



Norges miljø- og  
biovitenskapelige  
universitet

**Masteroppgave 2021 30 stp**

Fakultet for Miljøvitenskap og naturforvaltning (MINA)

## **Bruk av hogstmaskindata som referansedata ved ressurskartlegging av salgbart volum og sagtømmervolum**

Use of harvester data as referensedata when mapping totalvolume and sawlogvolume.

Mikal Råheim

Skogfag



## Forord

Denne masteroppgaven markerer slutten på min fem år lange skogfagutdannelse ved Norges Miljø- og Biovitenskaplige Universitet (NMBU). Oppgaven er skrevet ved Fakultetet for miljøvitenskap og naturforvaltning (MINA). Oppgaven er skrevet i samarbeid med det pågående forskningsprosjektet Precision. Hovedmål for prosjektet er: «å redusere tap forårsaket av råte gjennom å bedre informasjonstilgang og skogforvaltning». (råteiskogen.no), mens noen av delmålene er å bedre utnyttelse av data fra hogstmaskiner og forbedre takseringsmetoder, det er i den forbindelse at grunnlaget for denne masteren er gjort mulig og tilgjengelig. Prosjektet ledes av NIBIO men er et samarbeid med NMBU, tømmeromsetningsorganisasjonene, Skogkurs og flere utenlandske forskningsmiljøer (råteiskogen.no). Prosjektet er finansiert av Norges forskningsråd.

Jeg vil benytte anledningen til å takke min hovedveileder Lennart Noordermeer for god veiledning og hjelp gjennom hele prosessen med denne oppgaven. Du har vært tilgjengelig for meg hele veien, noe som har vært sentralt for at jeg har skullet få til dette, og jeg er takknemlig for all lærdom! Jeg vil også takke veileder Terje Gobakken som har vekket min interesse for volumpredikering og hadde ideen som førte til denne oppgaven, og til Ole Martin Bollandsås for hjelp med å fremgangsmåten i denne oppgaven. En stor takk til alle for uvurderlig hjelp og konstruktive tilbakemeldinger.

Takk til søster Ingvild Råheim og bestefar Erik Råheim og foreldre for korrekturlesning gode innspill. Setter stor pris på at dere tok dere tid til å lese oppgaven og komme med tilbakemeldinger, og at dere har hatt tålmodighet med meg mens oppgaveskrivingen har pågått.

Til slutt vil jeg gi en stor takk til alle mine medstudenter på NMBU som har gitt meg fem gode år med et begeistrende skogmiljø!

Norges Miljø og Biovitenskaplige universitet

Fønhus, 30. Mai 2021

Mikal Råheim



## Sammendrag

Oversikt over skogens ressurser og tilstand er en viktig forutsetning for å drive rasjonelt skogbruk. En skogbruksplan er en viktig kilde til slik informasjon, og gir grunnlag for fremtidig skogbruksplanlegging. Siden 2000-tallet har den dominerende og mest kostnadseffektive måten å lage skogbruksplaner på vært ved å bruke Flybåren Laser Skanning (FLS). Data fra FLS vist seg å være svært nyttig innen skoginventering og ressurskartlegging, for eksempel for estimering av volum. En stor del av kostnaden ved FLS er forbundet med manuelle prøveflatetakseringer, som er nødvendige for å ha referansedata til å kalibrere laserdataene. I nyere tid har det blitt sett på muligheten for å erstatte de manuelle feltmålingene, med nøyaktige posisjonerte hogstmaskindata. Det er lite informasjon i dagens skogbruksplaner om kvalitet på tømmervolumet. Selv om informasjon om sagtømmervolum er avgjørende for å ta riktige og lønnsomme beslutninger, finnes det i dag lite av denne informasjonen i en skogbruksplan. Hovedmålet i denne oppgaven var å undersøke om hogstmaskindata fra hogstmaskin i kommersiell drift, kan brukes som referansedata for å predikere totalvolum og sagtømmervolum for geografien maskinen arbeider i.

Random Forest (RF), K nearest neighbour (Knn) og lineær modellering (LM) ble brukt for å modellere og predikere sagtømmer og totalvolum ved bruk av hogstmaskin data og FLS data. Den såkalte Enhanced area based approach (EABA) ble brukt for å redusere kanteffektene. Modelleringsteknikkene ble vurdert etter Root Mean Squared Error Prosent (RMSE%), i tillegg til gjennomsnitt, standardavvik og maks feil i estimert volum. En hogstmaskin som går i kommersiell drift ble brukt for å samle inn referansedata. Maskinen var utstyrt med ekstraustyr som ga nøyaktig posisjonering på hogde trær.

For totalvolum ble verdier av  $RMSE\% = 39\%$  (RF),  $48\%$  (Knn) og  $46\%$  for LM oppnådd, mens for sagtømmervolum ble verdier på  $64\%$  (RF),  $77\%$  (Knn) og  $71\%$  for LM oppnådd. Feil på cellenivå utjevnet hverandre, og med gjennomsnittlige feil på  $-3,5 \text{ m}^3/\text{ha}$  (RF),  $6,11 \text{ m}^3/\text{ha}$  (Knn) og  $0,57 \text{ m}^3/\text{ha}$  (LM) for totalvolum. Gjennomsnittlig feil for sagtømmervolum ble  $-2,24 \text{ m}^3/\text{ha}$  (RF),  $1,68 \text{ m}^3/\text{ha}$  (Knn) og  $-0,08 \text{ m}^3/\text{ha}$  (LM). Gjennomsnittlig feil for både sagtømmervolum og totalvolum anses svært lav.

Denne oppgaven demonstrerer en ny og praktisk tilnærming på bruk av hogstmaskindata innen FLS-basert skoginventering. Lav gjennomsnittlig feil tyder på at nøyaktige bestandsestimater kan oppnås. Dette indikerer at det finnes en mulighet for at hogstmaskindata i fremtiden helt eller delvis erstatte manuelle feltmålinger.



## Abstract

An overview of the forest resources and condition is an important prerequisite for sustainable forest management. A forest management plan is a key source for this type of information, and is the basis for future forest planning. Since the 2000s, the dominant and most cost efficient way of carrying out forest inventories has been the use of Airborne Laser Scanning (ALS). Data from ALS have proven to be very useful within forest inventory and resource planning. Much of the costs tied to the use of ALS is associated with laborious field work, which is required to gather reference data used to calibrate the ALS data. In recent times, the opportunity of replacing the laborious field work with accurate positioned harvester data has shown promising. There is little information in today's forest inventory plans about timber quality such as saw timber. Information about the saw timber volume is crucial to make correct and profitable decisions. However, it exists little of this information in a forest management plan. The main goal in this thesis is to investigate if harvester data from a harvester in commercial operation, can be used as reference data to predict total volume and saw timber volume for the region it is working in.

Random Forest (RF), K-nearest neighbor (Knn) and linear modeling (LM) were used to model and predict saw timber and total volume with the use of harvester data and ALS. The method Enhanced Area Based Approach (EABA) was used to reduce the edge effects. The modelling techniques were assessed according to Root Mean Squared Error Percentage (RMSE%), plus mean, standard deviation and maximum error in estimated volume. A harvester in commercial use, was used to collect reference data. The machine was equipped with accessories that provided the exact (<1m) position of harvested trees.

The RMSE% values for total volume was 39% (RF), 48% (Knn) and 46% (LM) achieved, while saw timber volume got an RMSE% på 64% (RF), 77% (Knn) and 0.57 (LM). Errors at cell level equalized each other and mean errors were -3,5 m<sup>3</sup>/ha for (RF), 6,11 m<sup>3</sup>/ha (Knn) and 0,57 m<sup>3</sup>/ha (LM) for total volume. The average error for sawn timber volume was -2,24 (RF), 1,68 (Knn) and -0,08 with LM. Mean error for both sawn timber and total volume is considered very low.

This thesis demonstrates a new and practical approach to use of harvester data within ALS based forest inventory. Low average errors indicate that accurate stock estimates can be obtained. This may indicate that harvester data can partly or completely replace manual field work.





## Innholdsfortegnelse

|   |           |
|---|-----------|
| Forord.....   | iii       |
| Sammendrag .....  | v         |
| Abstract .....  | vii       |
| <b>1. Innledning .....</b>                                | <b>1</b>  |
| <b>1.1 Skog i Norge.....</b>                              | <b>1</b>  |
| <b>1.2.1 Ressursoversikt.....</b>                         | <b>1</b>  |
| <b>1.2.2 Utvikling av ressurskartlegging i skog .....</b> | <b>2</b>  |
| <b>1.3.1 Hogstmaskindata.....</b>                         | <b>5</b>  |
| <b>1.4 Hogstmaskindata som referansedata.....</b>         | <b>7</b>  |
| <b>1.5 Kanteffekter .....</b>                             | <b>8</b>  |
| <b>1.5.1 Enhanced Area Based Approach .....</b>           | <b>8</b>  |
| <b>1.6 Forskning .....</b>                                | <b>9</b>  |
| <b>1.7 Denne oppgaven.....</b>                            | <b>10</b> |
| <b>1.7.1 Problemstilling.....</b>                         | <b>10</b> |
| <b>2. Material og metode.....</b>                         | <b>11</b> |
| <b>2.1 Studieområde .....</b>                             | <b>11</b> |
| <b>2.2 Hogstmaskindata .....</b>                          | <b>11</b> |
| <b>2.3 Laserdata.....</b>                                 | <b>13</b> |
| <b>2.4.2 Beregningsceller.....</b>                        | <b>15</b> |
| <b>2.5.1 Modellerings teknikker .....</b>                 | <b>17</b> |
| <b>2.5.2 Random Forest .....</b>                          | <b>20</b> |
| <b>2.5.3. K-nærest neighbour.....</b>                     | <b>20</b> |
| <b>2.5.4 Lineær modellering .....</b>                     | <b>20</b> |
| <b>2.6.1 Versjoner og pakker.....</b>                     | <b>21</b> |
| <b>3 Resultat.....</b>                                    | <b>22</b> |
| <b>3.1 Totalvolum .....</b>                               | <b>22</b> |
| <b>3.1.1 Random Forest .....</b>                          | <b>22</b> |
| <b>3.1.2 K-Nearest neighbor .....</b>                     | <b>24</b> |
| <b>3.1.3 Lineær modellering .....</b>                     | <b>25</b> |
| <b>3.2 Sagtømmerevolum.....</b>                           | <b>26</b> |
| <b>3.2.1 Random Forest .....</b>                          | <b>26</b> |
| <b>3.2.2 K-Nearest neighbor .....</b>                     | <b>28</b> |
| <b>3.2.3 Lineær modellering .....</b>                     | <b>29</b> |
| <b>4.Diskusjon.....</b>                                   | <b>31</b> |
| <b>4.1 Hogstmaskindata .....</b>                          | <b>31</b> |

|  |    |
|--|----|
| 4.1.2 Hogstmaskindata som referansedata..... | 33 |
| 4.2 Celler .....                             | 33 |
| 4.3 Modelleringsteknikker .....              | 34 |
| 4.4 Totalvolum .....                         | 35 |
| 4.5 Sagtømmervolum.....                      | 36 |
| 5. Konklusjon.....                           | 38 |
| Referanser .....                             | 39 |

# 1. Innledning

## 1.1 Skog i Norge

Omtrent 38% av det landbaserte arealet i Norge er dekket av skog (Dalen, 2017). I 2020 ble det høstet 10,36 millioner m<sup>3</sup> med rundvirke i Norge (Landbruksdirektoratet, 2021). Dette er noe mindre enn hva som har vært vanlig de siste årene, men er fortsatt betraktelig mindre enn den totale tilveksten som er på over 24 millioner m<sup>3</sup> årlig (SSB, 2020). Mesteparten av skogen som hogges i Norge, hogges med Cut To Length (CTL) metoden. Ved CTL blir trærne hogd og aptert i skogen etter lengde og kvalitetskrav satt av industri og tømmerkjøpsorganisasjonene. Aptering er prosessen med å dele stokken til passende lengder og toppmål, for sikre best økonomisk utnyttelse av stokken (Leksikon, 2018). Det er stammevolumet som er den økonomisk viktigste delen av et tre. Sagtømmeret er det best betalte virket, når det ses bort fra nisje og spesialprodukter. Sagtømmeret går til sagbruk hvor det blir sagt til plank. Virke som ikke holder kravene til sagtømmerkvalitet går til massevirkeindustrien. Massevirkeindustrien har en større produktportefølje, mye av massevirkevolumet går til papir, og eksport.

### 1.2.1 Ressursoversikt

Virkeproduksjon av skogressurser er både areal og tidkrevende. Ett omløp med virkeproduksjon tar minst 50 år (Andreassen et al., 2018). For at skogforvalteren skal klare å holde tilstrekkelig oversikt over skogområdene, er det nødvendig med en skogbruksplan. Skogbruksplanen er et viktig verktøy som skogforvalteren bruker for å opprettholde en langsiktig og bærekraftig forvaltning av skogarealet. Skogbruksplanen er også et beslutningsverktøy, og kan brukes når det skal avgjøres hvilke tiltak som er aktuelle for forskjellige skogområder. Eksempler på tiltak er hogst, ungskogpleie og planting, samt tynning. Det er viktig at informasjonen som finnes i skogbruksplanen er riktig og oppdatert (Eid et al., 2004). Dette for å unngå økonomiske tap som følge av avgjørelser som blir gjort på feil grunnlag (Eid et al., 2004). Standardenheten i skogbruksplaner er individuelle bestand (Naesset, 1997a). Et bestand er definert som et geografisk sammenhengende område av like karakteristikk og homogen skog (Koivuniemi & Korhonen, 2006). Skogbruksplanen består

av informasjon om skogbestandene, som alder, høyde, treslagsfordeling, volum og grunnflate (Næsset et al., 2004). Alder, høyde, treslagsfordeling, Diameter i Brysthøyde (DBH), blir senere i denne oppgaven omtalt som egenskaper.

Volumoversikten i en skogbruksplan er viktig fordi den representerer en stor del av den økonomiske verdien for et skogområde (Naesset, 1997b). Ved kjøp og salg av rundvirke er det viktig å ha oversikt over virket som høstes fra skogen, for å ha oversikt over kostnader og inntekter ved skogsdriften. For industrien og tømmerkjøpsorganisasjonene er det viktig å vite hvor mye virke som kommer, når det kommer, og hvor det kommer fra, for å kunne optimalisere produksjonen (Kemmerer & Labelle, 2021; Lindroos et al., 2015).

