

## Krav til kvalitet på digitale bilde data for automatiske analyser.

Nils Erik Jørgensen i TerraNor har arbeidet med programvare for GIS og bildebehandling siden 1985. Nå arbeider TerraNor med et nytt system for skoglig kartlegging basert på teknologi utviklet av Canadian Forest Service, ved François A. Gougeon, Ph.D. Den nye teknologien registrerer alle enkelttrær i skogen og bestemmer treslag og størrelse semi-automatisk. Metoden åpner for nye store muligheter, men er samtidig krevende til kvalitet på data. Artikkelen beskriver hvorfor valg av riktig kvalitet på digitale bilde data er så viktig.

TerraNor utførte i 2006/2007 skogkartlegging prosjekt for Bergvik skog / Stora Enso, Sveaskog og Statens Fastighetsverk. Målene var skoglig kartlegging, kartlegging av skogskader (Gremeniella) og åtgjerder i ungskog. Kvalitet på digitale bilde data var bestemmende for resultat og total kostnader i alle prosjektene.

Testene omfattet nye digitale bilde data og scannede bilder fra film fra Lantmäteriet.

### Lantmäteriets nye digitale produksjon:

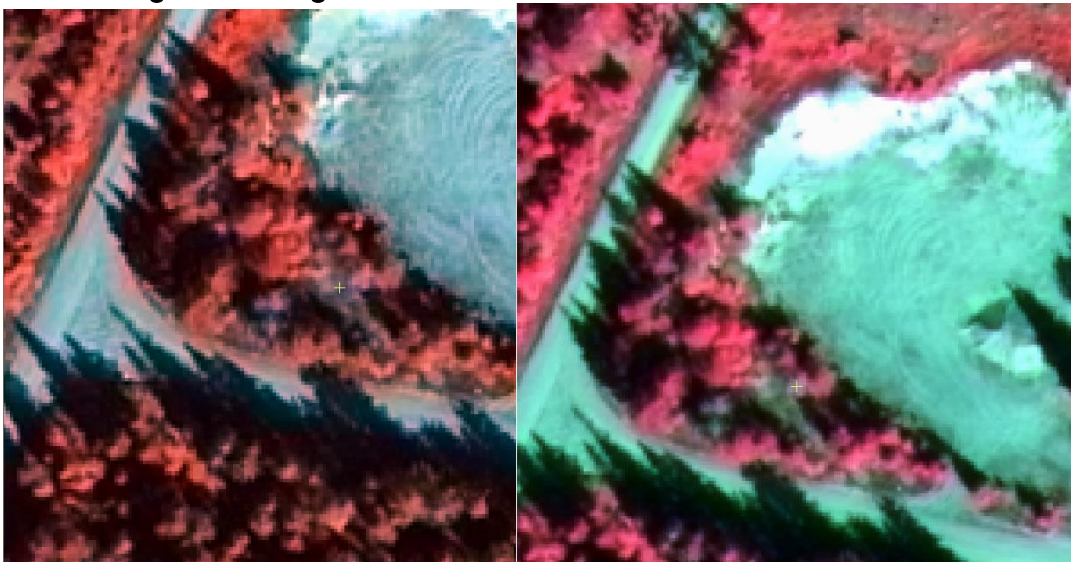
Lantmäteriet (LM) flyr 1/3 av Sverige hvert år med et nytt digitalt DMC 30 kamera/sensor fra Intergraph. DMC 30 leverer data i 16 bit format (sensoren registrerer 12 bit) i kanalene Rødt Grønt Blått (RGB), InfraRødt (IR) og gråtoner (PAN). Fargekanalene kalles Multispektrale. LM ortorektifiserer data og foretar fargejustering for å få mest mulig like farger over bildene. Data blir levert i 8 bit format tilpasset GIS. LM flyr med sensoren/kameraet i 4 800m høyde. Dette gir en oppløsning på data på 0,48 m i Pan (gråtoner) og 2,4 m i MultiSpectral (RGB + IR). Ved hjelp av spesielle behandlingsteknikker kan LM produsere 'falske' fargebilder med 0,48 m oppløsning. Målet for LM's produksjon er å produsere ortofoto til bruk i GIS.

