til å benyttes på andre uavhengige datasett.

## 6. Litteratur

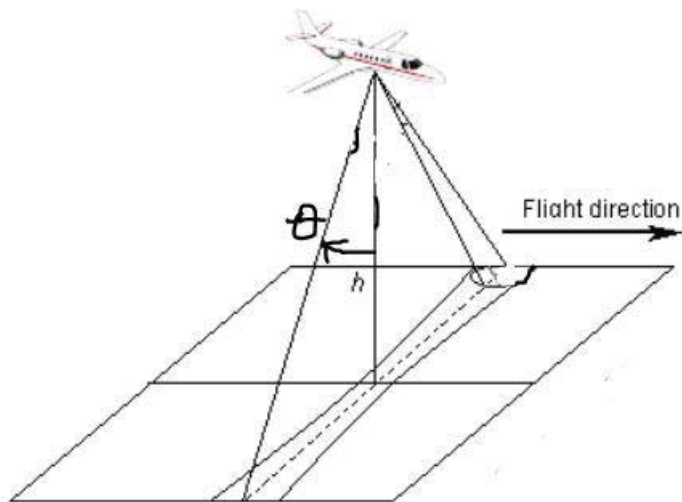
- Barker, T., Allali, A., Bojariu, R., Diaz, S., Elgizouli, I., Griggs, D. & Hawkins, D. (2017). *Climate Change 2007: Synthesis Report*.
- Bollandsås, O. M., Hanssen, K. H., Marthiniussen, S. & Næsset, E. (2008). Measures of spatial forest structure derived from airborne laser data are associated with natural regeneration patterns in an uneven-aged spruce forest.
- Bollandsås, O. M., Gregoire, T. G., Næsset, E. & Øyen, B.-H. (2013). Detection of biomass change in a Norwegian mountain forest area using small footprint airborne laser scanner data. *Statistical Methods & Applications*, 22 (1): 113-129. doi: 10.1007/s10260-012-0220-5.
- Bull, H. (1931). The use of polymorphic curves in determining site quality in young red pine plantations. *Journal of agricultural Research*, 43, 1-28
- Børset, O (1985). *Skogskjøtsel 1: Skogøkologi*. Oslo: Landbruksforlaget.
- Cohen, J. (1960). A coefficient of agreement for nominal scales.
- Dold, C. (2000). *Analyse und Implementierung eines Verfahrens zur Referenzierung geographischer Objekte*. Institut für Photogrammetrie. Tilgjengelig fra: <http://www.ifp.uni-stuttgart.de/lehre/diplomarbeiten/Dold/index.html>.
- Dymond, C. C., Beukema, S., Nitschke, C. R., Coates, K. D. & Scheller, R. M. (2016). Carbon sequestration in managed temperate coniferous forests under climate change. *Biogeosciences*, 13 (6): 1933- 1947.
- Dæhlen, G. (2017). *Skogbruket i Krødsherad*. Tilgjengelig fra: <https://www.krodsherad.kommune.no/tjenester/landbruk-natur-og-miljo/skogbruk-og-utmark/skogbruket-i-krodsherad/> (lest 13.04.2018).
- Eid, T. & Tuhus, E. (2000). Models for individual tree mortality in Norway. Kartverket. (2017). CPOS.
- KF/StoreNorskeleksikon. (u.å.). *Kart Buskerud fylke*.
- Korsvold, G. (2018). *Bonitering: Mjøsen Skog SA* (E-mail 07.05.2018).
- Kwak, C. & Clayton- Matthews, A. (2002). Methods: Multinomial Logistic Regression. *Nursing Research*.
- Maltamo, M., Mustonen, K., Hyyppä, J., Pitkänen, J. & Yu, X. (2004). The accuracy of estimating individual tree variables with airborne laser scanning in a boreal nature reserve.
- McRoberts, R. E., Næsset, E., Gobakken, T. & Bollandsås, O. M. (2015). Indirect and direct estimation of forest biomass change using forest inventory and airborne laser scanning data.
- Mendenhall, W. & Sincich, T. (2011). *A second course in Statistics: Regression analysis. Seventh edition*. USA, Pearson Education, Inc., s.315-318, 364- 365, 494- 502.
- Miljødirektoratet. (2017). *Kyotoprotokollen*. Tilgjengelig fra: <http://www.miljostatus.no/tema/klima/internasjonalt-klimatekonomi/kyotoprotokollen/> (lest 18.03.2018).
- Noordermeer, L., Bollandsås, O. M., Gobakken, T. & Næsset, E. (vurderes for publisering). Direct and indirect site index determination for Norway spruce and Scots pine 2 using bitemporal airborne laser scanner data.
- Norgeskart. (u.å.). Statens kartverk, Geodata AS. Tilgjengelig fra: <https://www.inatur.no/fiske/50f336c8e4b0adea30743bc7> (lest 01.05.2018).
- Næsset, E. (2002). Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data.
- Næsset, E. (2004). Practical large-scale forest stand inventory using a small-footprint airborne scanning laser. *Scandinavian Journal of Forest Research*, 19 (2): 164-179. doi: 10.1080/02827580310019257.