I tillegg til å produsere virke til industri, skal skogen produsere andre økosystemtjenester som rekreasjonsområder, karbonlager og sikre bevarelse av biologisk mangfold. Det er viktig at skogforvalteren har oversikt og kunnskap til å forvalte skogområdene, for å drive et bærekraftig skogbruk der skogforvalteren får oppfylt sine mål. Noen skogforvaltere vil fokusere på å maksimere den økonomiske avkastningen, mens andre vil maksimere biologisk mangfold ved eksempelvis å øke andelen død ved. Uansett er det nødvendig å ha oversikt over skogressursen. En slik oversikt kommer som regel gjennom en skogbruksplan. I denne oppgaven blir begrepet skogforvalter brukt som samlebegrep for de som informasjonen er aktuell for. Oftest vil dette være skogeieren, noen skogeiere velger å sette bort forvaltningen av skogeiendommen, og flere skogeiendommer er eid av kommune, stat, allmenninger eller lignende.

### **1.2.2 Utvikling av ressurskartlegging i skog**

Før moderne metoder ble tilgjengelig ble all skogressurskartlegging gjort ved manuelle registreringer, og inneholdt vesentlig mindre informasjon enn ved dagens kartleggingsmetoder. De første skogressurskartene er datert helt tilbake til 1870-1880 (Næsset, 2014). Informasjonen ble fremstilt på kart og inneholdt lite spesifikke egenskaper, men var på et generelt nivå med topografi, eiendomsgrenser og stedsnavn, veier og lignende. Etter hvert som bestandsskogbruket ble den dominerende driftsformen, ble det større heterogenitet i skogen som økte behovet for detaljkunnskaper om enkeltbestandene.

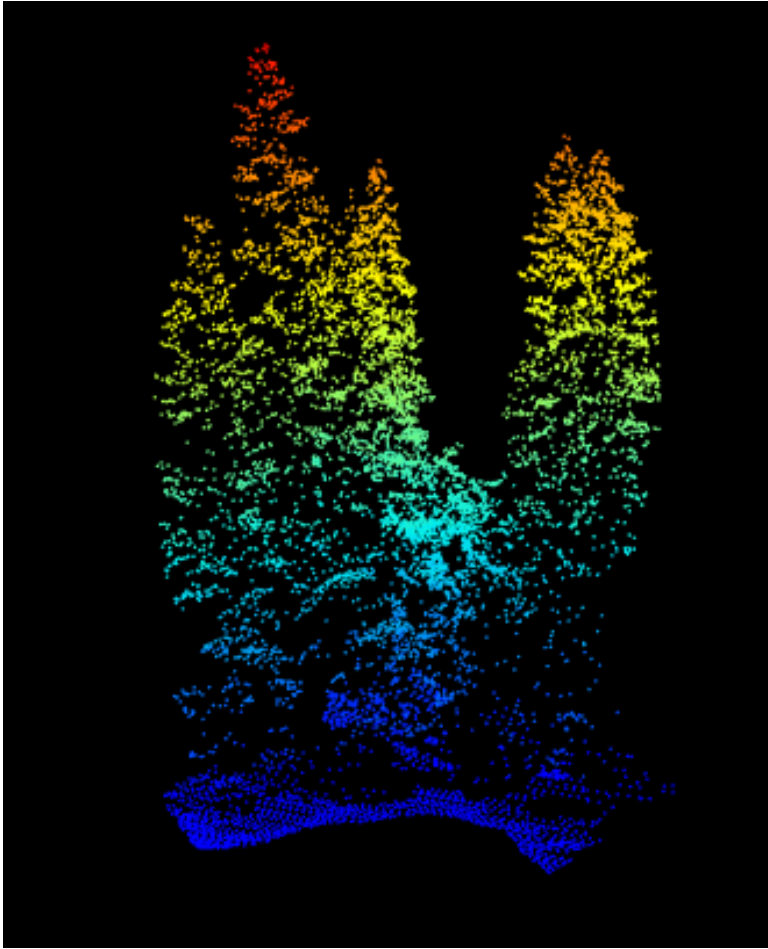
Norge var det første landet i verden som startet med nasjonal skogkartlegging, allerede i 1919-, men skogressurskartlegging for skogforvalteren ble ikke vanlig før på 1950 og 1960

tallet (Næsset, 2014). Tidligere skogressurskartlegging foregikk ved at enkeltbestand ble definert fra flybilder og topografiske kart (Næsset, 2014). Bestand ble befart, og enkle feltmålinger kombinert med erfaring fra feltmåleren ble brukt for å fastsette egenskaper og tiltak for bestandene. I perioden fra omtrent 1975 til rundt 1990 ble det et metodisk skifte hvor stereo fotogeometri, kombinert med Geografiske informasjons system (GIS) ble standard metode ved skogressurskartlegging (Næsset, 2014). Stereo Fotogeometri er en målemetode hvor bildepar fotografert med overlapping blir betraktet i stereo, som gir bildet rom og dybde (Leksikon, 2019). Skogegenskaper som volum ble predikert ved hjelp av tolkningen av fotogeometrien og etablerte volummodeller (Næsset, 2014). Dette reduserte behovet for å gjøre manuelle feltmålinger. Volumpredikering ble gjort ved hjelp av modeller (Næsset, 2014). Undersøkelser i felt gikk ut på å bekrefte volummodellene som ble brukt, eller bestemme spesifikke skjøtselstiltak (Næsset, 2014). Dette la grunnlaget for at den areal-baserte metoden, som går ut på at skogbestandene er basisenheter i en skogbruksplan (Næsset, 1997a). Sent på 1990 tallet ble digital fotogeometri innført, som produserer tre-dimensjonale punktskyer og ble brukt til volumpredikering og bestandsinndeling (Næsset, 2014). Fra 1990 tallet begynte skogsektoren også se på muligheten for å bruke Flybåren laser skanning (FLS) til skogressurskartlegging (Næsset, 1997a; Næsset, 1997b).

FLS er metode hvor en laserskanner er montert på fly og koblet til et Global Navigation Satellite System (GNSS). Laserskanneren sender ut laserpulser i en gitt retning og måler tiden fra laserpulsen blir sendt, til den samme laserpulsen har returnert til laserskanneren (Evans et al., 2009; Næsset, 1997a). Laserpulsen har en hastighet lik lysets hastighet og den klarer å regne avstanden til objektet som reflekterte laserpulsen (Evans et al., 2009). Laserskanneren kjenner sin egen posisjon gjennom GNSS-systemet den er koblet til, retningen og avstanden til laserpulsen, klarer laserskanneren å opprette georefererte punkter til objektet som reflekterte laserpulsen (Evans et al., 2009). Laserpulser som treffer skog kan gi flere refleksjoner, både når de treffer vegetasjon og bakken (Evans et al., 2009; Næsset, 2002). Refleksjonene som er lengst unna sensoren klassifiseres som bakketreff. Refleksjonene som er høyere enn bakketreff kalles vegetasjonstreff (Næsset, 2002). Hver refleksjon kan bearbeides til å bli georeferert punkt. Summen av punktene fra en FLS kalles en punktsky og kan sees i Figur 1. Punktskyen kan bearbeides videre for å klassifisere skogegenskaper (Evans et al., 2009). Egenskaper for punktsky kan sees i Tabell 1, og eksempel på Punktsky kan sees i Figur 1.

Med innføringen av FLS ble også metoden hvor man deler områdene inn i celler introdusert (Naesset, 1997a; Naesset, 1997b). Slik fremgangsmåte er vanlig ved skogressurskartlegging i dag og kalles Area Based Approach (ABA) (Naesset, 1997a; Naesset, 1997b). Volum, og andre egenskaper på bestandsnivå er summen av egenskapene til cellene innenfor bestandet. Verdiene til cellene predikeres ved hjelp modeller som er basert på korrelasjonen mellom laservariablene, og referansedata samlet inn fra et utvalg av cellene. Referansedata samlet inn ved hjelp av manuelt feltarbeid er ressurskrevende, og er vesentlig del av kostnaden knyttet til skogressurskartlegging. Ved innsamlingen av referansedataene er det viktig med nøyaktig georeferert posisjon til referansedataene for at de skal kunne sammenstilles med FLS dataene.

FLS er den dominerende måten å volumpredikere skog i skandinaviske land (Maltamo et al., 2014). FLS er dyrt å gjennomføre, men har vist seg å være det mest kostnadseffektive alternativet (Eid et al., 2004). En stor del av gevinsten ved bruk av FLS kommer fra at kostnaden ved å gjøre feil valg i fremtiden reduseres, ettersom laserdataene gir mer detaljert, og nøyaktige data om skogen og skogegenskapene. Dette gjør at FLS kan gi bedre estimater på bestandsegenskapene (Eid et al., 2004). laserdata har ofte blitt tilgjengelige fra staten i forbindelse med lagring av digitale terrengmodeller, og derav blitt tilgjengelig forholdsvis billig for skogbruket.



*Figur 1. Eksempel på en punktsky fra siden. Et punkt på er en laserpulsrefleksjon, samlingen av alle punktene er en punktsky. For å gjøre det enklere å visualisere punktskyen er det fargekoder på figuren. Blå punkter som er nederst i figuren er bakketreffene. Punktene som gradvis skifter farge fra blå, grønn, gul og rød farge, er vegetasjonstreff. De røde punktene er de høyeste punktene fra bakken.*

### **1.3.1 Hogstmaskindata**

Moderne hogstmaskiner som hogger etter CTL metoden, opparbeider og lagrer store mengder data (Maltamo et al., 2019; Noordermeer et al., 2021). I apteringsprosessen måler hogstaggregatet kontinuerlig lengde og diameter av stammen. Føreren gjør de visuelle vurderingene og registrerer treslag og kvalitet til apteringsdatamaskinen.

Apteringsdatamaskinen kommer med forslag til kappepunkter, som hogstmaskinen kapper på, med mindre føreren overstyrer maskinen. Dataene som blir opparbeidet gjennom

apteringsprosessen lagres på maskinens datamaskin (Noordermeer et al., 2021; Stendahl & Dahlin, 2002).

Informasjonen blir lagret i en filstruktur som er beskrevet av Standard for Forest Machine data 2010 (Stanford 2010) (Johan J. Moller, 2013) StanFord 2010 sikrer at dataene som kommer fra moderne skogsmaskiner er lesbare, og har lik oppbygging på tvers av maskinmodeller. Standarden inneholder flere filtyper, men i denne oppgaven er det den mest detaljrike fila, Harvester Production Report (HPR) som blir brukt. HPR-fila inneholder informasjon om hogde trær, som stamme ID, treslag, DBH, antall stokker, tidsstempel, kommersiell høyde og diameter for hver 10 cm langs stammen (Noordermeer et al., 2021). Data om hver stokk blir også lagret, da som stokk ID, stamme ID, sortiment, dimensjoner og volum over og under bark (Noordermeer et al., 2021). I tillegg er det mulig, dersom GNSS funksjonen er koblet inn, å hente ut maskinens geografiske posisjon (Noordermeer et al., 2021), retningen til kranen, vinkelen til bom og stikke for å posisjonere trær med posisjon på <1 meter feilmargen på snitt (Noordermeer et al., 2021).





*Figur 2: Sirkel 1: GNSS-antenner plassert på toppen av førerhytten. Sirkel 2: ekstra kran sensor montert mellom bom og stikke for å måle vinkel mellom bom og stikke kontinuerlig. Sirkel 4: teleskop fra stikke, på halv teleskoplengde. Sirkel 4 og 5: øvre kvistkniver, er montert på aggregatet og måler diameter av stammen kontinuerlig. Sirkel 6: målehjulet, målehjulet er plassert inne i rammen av aggregatet og måler lengden av stammen. Foto: Lennart Noordermeer*

#### **1.4 Hogstmaskindata som referansedata**

Muligheten for å bruke hogstmaskindata som referansedata til predikering av skogegenskaper, har lenge blitt vurdert (Lindroos et al., 2015). Blant annet forsøkte Stendahl (2002) allerede i 2002 å bruke hogstmaskindata til å estimere egenskaper på gjenstående skog etter tynning, basert på trærne som ble tatt ut i tynningen (Stendahl & Dahlin, 2002). Lindroos (2015) beskrev at for å bruke hogstmaskindata som referansedata til å utvikle FLS modeller trengtes det en nøyaktighet til treets plassering på bedre enn 1 meter (Lindroos et al., 2015). Dette nøyaktighetskravet ble nådd på maskinen som Noordermeer et al. (2021) og Hauglin et al. (2018) brukte i sine forsøk. Det er den samme maskinen som blir brukt som referansedata i denne oppgaven, og kan sees i Figur 2.

Hauglin et al. (2018) sammenlignet volumpredikering med FLS og referansedata fra hogstmaskindata med aggregatpresisjon på bedre enn 1 meter feilposisjonering. Med referansedata samlet inn med manuelle feltmålinger. Resultatene viste at modellene basert på hogstmaskinens referansedata inventerte volum med tilsvarende nøyaktighet som de manuelle feltmålingene, og viste at moderne hogstmaskiner kan brukes som referansedata ved skogressurskartlegging.