Data fra LM's nåværende produksjonslinje er uegnet for vår digitale analyse. I tillegg til datakomprimeringen, bearbeider LM data på flere måter som vi ikke ønsker. Vi fikk derfor LM til å sende oss de opprinnelige data fra sensoren sammen med en fil med orienteringsparametere for ortorektifisering. Datasettene består av en fil med Pankromatiske og en fil med Multispektrale data for hver scene/bilde som er tatt opp. Før analyse ble data ortorektifisert til det svenske koordinatsystemet RT90 som kunden fortsatt bruker.

Som del av leveransen til kunden ble data pansharpened, mosaikkert og fargejustert. Se beskrivelse nedenfor.

Scannede IR bilder blir ikke produsert lenger. Kvalitets og produksjonsmessig gir de samme resultat som digitale data levert i 8 bit format. I våre prosjekt brukte vi scannede data fra 2004.

### Kvalitet og utfordringer:

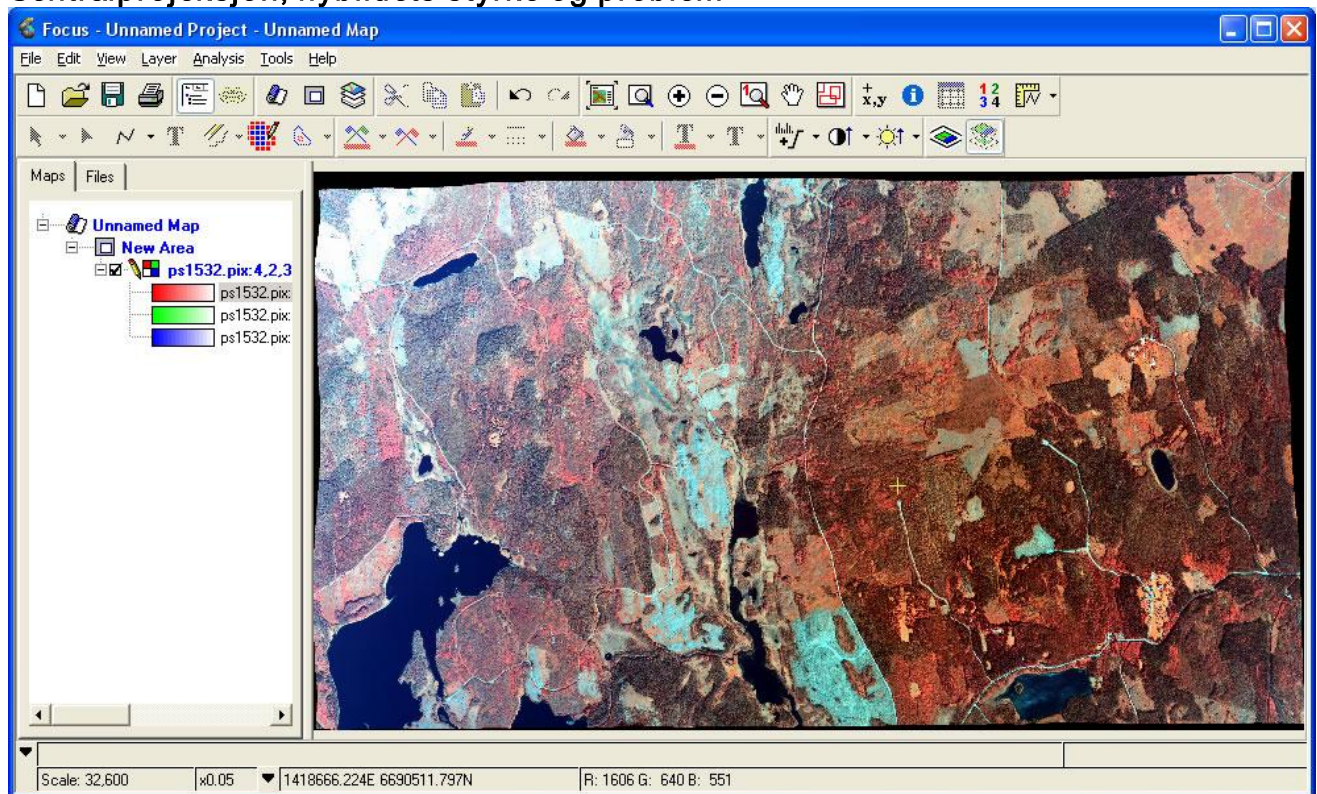