- Næsset, E. (2004b). 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. (2005a). Assessing sensor effects and effects of leaf-off and leaf-on canopy conditions on biophysical stand properties derived from small-footprint airborne laser data. *Remote Sensing of Environment*, 98 (2-3): 356-370.
- 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*.
- Næsset, E., Bollandsås, O. M., Gobakken, T., Gregoire, T. G. & Ståhl, G. (2013). Model-assisted estimation of change in forest biomass over an 11 year period in a sample survey supported by airborne LiDAR: A case study with post-stratification to provide “activity data”. *Remote Sensing of Environment*, 128: 299-314. doi: 10.1016/j.rse.2012.10.008.
- Olerud, K. (2014). *Klimakvoter*. Tilgjengelig fra: <https://snl.no/klimakvoter> (lest 22.03.2018). *RIEGL LMS-Q1560*. Horn, Austria: RIEGL Laser Measurement Systems GmbH.
- Rstudio inc. (2017). *Rstudio desktop* (Versjon 1.1). Programvare. Tilgjengelig fra: <https://www.rstudio.com/products/rstudio/> (lest 01.12.2017).
- TerratecAS. (u.å.). *Flybåren laserscanning*. Tilgjengelig fra: [https://www.terratec.no/flybaaren\\_laserskanning/cms/59](https://www.terratec.no/flybaaren_laserskanning/cms/59) (lest 07.04.2018).
- Tveite, B (1977). *Boniteringskurver for gran*. Ås: Meddelelser fra Norsk Institutt for skogforskning.
- Tveite, B., Braastad ,H. (1981). *Bonitering for gran, furu og bjørk*. *Norsk skogbruk nr.4: 17- 22*
- Thorsnæs, G. (2017). *Krødsherad*. Tilgjengelig fra: <https://snl.no/Kr%C3%B8dsherad> (lest 12.03.2018).
- UNREDD. (2017). *About REDD+*. Tilgjengelig fra: <https://www.unredd.net/about/what-is-redd-plus.html> (lest 18.03.2018).



## 7. Vedlegg

### Vedlegg 1



Figur 2: Prinsippskisse for FLS.  $h$ : høyden over bakken.  $\theta$ : skannevinkel. (Dold, 2000)

### Vedlegg 2

Med hensyn på likning (2) ble det beregnet en korrelasjonskoeffisient,  $r$ , mellom hver forklaringsvariabel. En  $r$  høyere enn 0,949 mellom to variabler, tilsvarte en VIF høyere enn 10. Denne verdien er merket med x i tabell 11, 12 og 13.

Tabell 11: Korrelasjonskoeffisientene til høydevariablene H10,...,Hmax

	H10	H20	H30	H40	H50	H60	H70	H80	H90	Hmax
H10	x	x	0,92	0,9	0,88	0,87	0,86	0,85	0,84	0,79
H20	x	x	x	x	x	x	0,94	0,92	0,91	0,87
H30	0,92	x	x	x	x	x	x	x	x	0,91
H40	0,9	x	x	x	x	x	x	x	x	0,93
H50	0,88	x	x	x	x	x	x	x	x	x
H60	0,87	x	x	x	x	x	x	x	x	x
H70	0,86	0,94	x	x	x	x	x	x	x	x
H80	0,85	0,92	x	x	x	x	x	x	x	x
H90	0,84	0,91	x	x	x	x	x	x	x	x
Hmax	0,79	0,87	0,91	0,93	x	x	x	x	x	x

Tabell 12: Korrelasjonskoeffisientene til tetthetsvariablene D0,...,D9

	<b>D0</b>	<b>D1</b>	<b>D2</b>	<b>D3</b>	<b>D4</b>	<b>D5</b>	<b>D6</b>	<b>D7</b>	<b>D8</b>	<b>D9</b>
<b>D0</b>	x	x	x	x	0,94	0,92	0,9	0,87	0,86	0,84
<b>D1</b>	x	x	x	x	x	0,94	0,92	0,9	0,88	0,86
<b>D2</b>	x	x	x	x	x	x	0,94	0,92	0,9	0,87
<b>D3</b>	x	x	x	x	x	x	x	0,94	0,91	0,88
<b>D4</b>	0,94	x	x	x	x	x	x	x	0,92	0,89
<b>D5</b>	0,92	x	x	x	x	x	x	x	0,94	0,9
<b>D6</b>	0,9	0,92	0,94	x	x	x	x	x	x	0,91
<b>D7</b>	0,87	0,9	0,92	0,94	x	x	x	x	x	0,93
<b>D8</b>	0,86	0,88	0,9	0,91	0,92	0,94	x	x	x	x
<b>D9</b>	0,84	0,86	0,87	0,88	0,89	0,9	0,91	0,93	x	x

Tabell 13: korrelasjonskoeffisientene til høydevariablene H10,..., Hmax, mot tetthetsvariablene D0,..., D9

	<b>H10</b>	<b>H20</b>	<b>H30</b>	<b>H40</b>	<b>H50</b>	<b>H60</b>	<b>H70</b>	<b>H80</b>	<b>H90</b>	<b>Hmax</b>
<b>D0</b>	0,64	0,71	0,73	0,75	0,76	0,77	0,76	0,76	0,76	0,79
<b>D1</b>	0,68	0,74	0,76	0,78	0,79	0,79	0,78	0,79	0,78	0,79
<b>D2</b>	0,71	0,77	0,78	0,79	0,8	0,8	0,8	0,8	0,79	0,79
<b>D3</b>	0,74	0,79	0,8	0,81	0,81	0,81	0,8	0,8	0,79	0,78
<b>D4</b>	0,77	0,81	0,82	0,82	0,82	0,82	0,81	0,8	0,79	0,78
<b>D5</b>	0,81	0,84	0,84	0,84	0,83	0,83	0,82	0,81	0,8	0,78
<b>D6</b>	0,83	0,86	0,86	0,86	0,85	0,85	0,83	0,82	0,81	0,79
<b>D7</b>	0,83	0,87	0,86	0,86	0,85	0,85	0,83	0,82	0,81	0,79
<b>D8</b>	0,79	0,84	0,83	0,83	0,83	0,83	0,82	0,81	0,79	0,77
<b>D9</b>	0,75	0,75	0,8	0,8	0,8	0,8	0,79	0,79	0,77	0,75





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

Postboks 5003  
NO-1432 Ås  
Norway