Spesielt for bruk av hogstmaskindata som referansedata er at de er totaldekkende, og har målt på alt volumet som er hogd i bestandet. Ikke alle trærne blir hogd når bestand hogges. Det er flere grunner for å sette igjen trær. Noen trær vil stå igjen av miljøhensyn. Disse kan eksempelvis stå igjen i kantsoner mot vann og vassdrag, enkeltvis spredt på driftsområdet eller samlet i grupper hvor de skal stå og dø en naturlig død (Andreassen et al., 2018). Det er et krav i miljøstandarden PEFC at det skal stå igjen minst ti livsløpstrær for hvert hektar (PEFC, 2015). Halvparten av livsløpstrærne kan stå igjen som høystubber, som er kappet høyt oppe på stammen så det blir stående igjen en grov død stamme (PEFC, 2015). Det er også

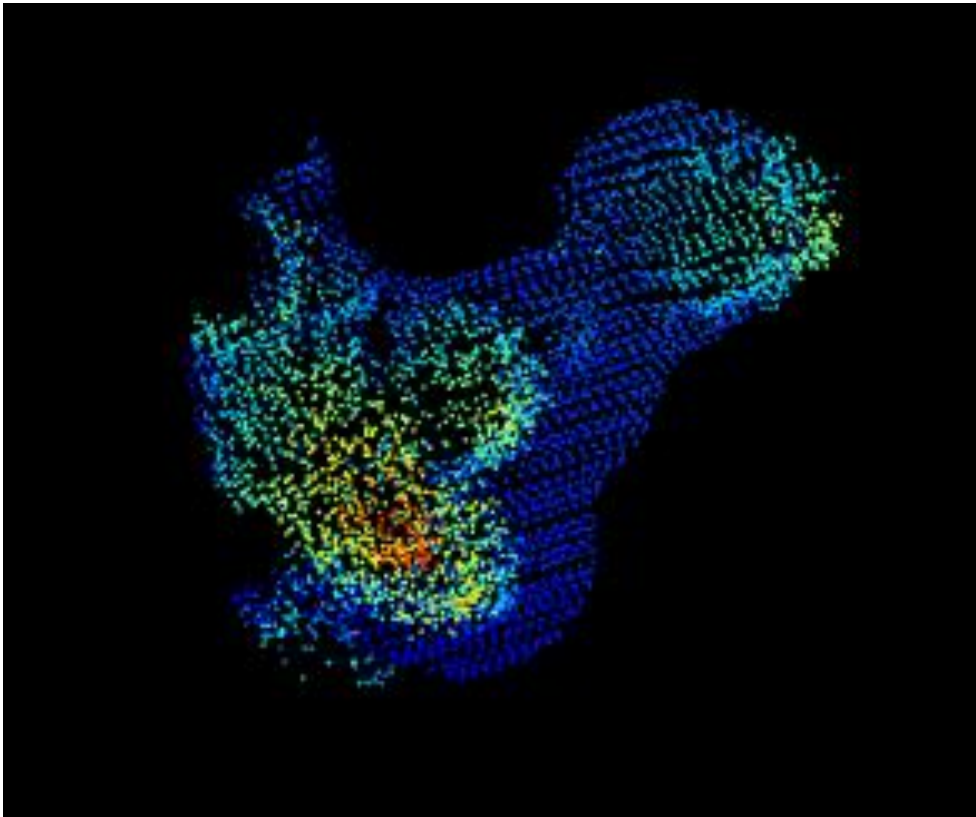
skogfaglige grunner for å sette igjen trær ved hogst, og for flere av driftene som er inkludert i dette datasettet, er det satt igjen frøtrær spredt på hogstflatene. Frøtrær blir satt igjen for å spre frø på hogstflatene for å legge grunnlaget for neste generasjon skog. Frøtrær er som oftest voksne godt utviklede furutrær.

## **1.5 Kanteffekter**

Når skogen blir delt inn i celler ved bruk av ABA, vil det for flere av cellene være punkttreff av trekronene der stammen står i nabocellen. Dette vil utløse såkalte kanteffekter som påvirker laseregenskapene til cellene. Cellene får andre laseregenskaper enn det volumet i cellen skulle tilsi. Ved volumpredikering på bestandsnivå er ikke kanteffektene så farlige, på grunn av at over-og underpredikeringen vil utjevne hverandre og være tilfeldig feil. Det er en risiko for at det kan bli en systematisk feil dersom cellene som blir brukt til referansedata har store kanteffekter. ABA tar ikke hensyn til kanteffekter.

### **1.5.1 Enhanced Area Based Approach**

Nyere studier av blant annet Packalen et al. (2015) og Pascual (2019) har vist gode resultater med bruk av Enhanced ABA (EABA). EABA er en metode som går ut på at cellene blir tilpasset trærnes posisjoner og kroneform for å redusere kanteffekter. Fra punktskyen er det mulig å lokalisere enkeltrær og dens kronedekke ved at man finner lokale maksimumshøydepunkter (Vauhkonen et al., 2012). EABA definerer hvert tre med dens kronedekke og klipper hele kronen til en celle. Pascual (2019) sammenlignet ABA med EABA og fikk bedre resultater for samtlige modeller som brukte EABA. Det ble også et skifte i hvilke laservariabler som hadde størst påvirkning på modellene, der høydevariablene fikk mindre påvirkning og tetthetsvariablene fikk større påvirkning ved bruk av EABA. Eksempel på hvordan en celle kan se ut med EABA kan sees i Figur 3. legg merke til at det kan skimtes krontreff til naboceller til venstre i figuren.



*Figur 3. Punktsky av en celle laget med EABA. Fargekodene i Figur 3 betyr det samme som i Figur 2. Fra Figuren kan vi se en celle med flere vegetasjonstrekk som ligner tre kroner. Disse er også tre kroner. Uthulingen som går inn i cellen fra cellekanten er tre kroner fra naboceller som er klippet ut og ligger i nabocellene.*

## **1.6 Forskning**

Det har tidligere blitt gjort flere vellykkede forsøk for å predikere totalvolum med bruk av hogstmaskindata som referansedata til FLS. Blant annet predikerte Maltamo et al. (2019) bestandsegenskaper som totalvolum med RMSE% på 9%. Hauglin et al. (2018) viste at det var mulig å bruke hogstmaskindata som referansedata til FLS, som erstatning eller supplementering til manuelle feltmålinger ved bruk av RF, Knn og LM, med tilsvarende like resultater.

## **1.7 Denne oppgaven**

Tidligere studier har ikke testet bruk av hogstmaskindata med flere laserskanninger, som det er i denne oppgaven. Hogstmaskinen som ble brukt i denne oppgaven har lite ekstrautstyr, men posisjonen til aggregatposisjoneringen var fortsatt gode nok (Noordermeer et al., 2021). Det er viktig å utvikle og teste modelleringsteknikker for volum for å forbedre ressursoversikten til skogforvalteren. Det er viktig for industrien og logistikselskapene å ha oversikt over volum av forskjellig kvalitet for å optimalisere virkestrømmer (Barth & Holmgren, 2013).

### **1.7.1 Problemstilling**

Hovedmålet i denne oppgaven var å undersøke om hogstmaskindata fra hogstmaskin i kommersiell drift, kan brukes som referansedata til FLS. For å predikere totalvolum og sagtømmervolum for geografien maskinen opererer i, med pålitelige estimater.

Det er også et delmål i denne oppgaven å evaluere hvilke prediksjonsmodeller som vil fungere best til å predikere totalvolum og sagtømmervolum på cellenivå, med hogstmaskindata som referansedata ved en arealbasert takst.

## 2. Material og metode

### 2.1 Studieområde

Dataene som ble brukt i dette oppgaven kommer fra 24 forskjellige drifter i kommunene Etnedal, Nord Aurdal og Nordre land i Innlandet fylke. For de fleste driftene var gran (*Picea abies*) det dominerende treslaget, med innslag av furu (*Pinus sylvestris*) og andre løvtreslag, hovedsakelig bjørk (*Betula pubescens*). For enkelte drifter var også furu det dominerende treslaget med innslag av gran og løvtreslag. Middels til lave boniteter er typisk for området. Skogen strekker seg fra omtrentlig 150 til 1000 meter over havet.



Figur 4: Kart over Norge som viser kommuner og fylkesgrenser. Kommunene som er med i oppgaven er markert i rødt.

### 2.2 Hogstmaskindata

Hogstmaskinen som ble brukt i denne oppgaven, var en Komatsu 931 XC med en 230 H kran med ti meters rekkevidde og et C144 hogstaggregat. Hogstmaskinen går i kommersiell drift og har gjort det siden maskinen var ny i oktober 2018. Maskinens datasystem er maxiXplorer.

Innledningsvis ble versjon 3.1.9 benyttet, men i løpet av prosjektet ble datasystemet oppgradert til versjon 3.2.1, uten at dette hadde påvirkning på aggregatposisjoneringen. På aggregattypen til denne maskinen skjer diametermåling ved de to øverste kvistknivene, (sirkel 4 og, - 5 Figur 2). Lengdemålingen blir gjort av lengdemålingshjulet som er plassert inne i rammen på hogstaggregatet (sirkel 6, Figur 2).

Maskinen ble innlemmet i prosjektet PRECISION siden den var ny. Maskinens standardutstyr var ikke tilstrekkelig for å oppnå en aggregatposisjon på bedre enn en meter fra sann posisjon. For å få en bedre aggregatposisjonering enn en meter, ble det gjort to tiltak for å forbedre maskinens aggregatposisjon.

1. Ekstra kransensor montert mellom bom og stikke (punkt 2, Figur 2). Sensoren måler kontinuerlig vinkelen mellom bom og stikke og lagrer vinkelen i HPR-fila for hvert kapp. Kransensoren ble installert som et tilleggutstyr av maskinleverandør.
2. Hogstmaskinens standard GNSS ble erstattet av to Real Time Kinematic (RTK) septentario AsteRx-U Differensial GNSS antenner (Kemmerer & Labelle, 2021). I maskinens HPR-filer ble lengde og breddegrad posisjoneringen til maskinen lagret fra erstatnings GNSS ved hvert kapp. Posisjoneringen til GNSS systemet er forventet å ha en horisontal nøyaktighet på omtrent en cm (Kemmerer & Labelle, 2021).

Maskinen er den samme som Noordermeer et al. (2021) brukte. Gjennomsnittlig avstand fra stubbeposisjonen i HPR fila og den virkelige posisjonen til senter av stubben i felt var 0,88 meter ved 288 kontrollmålinger. Feilen varierte fra 0,14 til 2,85m med et standardavvik på 0,38m (Noordermeer et al., 2021). Det skal også nevnes at det ikke var måleutstyr på teleskopen (Sirkel 3, Figur 2). Teleskopen har en rekkevidde fra 0 til 2,14 meter. I aggregatposisjonerings-beregningene ble teleskopens lengde satt til 1,07 m.

Hogstmaskindata som ble brukt i denne oppgaven omfatter 72 671 trær. Hogstmaskinen har samlet inn data fra maskinen var ny i oktober 2018. Den hogger årlig cirka 30 000 m<sup>3</sup>, der alt av data har blitt disponert til prosjektet. Det har vært noen problemer med en av de ettermonterte antennene (Figur 2) løsnet og gjorde at det ble en dårlig nøyaktighet på aggregatets posisjon. Dataene som ble samlet inn i perioder som den nevnte antennen var løs, er blitt fjernet fra datasettet. Maskinen hogger over en stor geografi, laserdata er ikke tilgjengelige for hele geografien. Drifter som er utenfor tilgjengelige laserdatasett, er derfor ekskludert fra denne oppgaven.

Definisjonen av sagtømmer har forandret seg i løpet av hogstperioden. Maskinen har hogd for forskjellige tømmerkjøpere som leverer til forskjellige sagbruk. Disse har varierende krav til kvalitet og dimensjoner for sagtømmeret. Dette reduserer nøyaktigheten til sagtømmerpredikeringen gjort i dette oppgaven. Sagtømmerdefinisjonene har som nevnt variert, men har vært lik eller tatt utgangspunkt fra gjeldende sagtømmerreglement der definisjonene har grunnlaget fra norsk virkesmåling (Virkesmåling, 1995) Minste godkjente diameter for sagtømmer har vært 12 og 13 cm under bark (Råheim, 2021), i forsøksperioden. Det har vært tider hvor de minste sagtømmerdimensjon i diameter og lengde har gitt dårligere pris enn for massevirke. Virket har blitt aptert til best betalte pris, til enhver tid.

### **2.3 Laserdata**

Laserdataene som ble brukt i dette prosjektet kommer fra flere prosjekter, og aktører. Det har ført til et datasett av ulik kvalitet og alder. Tettheten til punktskyene varierer fra minimum 0,7 til 10 pkt/m<sup>2</sup>. Totalt er det fem laserdatasett i denne oppgaven. Data om prosjektene kan sees i Tabell 2.

Tabell 2. Data om FLS og andel av datasettet i treantall.

| År   | Instrument                | Pulsrate<br>(KHz) | Scan<br>Rate<br>(Hz) | Flyhøyde | Tids-<br>periode | Punktetthet<br>Gjennomsnittlig | treantall |
|------|---------------------------|-------------------|----------------------|----------|------------------|--------------------------------|-----------|
| 2013 | TopEye<br>S/N 444         | 200               | 92                   | 1500     | Mai-Juli         | 7,7                            | 28 169    |
| 2016 | Riegl<br>LMS Q-<br>1560   | 400               | 100                  | 2900     | September        | 3,2                            | 12 955    |
| 2017 | Riegl<br>LMS Q-<br>1560 I | 700               | 240                  | 2300     | Juli             | 6,8                            | 4 784     |
| 2019 | Leica<br>ALS70-<br>HP     | 495               | 69                   | 1150     | August           | 5,9                            | 22 163    |
| 2020 | Riegl VQ-<br>1560 II      | 749               | 158                  | 1100     | Juni             | 10,4                           | 4 600     |

Det har ikke blitt sett på forskjell mellom nye og eldre datasett i denne oppgaven.