Bildet til venstre er 12 bit data (16 bit format) slik de opprinnelig er registrert i DMC 30 sensoren. Til høyre er bildene i 8 bit format etter bearbeiding hos LM. Man kan tydelig se hvordan detaljer forsvinner i 8 bit. Ved reduksjon fra 12 til 8bit er datamengden redusert 16 ganger. Det gjør det vanskelig å klassifisere alle treslag. Tilsvarende problem gjelder for analyse av landbruksarealer, parker, forurensing, utbygginger, flom osv.

Ved å konvertere fra 12 til 8 bit, reduserer man plassmengden og øker hastigheten. Til bruk i GIS kan dette være gunstig. Informasjonsmengden i bildet blir også redusert 16 ganger. Ønsker man på et senere tidspunkt å kunne bruke data til analyse, må man lagre de opprinnelige data i 16 bit format før man bearbeider de for mye.

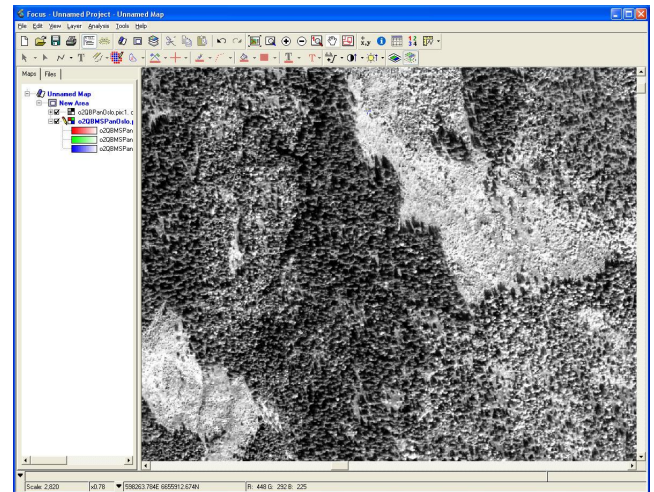
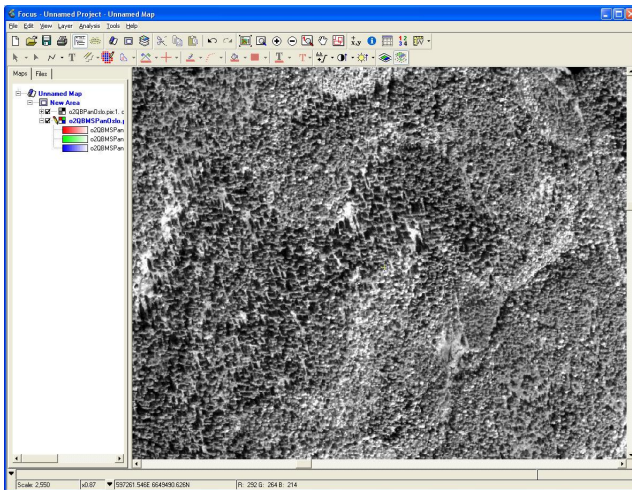
Litt oppklaring rundt 12 bit /16 bit: Sensoren registrerer data med 12 bit. Lagringsformatet på maskinen er 16 bit fordi dette er et standard dataformat som kan leses i de fleste program.