## 2.4 Dataanalyse

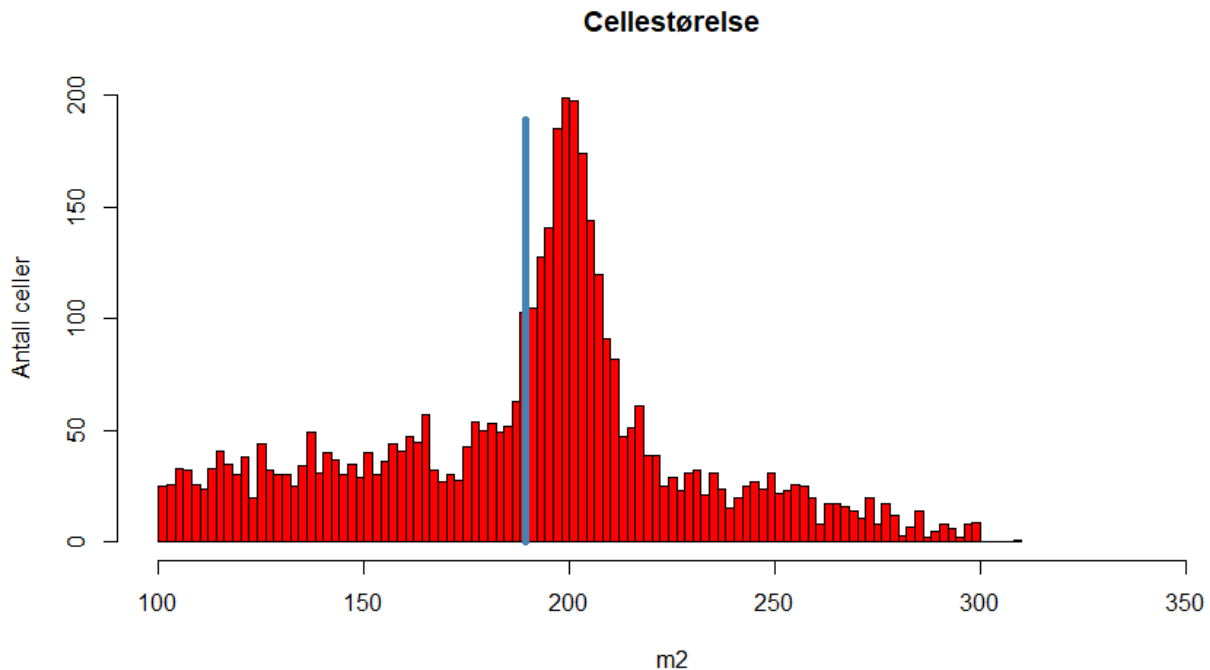
Tabell 1. Egenskapstabell for laservariabler.

| Variabel  | Beskrivelse  |
|---|--|
| Hmean   | Gjennomsnittlig høyde over bakken.   |
| Hmax  | Høyeste punktet i punktskyen.  |
| Hsd   | Standardavviket til høydefordelingen.  |
| Hcv   | Variasjonskoeffisienten til høydefordelingen.  |
| Hkurt   | Beskriver høydefordelingen av punktskyen.  |
| Hskewness   | Beskriver skjevheten i høydefordelingen.   |
| H10, H20, H30, H40;<br>H50, H60, H70, H80<br>og H90 | Høydepercentiler, hver Høydepercentil representer gjennomsnitthøyde til laserpulsrefleksjonene innenfor percentilen. |
| D0-D10  | Beskriver tettheten av laserpulser som er lik eller høyere tilsvarende høydepercentil.                               |
| n   | Antall laserpulsrefleksjoner innenfor cellen.  |
| year  | Årstall for gjennomføringen av FLS.  |

### 2.4.2 Beregningsceller

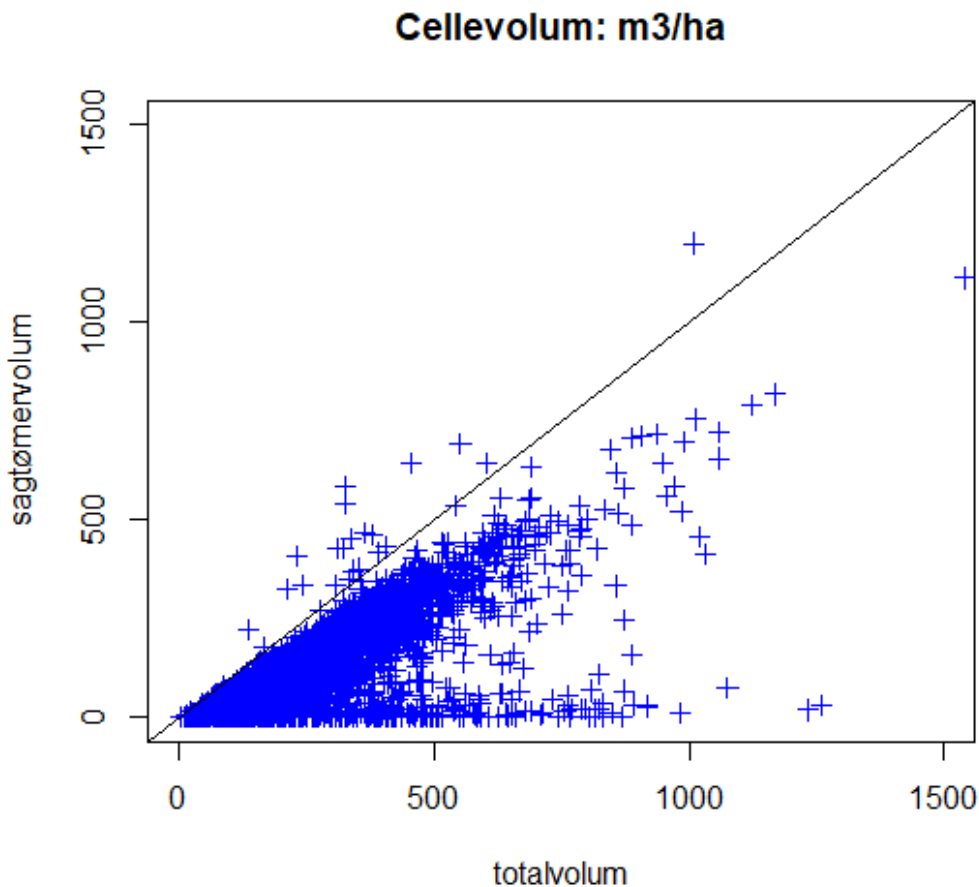
Cellene var i utgangspunktet på 200 m<sup>2</sup>, men, med EABA metoden som ble brukt i denne oppgaven, blir cellestørrelsen endret når kronetreffene til cellene klippes inn eller ut av cellene. For å unngå for store forskjeller ble de minste cellene satt sammen med naboceller, celler større enn 300 m<sup>2</sup> blir delt opp til to individuelle celler. Fra Figur 5 kan man se et histogram over cellestørrelsene. Det er en celle som er over 300 m<sup>2</sup>, av ukjente grunner. Det var ingen celler på under 100 m<sup>2</sup>, og gjennomsnittsstørrelsen på cellene var 189 m<sup>2</sup>,

representert av blå linje på Figur 5. Valget av 200 m<sup>2</sup> store celler som utgangspunkt kommer fra tidligere studier av Maltamo et al. (2019) og masteroppgaven (Dieset, 2017) som begge fant at 200 m<sup>2</sup> store celler var passende for å kartlegge skogegenskaper ved bruk av hogstmaskindata.



Figur 5. Cellestørrelse oppgitt i m<sup>2</sup>, gjennomsnittlig cellestørrelse var 189 m<sup>2</sup> og er representert av blå linje i Figur 5.

Totalvolum og sagtømmervolum er hentet fra maskinens beregninger som er lagret i maskindatamaskinen og presentert i HPR filene. Sagtømmervolumet er volumet av alle sagtømmerstokker som er hogd innenfor den aktuelle cellen. Sagtømmervolumet er regnet ut stokkvis av hogstmaskinen som bruker formel 4, som maskinen får oppgitt av prislista i maskinen. Totalvolumet er regnet som alt volum som blir hogd innenfor en kvalitetsklasse. Det vil si at volum som blir kappet bort i aptering, eller ikke holder lengde og diameterkrav, er ikke medregnet volum i denne oppgaven. Totalvolumet blir regnet ut stokkvis med Formel 5. Alle volumberegninger og målinger er gjort på bark, alle volumverdier som er presentert i denne oppgaven er over bark. Volum i m<sup>3</sup>/ha for hver celle kan sees i Figur 6. Det finnes flere ekstreme målinger i datasettet. Blant annet kan det sees i Figur 6 at flere celler har over 1000 m<sup>3</sup>/ha, som virker svært usannsynlig for området som maskinen har arbeidet i. Det er også flere av cellene som har mer sagtømmervolum enn totalvolum.



*Figur 6. Volum oppgitt i m<sup>3</sup>/ha målt på cellenivå. Svart heltrukket linje gjennom Figuren har stigningstall lik 1, som vil si at punktene som ligger over linjen har mer sagtømmervolum enn totalvolum.*

### **2.5.1 Modelleringssteknikker**

I denne oppgaven er det brukt tre forskjellige modelleringssteknikker for å estimere totalvolum og sagtømmervolum. Metodene er Random Forest (RF), K Nearest Neighbour (Knn) og Lineær modellering (LM). Knn og RF er ikke parametriske metoder, mens LM er en parametriske metode. Fordelene ved å bruke parameterfrie metoder i denne oppgaven er at antagelsen om uavhengighet for parameterne ved parametriske metoder er brutt da cellene ligger inntil hverandre, og derav påvirker hverandre. Bruk av data som ikke er uavhengige går imot prinsippet om individuelle uavhengige observasjoner.

For å teste datasettet ble det brukt metoden: «Random subsampling cross validation». Metoden trekker et tilfeldig utvalg av de 4216 cellene. Modellerings-teknikkene trener modellen på en del av datasettet, og tester på den andre delen av datasettet. Prosessen kan gjentas uendelig mange ganger, og hver gjentakelse kalles for en iterasjon. I denne oppgaven ble 10 % brukt som kalibreringsdata og 90 % som treningsdata. Alle resultater vist i denne oppgaven er fra resultater med 1 000 gjentakelser. Alle laseregenskaper som er nevnt i Tabell 1, var tilgjengelig for modellerings-teknikkene. Modellerings-teknikkene ble målt på hvor gode de var på RMSE%, SD, gjennomsnittlige verdier og maks feil av predikert volum.

### 2.5.2 Formler

Root Mean Square Error (RMSE%) beskriver gjennomsnittsfeilen i prosent av gjennomsnittlig volum i kalibreringsdatasettet. RMSE er summen av predikert volum minus sant volum, og lagt sammen, kvadrert og delt på antall observasjoner. Lav RMSE% betyr at modellen er god.

RMSE% er blir regnet ut som vist i Formel 1.

Formel 1.

$$RMSE\% = \left( \sqrt{\frac{\sum_i^n 1(v_i - \hat{v}_i)^2}{n}} \right) / \hat{v}_i * 100$$

*n*: antall celler. *v<sub>0</sub>*: observert volumverdi i celle *i*, *ŷ<sub>i</sub>*: inventert volum i celle *i*.

SD beskriver spredningen av predikasjonene og gir den gjennomsnittlige avstanden til gjennomsnittet. Lav SD betyr at predikasjonene har bommet lite.

Standardavvik regnet ut som vist i formel 2.

Formel 2.

$$SD = \sqrt{\frac{\sum_{i=1}^n (v_i - \bar{v})^2}{n - 1}}$$

$n$ : antall celler,  $v$  volumverdien for celle  $v_i$ ,  $\bar{v}$ : gjennomsnittlig volumverdi for alle celler.

Gjennomsnittlig feil beskriver gjennomsnittet av alle predikasjonene.

Gjennomsnitt har blitt tatt med i denne oppgaven fordi den kan indikere om det er systematiske feil i datasettet.

Gjennomsnittet blir regnet ut som vist i formel 3.

Formel 3.

$$\text{Gjennomsnittlig feil} = \bar{v} = \frac{1}{n} \sum_{i=1}^n v_i$$

$n$ : antall celler.  $v_i$  volum for celle  $v_i$

Formel 4 viser hvordan maskinen regner oppgjørsvolum for sagtømmervolum. Volumet regnes ut fra diameter av toppsylynderen, pluss en halv cm i diameter for hver meter i lengde.

Formel 4 blir kalt for  $m^3fmiNO$  i apteringssystemet.

Stokkvis sagtømmervolum blir regnet ut som vist i Formel 4 (FUNT, 1997).

Formel 4.

$$V = 3,14 * \left( D_t + 0,5 + \left( \frac{L}{20} \right) \right)^2 * L * 1/400$$

$V$ : volum i  $dm^3$ .  $D_t$  toppdiameter i cm.  $L$  stokkens lengde i dm.

Formel 5 Viser hvordan maskinen regner oppgjørsvolum for massevirke. Volumet regnes fra diameter midt på stokken og stammens lengde. Formel 5 blir kalt  $m^3f$  i apteringssystemet.

Stokkvis ikke-sagtømmervolum blir regnet ut på følgende måte (FUNT, 1997)

Formel 5.

$$V = 3,14 * (D_m + 0,5)^2 * L * 1/400$$

*V: volum i dm<sup>3</sup>, D<sub>m</sub> er diameteren midt på stokken, L er lengden*

### **2.5.2 Random Forest**

RF er en datalæringsmetode som går ut på at det blir bygget et stort antall ukorrelerte beslutningstrær med utgangspunkt i datasettet. Regresjonstrærne vokser ved at RF algoritmen tar prøver av datasettet, for å dyrke frem beslutningstrær bestående av noder og blad. Noder er valg RF lærer seg, og blad er sluttverdier. RF gjentar seg selv flere ganger med tilfeldig prøvetaking fra datasettet. Prosessen med tilfeldig prøvetaking er kjent som bootstrapping. Bootstrappingsprosessen blir gjentatt et definert antall ganger, og den kan lage stort antall beslutningstrær. Når det kommer nye data som skal predikeres blir de predikert ved at de følger nodene i beslutningstreet til de kommer frem til bladene. Bladene er verdien som cellen vil få, og siden prosessen gjentas flere ganger vil sluttverdien til cellen være gjennomsnittet av verdiene til bladene for den aktuelle cellen.

### **2.5.3. K-nærest neighbour**

Knn er en relativ enkel, og intuitiv ikke-parametrisk datalæringsmetode som kan brukes til regresjonspredikering. Metoden bygget opp ved at modellen kjenner igjen de nærmeste naboene til en ny observasjon, og gjør en predikering basert på verdien til de nærmeste naboene. K er antall naboer som det skal sammenlignes med. Det ble kjørt flere runder med flere forskjellige K både 1, 3, 5 og 10.

### **2.5.4 Lineær modellering**

LM er som nevnt i innledningen den eneste parameterbaserte metoden som er brukt i denne oppgaven. Parameterne som ble brukt til LM ble valgt ut basert på hvilke parametere som hadde størst påvirkning fra RF. Det ble derfor valgt nye parametere for hver av de 1 000 gjentakelsene i Random subsampling cross validation.

LM lager så en funksjon basert på de tre parameterne som blir valgt ut. Denne funksjonen inneholder krysningspunkt på y-aksen, og stigningstall for hver av de tre predikatorvariablene.

### **2.6.1 Versjoner og pakker**

Programmet R-studio har blitt brukt til all databearbeiding i denne oppgaven. Scriptene som har blitt brukt har i stor grad blitt laget av veileder. Hogstmaskindata er ikke vanlig å bruke innen skoginventering enda, så det er ikke noen standard pakker tilgjengelige. For bearbeiding av FLS data ble det brukt lidR pakken av (Romain, 2021) versjon 3.1.3 og lasR pakken Ørka (2018) versjon 0.3.3.

RF modellen kommer fra Ranger pakken i R og versjonen 0.12 ble brukt. Knn som ble kommer fra pakken FNN i R og versjonen var 1.1.3 LM ble utregnet i R og kommer fra pakken stats, versjon 4.0.4. For å kjenne igjen kronetreffene fra FLS dataen som skulle legges til eller fjernes fra en celle ble det brukt R pakken itcSegment Dalponte (2016)

## 3 Resultat

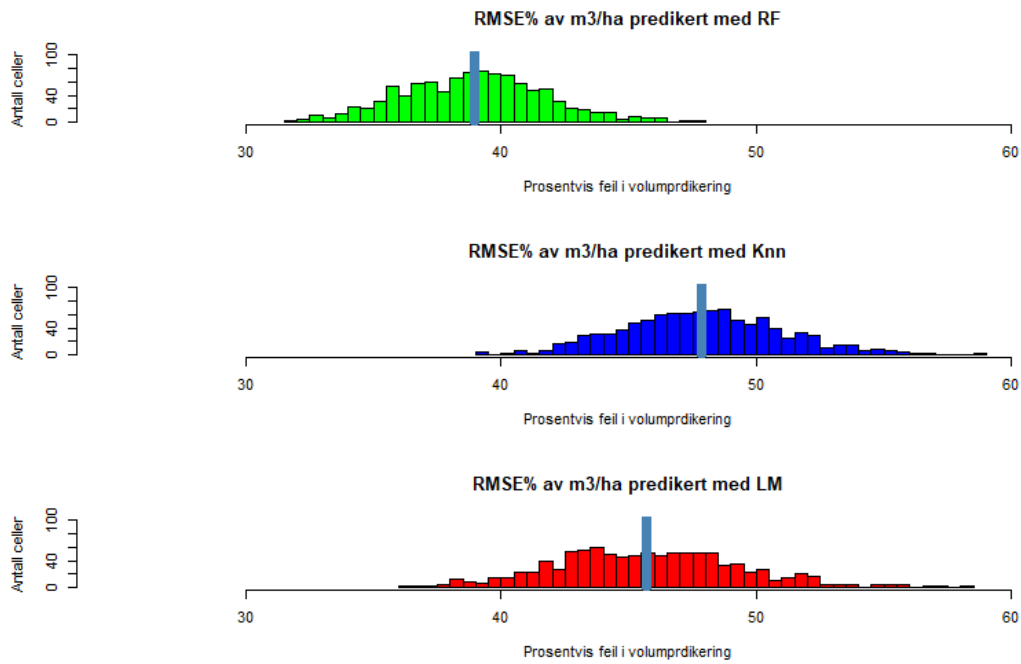
Det har blitt utregnet resultater som blir i både prosentvis av volumpredikering, absolutte verdier og gjennomsnittlige verdier..

### 3.1 Totalvolum

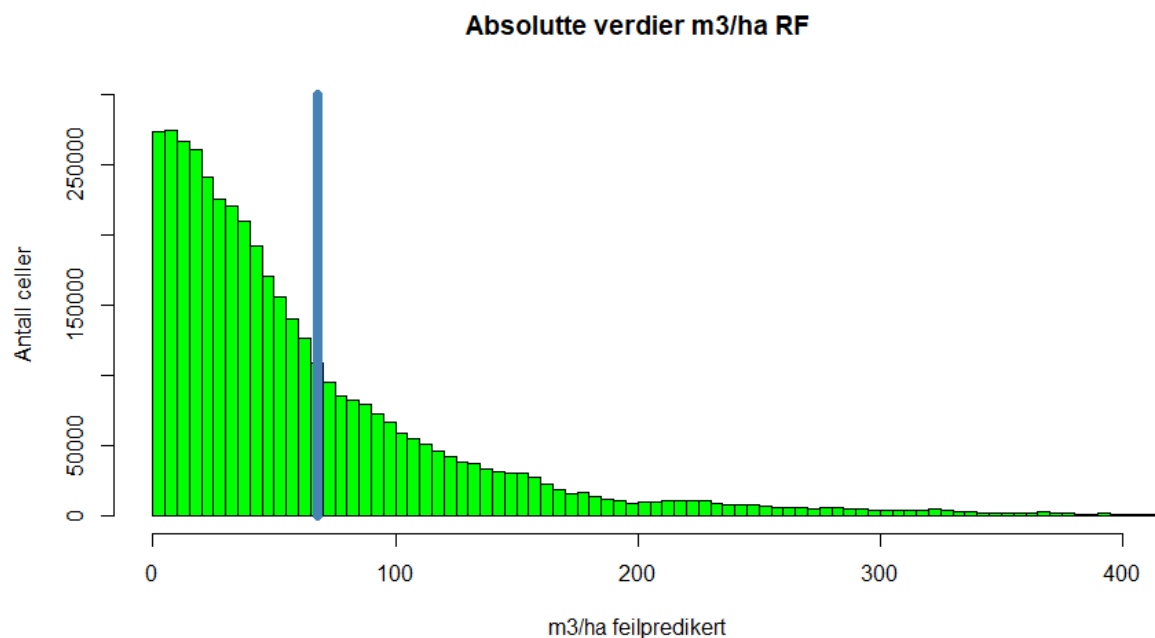
#### 3.1.1 Random Forest

Av metodene brukt i denne oppgaven fikk RF laveste verdi i gjennomsnittlig RMSE% på 38,9% (Figur 7), som er 6,8% lavere enn gjennomsnittet til LM som er nærmest RF i RMSE%. Fra Figur 7 kan vi se at RF har en vesentlig lavere RMSE% enn de to andre modelleringsteknikkene Knn og LM. For absolutte verdier som i Figur 8 sammenlignes med Figur 9 og 10, har RF den laveste feil i inventeringen for gjennomsnittlige absolutte verdier, med 68 m<sup>3</sup>/ha fra sant volum. Dette er 16 m<sup>3</sup>/Ha lavere enn LM, som er nærmest RF i absolutte verdier. Når Figurene 8, 9 og 10 sammenlignes er det tydelig å at RF har flest celler som det blir svært liten feil i m<sup>3</sup>/ha. RF har også den minste maks feil i predikert volum som er på 991 m<sup>3</sup>/ha som er hele 179 m<sup>3</sup>/ha lavere enn de andre modelleringsteknikkene. RF fikk også det laveste standardavviket på 104 m<sup>3</sup>/ha som er 18 m<sup>3</sup>/ha, lavere enn laveste av de andre modelleringsteknikkene. RF fikk beste resultat på alle målte variabler utenom en, som var gjennomsnittlig feil m<sup>3</sup>/ha hvor RF fikk -3,5 m<sup>3</sup>/ha som var laveste verdi av modelleringsteknikkene, men LM fikk -0,57 m<sup>3</sup>/ha som er nærmere null og dermed bedre gjennomsnittsverdi.





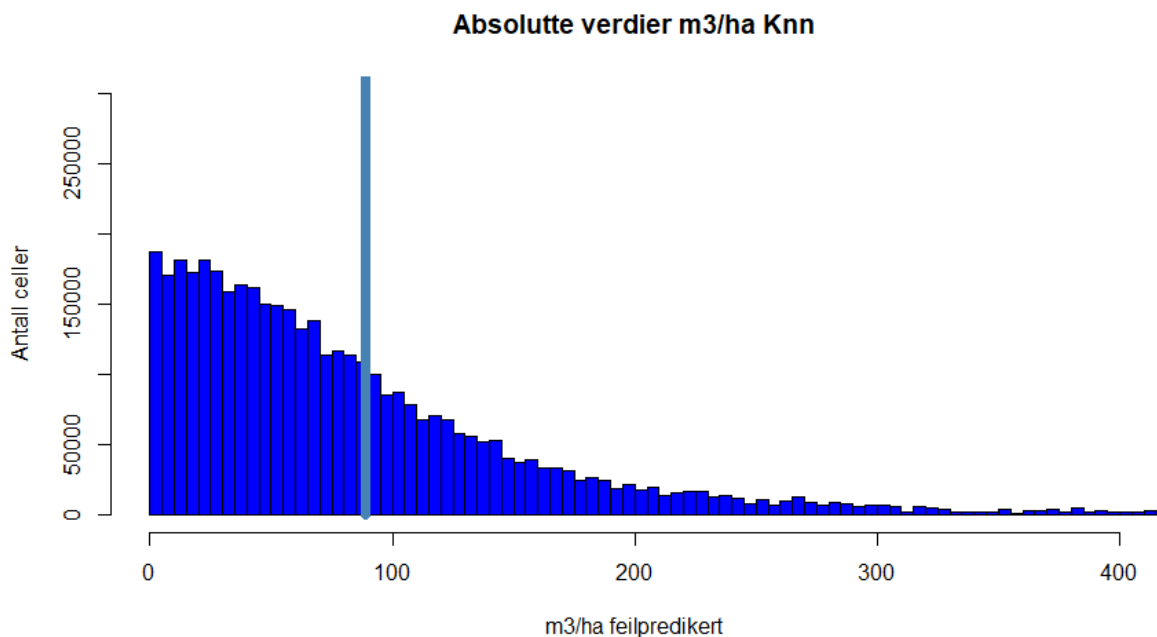
Figur 7: RMSE% av volumpredikering på cellenivå basert på 1 000 iterasjoner. Gjennomsnittlig RMSE% er presentert med blå linje. Gjennomsnittlig RMSE% ble 38,9% (RF), 47,9% (Knn) og 45,7% for (LM)



Figur 8. Absolutte verdier for feil for predikert volum i  $m^3/ha$  predikert med RF. Blå linje representerer gjennomsnittlig feil i predikert volum og er på  $67,9 m^3$ .

### 3.1.2 K-Nearest neighbor

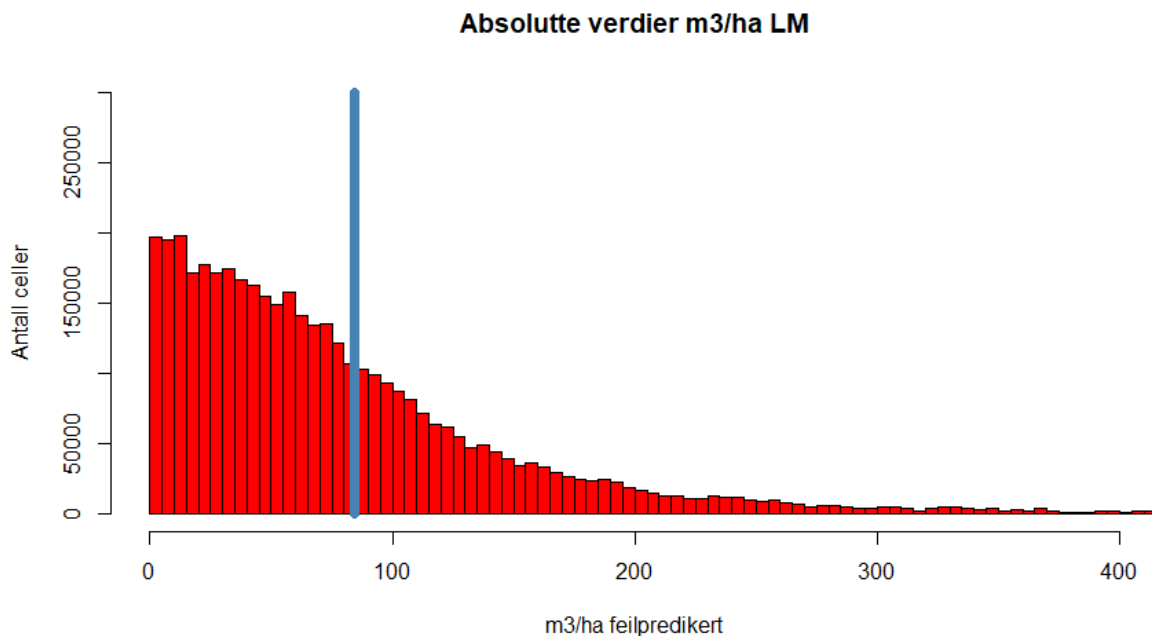
Av metodene brukt i denne oppgaven, har Knn metoden høyeste verdi RMSE% på 47,9% (Figur 7). Selv om den var høyest, var den ikke mer enn 2,2 % høyere enn med LM (Figur 7). I absolutte verdier predikert med Knn ble gjennomsnittet  $88 \text{ m}^3/\text{ha}$  (Figur 9) fra sant volum. Med et standardavvik på  $129 \text{ m}^3/\text{ha}$ , som er høyest av modelleringsteknikkene, men ganske nærme LM, og har et standardavvik som kun er  $7 \text{ m}^3/\text{ha}$  høyere enn for LM. Ved å sammenligne Figur 9 med 8 og 10 er det tydelig at Knn har en flatere kurve en RF i Figur 8. Knn har høyeste gjennomsnittlige feil, og var den eneste som ble positiv med  $6,1 \text{ m}^3/\text{ha}$ . Maksimalt feil volum for Knn ble på  $1171 \text{ m}^3/\text{ha}$ . Lavest RMSE% ble det når  $K = 5$ .



Figur 9. Absolutte verdier for feil i prediksjonen oppgitt i  $\text{m}^3/\text{ha}$  predikert med Knn. Blå linje representerer gjennomsnittlig feil volum som er  $88 \text{ m}^3/\text{ha}$ .

### 3.1.3 Lineær modellering

Av metodene brukt i denne oppgaven så er resultatene for LM mellom resultatene til RF og Knn for de fleste observasjoner. LM fikk en RMSE% på 45,7% (Figur 7) som er litt lavere enn for Knn som er nærmest LM. For absolutte verdier kan man se fra Figur 10 at LM også har et gjennomsnittlig feil i predikert volum på 84 m<sup>3</sup>/ha, som kun er 4 m<sup>3</sup>/ha lavere enn med Knn. Fra Figur 10 kan man også se at Figuren ligner på Figur 9 for Knn. Maks feil i predikert volum ble 1253 m<sup>3</sup>/ha som er høyeste observerte feilinventering i denne oppgaven (Tabell 3). LM fikk et standardavvik på 122 m<sup>3</sup>/ha som var kun 7 m<sup>3</sup>/ha lavere enn Knn (Tabell 3). LM fikk beste resultat på gjennomsnittlig feil i m<sup>3</sup>/ha som var på – 0,57 m<sup>3</sup>/ha.



Figur 10. Absolutte verdier for feil i predikert m<sup>3</sup>/ha predikert med LM. Blå linje representerer gjennomsnittlig feil i predikert volum og er på 84 m<sup>3</sup>/ha.

Tabell 3: Statistikk om totalvolum i m<sup>3</sup>/ha.

| Prediksjoner m <sup>3</sup> /ha          | RF    |  | Knn  | LM    |
|--|-------|--|------|-------|
| Standardavvik                            | 104   |  | 129  | 122   |
| Gjennomsnittlig feil                     | -3,5  |  | 6,11 | -0,57 |
| Maksimalt feil for modelleringsteknikken | 991.6 |  | 1171 | 1253  |

Tabell 3: Standardavvik, gjennomsnitt og maksimale verdier til volum predikert med metodene RF, Knn og LM. Alle verdier er oppgitt i m<sup>3</sup>/ha. Alle oppgitte verdier i tabellen er regnet ut fra alle inventert celler fra 1000 iterasjonene som ble kjørt og består dermed av 4 220 000 forsøk.

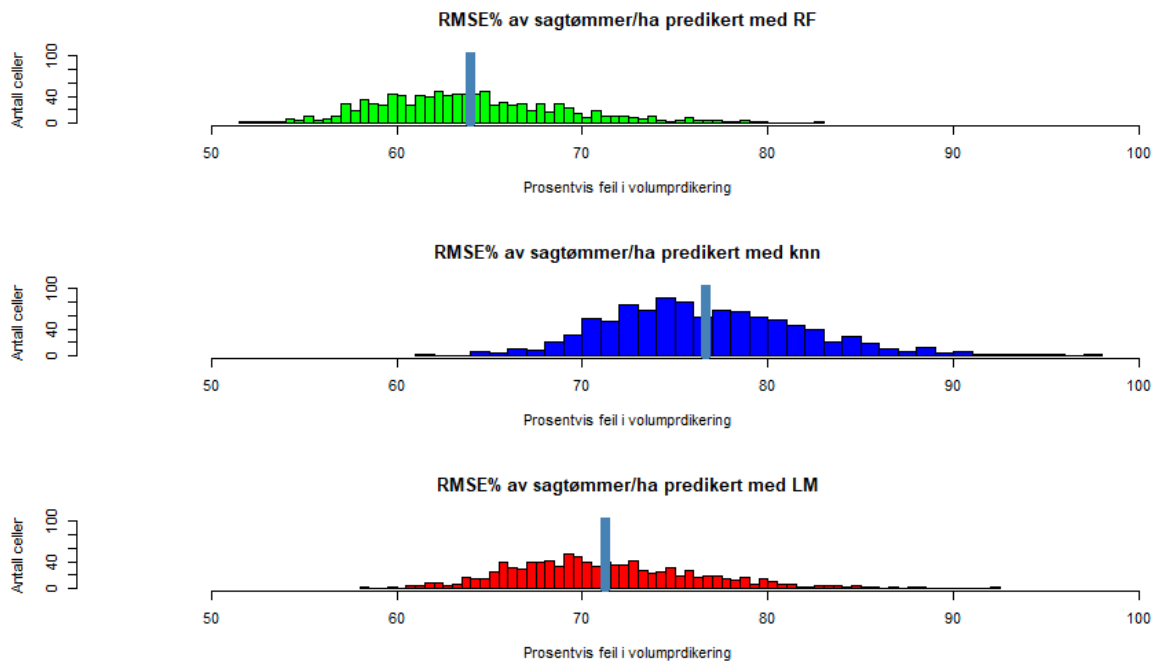
Etter hvert som scriptet ble kjørt flere ganger, forandret verdiene til max og minimum seg ganske mye, mens gjennomsnittlige verdier var stabile og forandret seg mindre en 1% mellom rundene.

## 3.2 Sagtømmerevolum

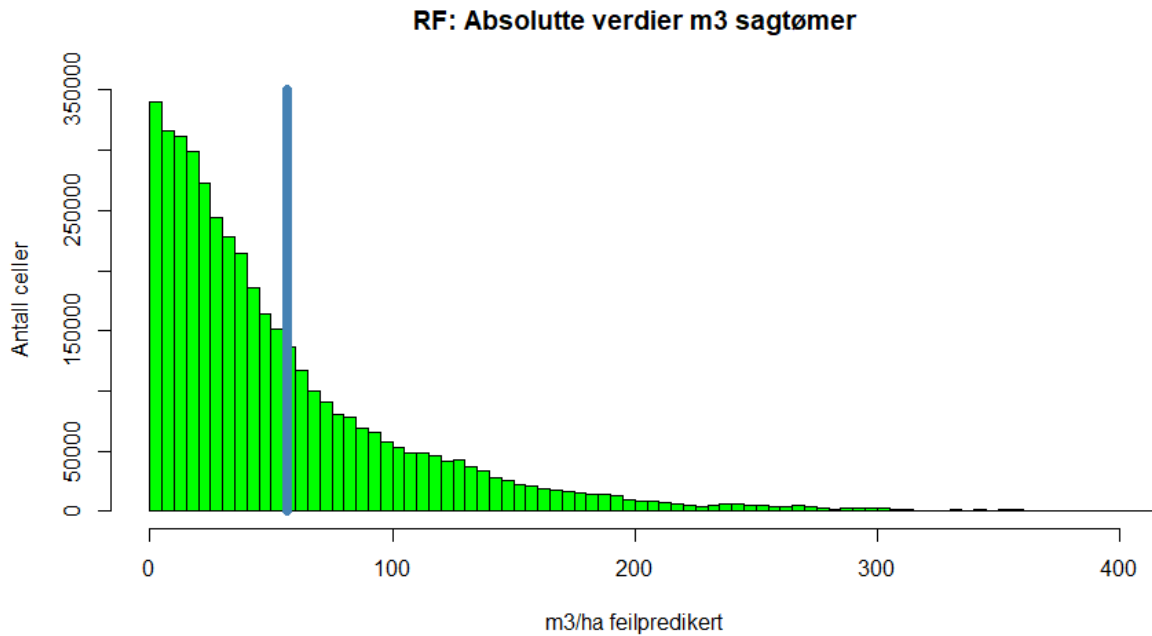
### 3.2.1 Random Forest

Av metodene som er brukt i denne oppgaven så fikk RF laveste verdien i gjennomsnittlig RMSE% også for sagtømmer som er på 64% (Figur 11). Dette er 7 % lavere enn gjennomsnittet til LM som er nærmest RF i RMSE%. Fra Figur 11 kan man se at RF har en vesentlig lavere RMSE% enn de to andre modelleringsteknikkene Knn og LM. RF har laveste feilinventering i gjennomsnittlige absolutte verdier, med 56 m<sup>3</sup>/ha gjennomsnittlig fra sant volum. RF har også minste maksimalt feil i predikert volum på 1 004 m<sup>3</sup>/ha (Tabell 4), men er kun 5 m<sup>3</sup>/ha bedre enn Knn som er lavest av de andre modelleringsteknikkene (Tabell 4). RF fikk laveste standardavvik med 84 m<sup>3</sup>/ha. RF fikk beste resultat på alle målte variabler utenom en, som var gjennomsnittlig predikert m<sup>3</sup>/ha hvor RF predikerte med -2,2 m<sup>3</sup>/ha som var

laveste verdi av modelleringsteknikkene. LM fikk  $-0,08 \text{ m}^3/\text{ha}$  som er nærmere null og dermed bedre gjennomsnittsverdi.



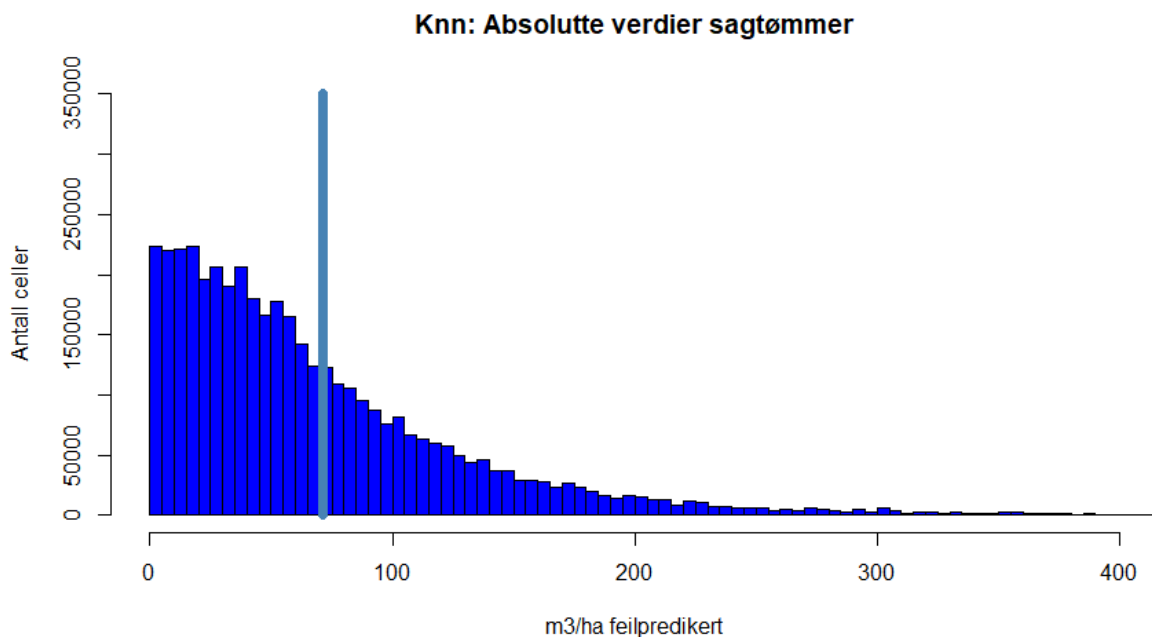
Figur 11: RMSE% av sagtømmervolumpredikering på cellenivå basert på 1 000 iterasjoner. Gjennomsnittlig RMSE% er presentert med blå linje. Gjennomsnittlig RMSE% ble 64% (RF), 77% (Knn) og 71% for (LM)



*Figur 12: Absolutte verdier for m<sup>3</sup>/ha og feil i predikert m<sup>3</sup>/ha blå strek er gjennomsnittlig feil for predikert sagtømmervolum, som er 56 m<sup>3</sup>/ha.*

### 3.2.2 K-Nearest neighbor

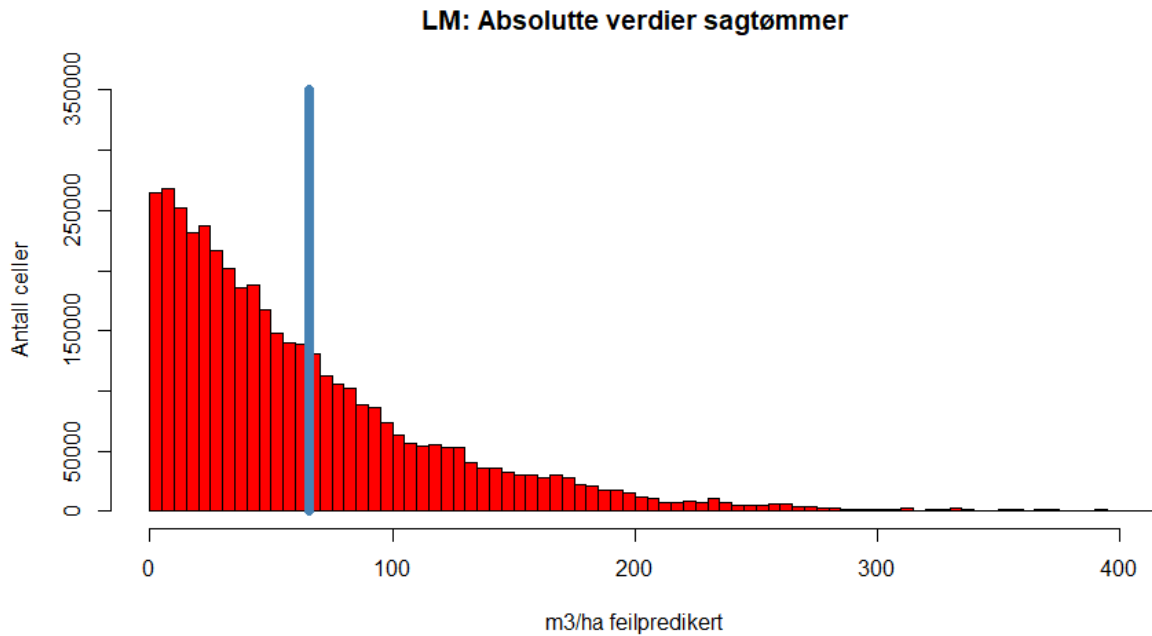
Av metodene som ble brukt i denne oppgaven hadde Knn metoden høyeste RMSE% som var på 77% (Figur 11). Selv om den var høyest var den ikke mer enn 6% høyere enn med LM (Figur 11). I absolutte verdier inventerte Knn gjennomsnittlig 71 m<sup>3</sup>/ha (Figur 13) feil med et standardavvik på 101 m<sup>3</sup>/ha (Tabell 4), som er høyest av modelleringsteknikkene, som er 8 m<sup>3</sup>/ha høyere enn for LM. Dette er nærmest av de andre modelleringsteknikkene. Knn var den eneste som ble positiv med 1,68 m<sup>3</sup>/ha gjennomsnittlig feil volum (Tabell 4). Maksimalt feil volum for Knn ble på 1 009 m<sup>3</sup>/ha, som kun er 5 m<sup>3</sup>/ha lavere enn RF.



*Figur 13. Absolutte verdier for feil i predikert sagtømmervolum i m<sup>3</sup>/ha predikert med Knn. Blå linje representerer gjennomsnittlig feil i predikert volum som er på 71 m<sup>3</sup>/ha.*

### 3.2.3 Lineær modellering

Av metodene brukt i denne oppgaven er resultatene for LM mellom resultatene til RF og Knn for de fleste observasjoner, likt som for totalvolum. LM fikk en RMSE% på 71% (Figur 11) som er 6% lavere enn for Knn. For absolutte verdier kan man se fra Figur 14 at LM fikk gjennomsnittlig feil i volum på 66 m<sup>3</sup>/ha, som er 5 m<sup>3</sup>/ha lavere enn med Knn. Maks feil i inventert sagtømmervolum ble 1 040 m<sup>3</sup>/ha, som er høyeste observerte verdi i denne oppgaven (Tabell 4). LM fikk et standardavvik på 93 m<sup>3</sup>/ha, kun 8 m<sup>3</sup>/ha lavere enn Knn (Tabell 4). LM fikk beste resultat på gjennomsnittlig feil i m<sup>3</sup>/ha som var på – 0,08 m<sup>3</sup>/ha.



Figur 14. Absolutte verdier for feil i predikert sagtømmervolum i m<sup>3</sup>/ha predikert med LM. Blå linje representerer gjennomsnittlig feil i predikert volum som er på 66 m<sup>3</sup>/ha.

Tabell 4. Statistikk om sagtømmervolum

| Prediksjoner m <sup>3</sup> /ha         | RF    | Knn   | LM    |
|---|-------|-------|-------|
| Standardavvik                           | 84    | 101   | 93    |
| Gjennomsnittlig feil                    | -2,24 | 1,68  | -0,08 |
| Maksimal feil for modelleringsteknikken | 1 004 | 1 009 | 1040  |

Tabell 4: Standardavvik, gjennomsnitt og maksimale verdier i feil for modelleringsteknikken sagtømmervolum predikert med metodene RF, Knn og LM. Alle verdier er oppgitt i m<sup>3</sup>/ha. Alle oppgitte verdier i Tabellen er regnet ut fra alle inventerte celler fra 1000 iterasjoner som ble kjørt og består dermed av 4 220 000 forsøk.



## 4.Diskusjon

Hovedmålet i denne oppgaven var å undersøke om hogstmaskindata fra hogstmaskin i kommersiell drift kan brukes som referansedata for å predikere totalvolum og sagtømmervolum for geografien maskinen arbeider i. I denne oppgaven ble det brukt EABA og random subsampling cross validation med modelleringsteknikkene RF, Knn og LM for å predikere totalvolum og sagtømmervolum på cellenivå med hogstmaskindata som referansedata. Resultatene fra denne oppgaven viser at det er mulig å inventere med lav gjennomsnittlig feil. Laveste gjennomsnitt i denne oppgaven ble  $-0,08 \text{ m}^3/\text{ha}$  for sagtømmervolum predikert med LM. Som nevnt i innledningen er tilgangen til fasitvolumet til predikerte celler for så store områder en unik egenskap, ved bruk av hogstmaskindata som referansedata.

### 4.1 Hogstmaskindata

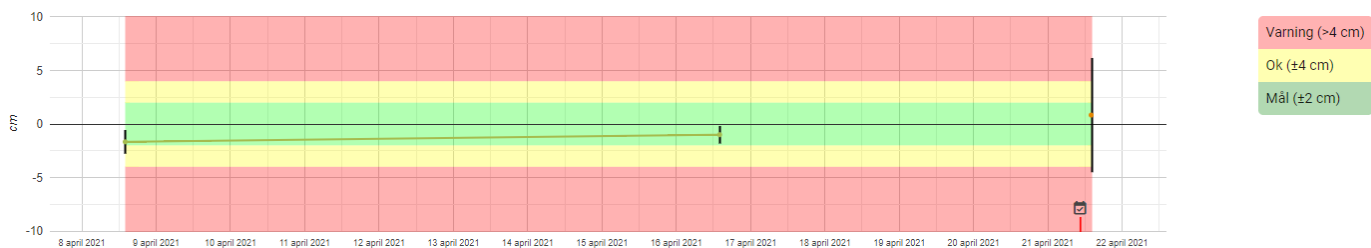
Sammenlignet med studiene til Hauglin et al. (2018) og Maltamo et al. (2019) som også monterte ekstra GNSS-antenner og sensorer, ble det enklere med denne maskinen da det kun var behov for en ekstra sensor og ikke syv som det var på den andre. Noe av dette kommer fra valg av maskintype og plassering av GNSS, men også at sensorer for å ha kjennskap til kranspissposisjonen blir mer vanlig da det kan brukes for å hjelpe førerne med å øke produksjonen (Lindroos et al., 2015). Noe som burde gjøre det enklere å få bedre posisjonerte hogstmaskindata i fremtiden. Selv om denne oppgaven, med andre studier av Hauglin et al. (2018), Maltamo et al. (2019) og masteroppgaven til Dieset (2017) viser at det er mulig å samle inn og bruke hogstmaskindata til skogresurskratlegging med tilsvarende feil som vanlig manuelle feltmålinger (Maltamo et al., 2019). Er det ikke sikkert at kostnaden ved å forbedre hogstmaskinen posisjoneringsutstyr blir dekket av sparte utgifter som kommer med å bruke hogstmaskindata som referansedata til skogbruksplaner laget med FLS alene.

I denne oppgaven det flest treantall fra FLS dataene som kommer fra 2013, som vil si det gått 5 til 7 år mellom FLS til hogst. Dette indikerer at en punktsky over større område kan brukes til å til å inventere skog i flere år etter innsamling. Og at det kan brukes til å predikere volum på mindre områder med lav gjennomsnittlig feil ved behov.

### 4.1.1 Hogstmaskindata

Hogstmaskindata blir behandlet og beskrevet som fasitvolum er det flere faktorer som kan påvirke hogstmaskinens målinger. Det er viktig at maskinens utstyr er kalibrert etter maskinleverandørens spesifikasjoner (Nieuwenhuis & Dooley, 2006; Noordermeer et al., 2021) For volumpredikering gjelder dette maskinens lengde, og diametermåling. Temperatur og trykk vil påvirke maskinens målinger; derfor er det viktig med jevnlige kalibreringer av maskinens utstyr for å unngå systematiske feil i volummålingene.

Fra Figur 15 kan vi se maskinens lengdekalibreringer for en periode på 22 dager i april. Sent i april ble det varmere vær som førte til at lengdene på målingene ble endert. Varmere vær fører til at barken og stammen «tinte» opp, dette gjør at målehjulet går dypere i barken. Når målehjulet går dypere blir lengdemålingene kortere, ettersom målehjulet går en runde, litt kortere på stammen. Derfor kan man se at kontrollmålingene i slutten av april er lengre enn for starten av april (Figur 15) og at standardavviket for lengdemålingene øker.



*Figur 15: Lengdekalibrering for april måned på hogstmaskinen som har skaffet referansedataene. På x-aksen er datoene 8 -22 april og på y-aksen er det antall cm forskjell på maskinmålt og kontrollmålte lengder. Målingene er gjennomsnittsmålingen for lengden som ble målt og den svarte diagonale streken er standardavviket.*

Hogstmaskinen kan jobbe med forskjellige volummålingsformler enn Formel 4 og 5 som er nevnt tidligere i denne oppgaven. Formlene kan variere mellom drifter, tømmerkjøperorganisasjoner og land. I denne oppgaven har jeg valgt å fokusere lite på eventuelle feilkilder i hogstmaskinens volummålinger eller utregninger. Jeg har antatt at det vil påvirke resultatet lite, og er en tilfeldig feil. Men, det er viktig å være bevisst på at det kan være en feilkilde.

### 4.1.2 Hogstmaskindata som referansedata

Resultatene fra denne oppgaven har en høy RMSE% på cellenivå. Det er flere feilkilder som kan være med å forklare resultatene for RMSE%. Bedre nøyaktighet av aggregatposisjon forventes å være den faktoren som vil forbedre RMSE% mest. Men også feil i referansedataene fra hogstmaskinens HPR-fil, som vil ha lagret flere trær, enn antallet trær som har blitt hogd. Maskinen registrerer nytt tre for hvert fellekapp. Kapp registreres som fellekapp dersom det ikke er registrert tre i hogstmaskinens dataprogram, eller ved første kapp etter treregistrering fra fører. Ved hogst av trær som har dobbelttopp og begge toppene blir aptert hver for seg vil det være to fellingskapp registrert, men kun en stubbe i skogen og kun punktsky av et tre. Toppen som blir aptert som tre nummer to får en helt annen posisjon enn dens egentlige posisjon.

Toppbrekk vil også føre til dobbeltregistrering av trær da toppen som knekker, ofte vil ligge utenfor arbeidsområdet til maskinen, og ikke bli hogd før den kommer inn i maskinens arbeidsområde igjen. Følgende vil høyde og volum bli for lite på rotstammen og det vil bli registrert en topp med volum hvor det ikke skulle vært noen trær. Feilkilden fra dobbeltregistrerte trær vil kunne reduseres med å implementere en funksjon som gjør at trær som ikke blir målt til en dimensjon lik minste massevirkedimensjon, søker etter trær senere i datasettet. Med dimensjon lik eller litt mindre siste kapp på det aktuelle treet.

Hogstmaskindata har målinger langs hele stammen, og volumberegninger som oppgjørsvolum til skogeier er en styrke med referansedata fra hogstmaskindata. Merk at denne oppgaven kun har predikert oppgjørsvolum ved kjøp og salg av rundvirke. Dersom det eksempelvis skal predikeres karbon bundet i skog, død ved, totalvolum inkludert bultinger og topper eksempelvis, er det ikke sikkert at det blir like resultat som i denne oppgaven.

## 4.2 Celler

Valg av cellestørrelse påvirker resultatene til volumpredikeringen (Dieset, 2017; Maltamo et al., 2019). I denne oppgaven ble det brukt 200 m<sup>2</sup> celler grunnet funn fra Maltamo et al. (2019) som sammenlignet flere cellestørrelser, og fikk gode resultater med små forskjeller for cellene som var på 200 m<sup>2</sup> og 400 m<sup>2</sup> (Maltamo et al., 2019). Dieset (2017) fikk også gode resultater for 200 m<sup>2</sup> celler. Store celler fører til mer gjennomsnittlige verdier (Maltamo et al.,

2014). Ved bruk av små celler kan det enklere forekomme store feil da kanteffektene og feil i aggregatposisjoneringen vil føre til dårlige resultat for volumpredikering (Dieset, 2017).

Selv om metoden EABA har blitt brukt., er kanteffekten redusert, men ikke borte (Pascual, 2019). Aggregatposisjoneringen klarer å lokalisere stubben med posisjon med en meter nøyaktighet som anbefalt av Lindroos et al. (2015). Vil det være flere trær som blir posisjonert i nabocellen. Som er en av årsakene til de ekstreme verdiene som blir vist i Figur 6. Ytterligere forbedring av posisjoneringsutstyret vil mest sannsynlig redusere feilene i dataene. Trær som sættes igjen av miljømessige eller skogfaglige grunner vil heller ikke bli måles og kan bli en feilkilde.

### **4.3 Modelleringssteknikker**

Det har ikke blitt ekskludert noen i denne oppgaven, dermed vil det være flere tilfeller av at celler har blitt predikert med verdier fra nabocellene. Dette gjelder alle modelleringssteknikkene, men er spesielt viktig for LM, da bruk av LM innebærer antagelsen om at utvalget er statistiske uavhengige. Dette kravet er ikke oppfylt da cellene som er ved siden av hverandre påvirker hverandre direkte. Selv om utvalget ikke er statiske uavhengige er det antatt at det vil påvirke resultatene lite. Gjennom metoden Random subsampling cross validation velges cellene tilfeldig fra de 4216 cellene. Dette medfører at påvirkningen som kommer fra nabocellene, blir liten. Hauglin et al. (2018) beskrev også at det var en romlig korrelasjon på stammevolumet på en avstand på opptil 200 meter. Ved inventering av skogressurser er det viktigere å ha gode modeller som også fungerer godt utenfor det oppsamlede området, enn at utvalget er statistisk uavhengige (Hauglin et al., 2018). Ved bruk av et så bredt datasett som det er i denne oppgaven, er det forventet at modellen vil fungere godt. Og at den romlige korrelasjonen har lite påvirkning.

Ved å sammenligne RMSE% for LM med de andre modelleringssteknikkene RG og KNN i denne oppgaven, med RMSE% for LM sammenlignet med RF og Knn i Hauglin et al. (2018), så er RMSE% for LM bedre i denne oppgaven. At RMSE% er lavere i denne oppgaven, relativt til de andre modelleringssteknikkene. Kan komme fra celler i sammen bestand er med i trening og valideringsdatasettet med metoden Random subsampling cross validation, mens Hauglin et al. (2018) brukte fivefold cross validation.

Resultatene indikerer at RF skiller seg ut som den beste predikasjonsmetoden i dette studiet. Hauglin et al. (2018) fikk også laveste RMSE% med RF. RF er en vanlig modelleringsteknikk (Breiman, 2001). En fordel ved bruk av RF er at det kan brukes et stort antall uavhengige variabler selv om de er i korrelasjon med hverandre.

Av metodene som ble brukt i denne oppgaven er LM den eneste metoden som ikke er parameterfri. Resultatene til LM ligger mellom resultatene til RF og Knn for samtlige resultat, utenom maks feil av predikert volum som LM ligger 82 m<sup>3</sup>/ha høyere enn Knn. Forskjellen anses å være liten da maksimalt volum er på 1 253 m<sup>3</sup>/ha med LM. Maks feil av predikert volum er lite viktig måleparameter, men den er inkludert i denne oppgaven for å se hvilken målemetode som kom til å gjøre den mest ekstreme predikasjonen. Som nevnt i er LM mest tilbøyelig til å komme med ekstreme prediksjoner som predikerer utenfor datasettet, noe maskinlæringsmetoder ikke vil gjøre og blir støttet av resultatene i denne oppgaven. Selv om LM ikke predikerer verdier som er høyere enn datasettet, har LM den mest ekstreme verdien.

Det er viktig å huske at parameterne som LM brukte i sin modellering ble valgt ut av RF modellen. Når scriptet ble kjørt flere ganger ble det observert at parameterne som ble valgt ut varierte. Dette vil si at LM modellen mest sannsynlig kunne blitt forbedret hadde det blitt undersøkt mer hvilke parametere som hadde vært best å bruke for å for LM til dette datasettet. Fordelen med å gjøre det på måten som er gjort i denne oppgaven er at resultatene også vli bli en tilpasningsdyktighetsbeskrivelse for å se hvilke av metodene som best tilpasset store datamengder av ulik kvalitet, alder, og egenskaper.

#### **4.4 Totalvolum**

Selv om jeg nå har pekt på flere svakheter ved oppgaven viser resultatene at det er mulig å bruke hogstmaskindata som referansedata til FLS. Dette stemmer med tidligere resultater fra Hauglin et al. (2018) og Maltamo et al. (2019).

Resultatene er oppgitt i m<sup>3</sup>/ha og kan virke mye, er det viktig å forstå at volumet er regnet ut i m<sup>3</sup>/ha og cellene er på omtrent 200 m<sup>2</sup> (Figur 5), Det vil si det trengs 50 celler før det blir en hektar. Resultatene styrker påstanden om at prediksjonsfeilene vil utjevne hverandre ved at gjennomsnittlig predikert volum fra sant volum - 3,5 m<sup>3</sup>/ha (RF), 6,1 (Knn) og - 0,57 med

LM. til  $6,1 \text{ m}^3/\text{ha}$  for totalvolum. Dette anses som svært lavt og indikerer at hogstmaskindata kan brukes som referansedata til skoginventering ved FLS.

Resultatene om gjennomsnittlig prediksjonsfeil for kan sees i sammenheng med resultatene til Maltamo et al. (2019). Det er flere forskjeller på studiet til Maltamo et al. (2019) og denne oppgaven. Blant annet inventerte Maltamo et al. (2019) på bestandsnivå, og brukte kun metoden Knn, men de fikk en RMSE på kun 9% på bestandsnivå.

Resultatene for totalvolum er på samme nivå som tidligere studier, spesielt Hauglin et al. (2018) som det vil være mest naturlig å sammenligne med. Datamaterialet er ikke det samme, men de er basert på hogstmaskindata og FLS. Hauglin et al. (2018) delte også datasettet med høy og lav bonitet, noe som ikke ble gjort i denne oppgaven. Hauglin et al. (2018) oppnådde lavere RMSE% både for RF og Knn, men ikke for LM i datasettet som var i middels boniteter. At denne oppgaven fikk bedre resultater for LM i middels boniteter kan skyldes at flere celler fra samme bestand blir brukt som treningsceller i denne oppgaven, eller EABA metoden. Forskjellen i RMSE% er ikke stor når det blir tatt i betraktning at det i denne oppgaven er flere FLS datasett, ikke skilt på middels og høye boniteter og flere år mellom FLS og hogst. Hauglin et al. (2018) brukte også  $400\text{m}^2$  celler, mens det i denne oppgaven er brukt  $200 \text{ m}^2$  celler som vil gi større varians. I tillegg ble datasettet i Hauglin et al. (2018) delt opp i gode og middels boniteter, noe som ikke ble gjort i denne oppgaven, og vil også være en del av forklaringen på hvorfor RMSE% er høyere i denne oppgaven.

#### **4.5 Sagtømmervolum**

Som nevnt innledningsvis så har definisjonen av sagtømmer har forandret seg i løpet av innsamlingsperioden og hogstmaskinen har hogd for forskjellige tømmerkjøpere som leverer til forskjellige sagbruk og industriforbrukere som har varierende krav til kvaliteten og dimensjon. Dette kan svekke resultatene noe, men er også en representasjon for at markedet av rundvirke er skiftende i etterspørselen, noe det også er forventet å være i fremtiden.

Resultatene i absolutte verdier, standardavvik og gjennomsnitt for sagtømmervolum er svært like resultatene for totalvolum, bare noe mindre som og vil være naturlig da det er mindre sagtømmervolumet enn totalvolum. Resultatene for gjennomsnittlig RMSE% varierte fra 64% (RF), 77% (Knn) og 71% (LM). Som er vesentlig høyere enn det Dieset (2017) oppnådde,

som at sagtømmerandel kunne estimeres med cirka 30%. Det er flere forskjeller på disse oppgavene. I denne oppgaven er en større spredning i datasettet som slår ut mye på prosentvis feil. Viktigste forskjellen er at det i denne oppgaven er predikert sagtømmer volum, og ikke sagtømmerandel som det i oppgaven til Dieset (2017). Sagtømmervolum er ett bedre estimat, og gir mer kunnskap om sagtømmervolumet som vil være vesentlig for å valg av bestand til hogst. Resultatene for denne oppgaven med lav gjennomsnittlig feil på  $-2,24 \text{ m}^3/\text{ha}$  (RF), 1,68 (Knn), og  $-0,08$  for LM er å anse som svert lavt. Dette tyder på at nøyaktige bestandsestimat for sagtømmer kan oppnås med bruk av hogstmaskindata som referansedata til skoginventering ved FLS. Resultatene vier også at sagtømmerprosenten kan være en dårligere indikator i forhold til sagtømmervolum.

## 5. Konklusjon

Denne oppgaven har vist at det er mulig å predikere totalvolum og sagtømmervolum med hogstmaskindata som referansedata til FLS. Resultatene tyder på at nøyaktige hogstmaskindata kan gi gode bestandsestimat for volumpredikering til skogbruksplaner og enkeltbestand. I geografien maskinen operer i med lav gjennomsnittlig feil.

Modelleringsteknikken som fungerte best i denne oppgaven var RF. Resultatene fra denne oppgaven viser at modelleringsteknikken RF vil gi laveste standardavvik og RMSE% på cellenivå både for sagtømmer og totalvolum, RF var modelleringsteknikken som ga best resultat sammenlagt i denne oppgaven.



## Referanser

- Andreassen, K., Allen, M., Hanssen, K. H., Kvaalen, H. & Støtvig, S. (2018). Langsiktig feltforsøk i skog ved NIBIO. *NIBIO Rapport*.
- Barth, A. & Holmgren, J. (2013). Stem taper estimates based on airborne laser scanning and cut-to-length harvester measurements for pre-harvest planning. *International Journal of Forest Engineering*, 24 (3): 161-169. doi: 10.1080/14942119.2013.858911.
- Breiman, L. (2001). *Machine Learning*, 45 (1): 5-32. doi: 10.1023/a:1010933404324.
- Dalen, L. S. (2017). *Nye rekordtall for skogen i Norge*. Tilgjengelig fra: <https://www.nibio.no/nyheter/nye-rekordtall-for-skogen-i-norge> (lest 19.05.2021).
- Dalponete, M. (2016). *R Package `itcSegment`: User Manual*.
- Dieset, T. (2017). *Estimering av sagtømmerandel ved hjelp av flybåren laserskanning*: Norwegian University of Life Sciences, Ås.
- Eid, T., Gobakken, T. & Næsset, E. (2004). Comparing stand inventories for large areas based on photo-interpretation and laser scanning by means of cost-plus-loss analyses. *Scandinavian Journal of Forest Research*, 19 (6): 512-523. doi: 10.1080/02827580410019463.
- Evans, J., Hudak, A., Faux, R. & Smith, A. M. (2009). Discrete Return Lidar in Natural Resources: Recommendations for Project Planning, Data Processing, and Deliverables. *Remote Sensing*, 1 (4): 776-794. doi: 10.3390/rs1040776.
- FUNT. (1997). *Volumtabeller i basis virkesystemer fra 01.01.98*: Tømmermålingsforeningens Fellesorgan.
- Hauglin, M., Hansen, E., Sørngård, E., Næsset, E. & Gobakken, T. (2018). Utilizing accurately positioned harvester data: modelling forest volume with airborne laser scanning. *Canadian Journal of Forest Research*, 48 (8): 913-922. doi: 10.1139/cjfr-2017-0467.
- Johan J. Moller, J. A., . Maria Nordstrom. (2013). *Test av StanForD 2010 implementation i skordare. StanForD 2010- Implementation and test of harvester. . Skogforsk*
- Kemmerer, J. & Labelle, E. R. (2021). Using harvester data from on-board computers: a review of key findings, opportunities and challenges. *European Journal of Forest Research*, 140 (1): 1-17. doi: 10.1007/s10342-020-01313-4.
- Koivuniemi, J. & Korhonen, K. T. (2006). Inventory by compartments. I: *Forest Inventory*, s. 271-278: Springer.
- Landbruksdirektoratet. (2021). *Tømmeravirkning og priser* Tilgjengelig fra: [https://www.landbruksdirektoratet.no/nb/statistikk-og-utviklingstrekk/utviklingstrekk-i-skogbruket/tommeravirkning-og-priser?fbclid=IwAR3NZitj5tlb-PWdpPRhyiP5nFzpLoyV3NkRxwc4RUroz\\_Yi4Y1PXYSkJM](https://www.landbruksdirektoratet.no/nb/statistikk-og-utviklingstrekk/utviklingstrekk-i-skogbruket/tommeravirkning-og-priser?fbclid=IwAR3NZitj5tlb-PWdpPRhyiP5nFzpLoyV3NkRxwc4RUroz_Yi4Y1PXYSkJM) (lest 19.05.2021).
- Leksikon, S. N. (2018). *Aptere*. Tilgjengelig fra: <https://snl.no/aptere> (lest 19.05.2021).
- Leksikon, S. N. (2019). *Stereofotogrammetri*. Tilgjengelig fra: <https://snl.no/stereofotogrammetri> (lest 19.05.2021).

- Lindroos, O., Ringdahl, O., La Hera, P., Hohnloser, P. & Hellström, T. H. (2015). Estimating the position of the harvester head—a key step towards the precision forestry of the future? *Croatian Journal of Forest Engineering: Journal for Theory and Application of Forestry Engineering*, 36 (2): 147-164.
- Maltamo, M., Næsset, E. & Vauhkonen, J. (2014). Forestry applications of airborne laser scanning. *Concepts and case studies. Manag For Ecosys*, 27: 2014.
- Maltamo, M., Hauglin, M., Naesset, E. & Gobakken, T. (2019). Estimating stand level stem diameter distribution utilizing harvester data and airborne laser scanning. *Silva Fennica*, 53 (3). doi: 10.14214/sf.10075.
- Naesset, E. (1997a). Determination of mean tree height of forest stands using airborne laser scanner data. *ISPRS journal of photogrammetry and remote sensing*, 52 (2): 49-56. doi: 10.1016/S0924-2716(97)83000-6.
- Naesset, E. (1997b). Estimating timber volume of forest stands using airborne laser scanner data. *Remote sensing of environment*, 61 (2): 246-253. doi: 10.1016/S0034-4257(97)00041-2.
- Nieuwenhuis, M. & Dooley, T. (2006). The Effect of Calibration on the Accuracy of Harvester Measurements. *International Journal of Forest Engineering*, 17 (2): 25-33. doi: 10.1080/14942119.2006.10702533.
- Noordermeer, L., Sørngård, E., Astrup, R., Næsset, E. & Gobakken, T. (2021). Coupling a differential global navigation satellite system to a cut-to-length harvester operating system enables precise positioning of harvested trees. *International Journal of Forest Engineering*: 1-9. doi: 10.1080/14942119.2021.1899686.
- Næsset, E. (2002). Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sensing of Environment*, 80 (1): 88-99. doi: 10.1016/s0034-4257(01)00290-5.
- Næsset, E., Gobakken, T., Holmgren, J., Hyypä, H., Hyypä, J., Maltamo, M., Nilsson, M., Olsson, H., Persson, Å. & Söderman, U. (2004). Laser scanning of forest resources: the nordic experience. *Scandinavian Journal of Forest Research*, 19 (6): 482-499. doi: 10.1080/02827580410019553.
- Næsset, E. (2014). Area-Based Inventory in Norway – From Innovation to an Operational Reality. I, s. 215-240: Springer Netherlands.
- Packalen, P., Strunk, J. L., Pitkanen, J. A., Temesgen, H. & Maltamo, M. (2015). Edge-Tree Correction for Predicting Forest Inventory Attributes Using Area-Based Approach With Airborne Laser Scanning. *IEEE journal of selected topics in applied earth observations and remote sensing*, 8 (3): 1274-1280. doi: 10.1109/JSTARS.2015.2402693.
- Pascual. (2019). Using Tree Detection Based on Airborne Laser Scanning to Improve Forest Inventory Considering Edge Effects and the Co-Registration Factor. *Remote Sensing*, 11 (22): 2675. doi: 10.3390/rs11222675.
- PEFC. (2015). *Norsk PEFC Skogstandard*. Tilgjengelig fra: <https://cdn.pefc.org/pefc.no/media/2021-01/c713871e-587a-4b0b-b367-64ea6a3ab5e0/16118998-c836-5851-aa38-eda2bed48004.pdf> (lest 30.05.2021).

- Romain, J.-. (2021). *lidR* (Versjon 3.1.3). Github. Tilgjengelig fra: <https://github.com/Jean-Romain/lidR>.
- Råheim, A. (2021). 23.05.2021.
- SSB. (2020). *Landsskogtakseringen* Tilgjengelig fra: <https://www.ssb.no/lst> (lest 19.05.2021).
- Stendahl, J. & Dahlin, B. (2002). Possibilities for Harvester-based Forest Inventory in Thinnings. *Scandinavian Journal of Forest Research*, 17 (6): 548-555. doi: 10.1080/02827580260417206.
- Vauhkonen, J., Ene, L., Gupta, S., Heinzl, J., Holmgren, J., Pitkanen, J., Solberg, S., Wang, Y., Weinacker, H., Hauglin, K. M., et al. (2012). Comparative testing of single-tree detection algorithms under different types of forest. *Forestry*, 85 (1): 27-40. doi: 10.1093/forestry/cpr051.
- Virkesmåling, N. (1995). *Reglement*. Tilgjengelig fra: <https://www.m3n.no/aktuelt/reglement/> (lest 23.05.95).
- Ørka, H. O. (2018). *lasR*. GitHub. Tilgjengelig fra: <https://github.com/hansoleorka/lasR>.



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

Postboks 5003  
NO-1432 Ås  
Norway