## Sentralprojeksjon, flybildets styrke og problem



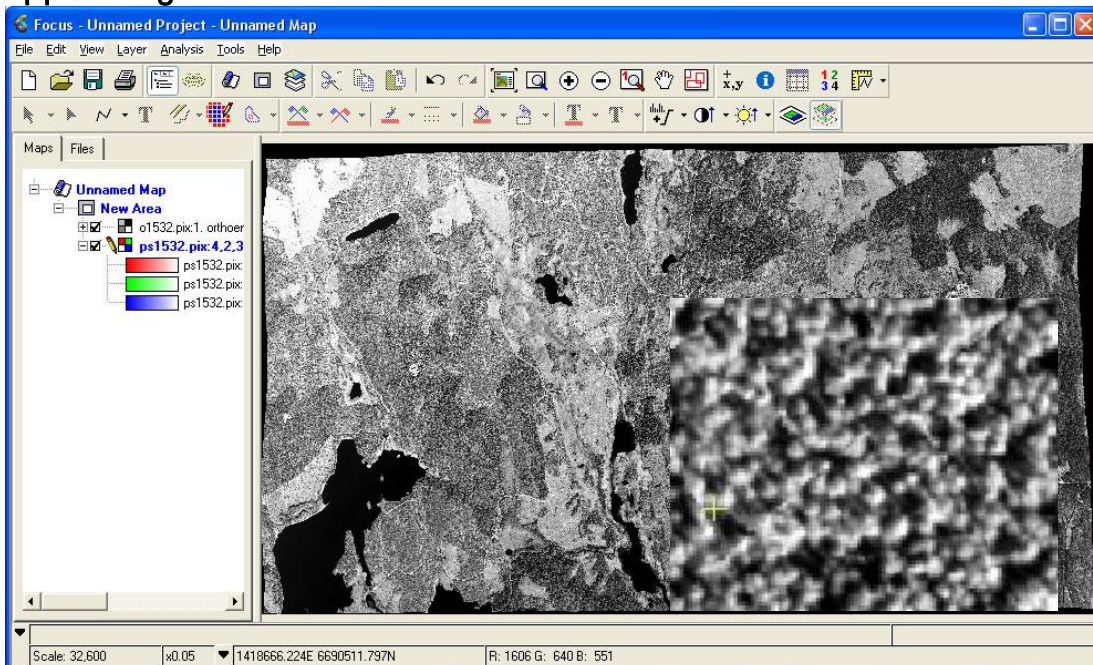
Dette IR bildet viser noen problemer med flybilder. Fargene og lysstyrken varierer med solinnstrålingen over bildet. Data er lyse i nordvest og mørke i sydøst. Ulempen med forskjellig farge over bildet, er at vi må hente inn svært flere prøveflater for å kunne analysere korrekt. Før klassifisering markerer vi et visst antall prøveflater på bildet/datsettet som vi kjenner egenskapene til og forteller programmet at dette er løv, tall/furu eller gran. Prøveflatene brukes for å klassifisere resten av bildet.





Data vi mottok fra LM har omtrent samme oppløsning som QuickBird. På disse to bildene fra et QuickBird datasett over skogene nord for Oslo, ser vi at trærne står i samme vinkel selv om det ene bildet er fra senter og det andre fra det nordvestre hjørne. I motsetning til data fra fly har høyoppløselige satellittdata samme fargetoning over hele bildet, og gir vesentlig bedre resultat ved analyse med adskillig færre prøveflater. Eller sagt på en annen måte: bedre kvalitet til lavere totalpris.

## Oppløsning:

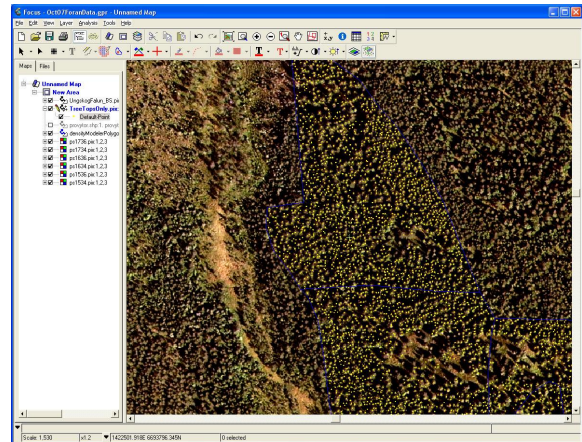
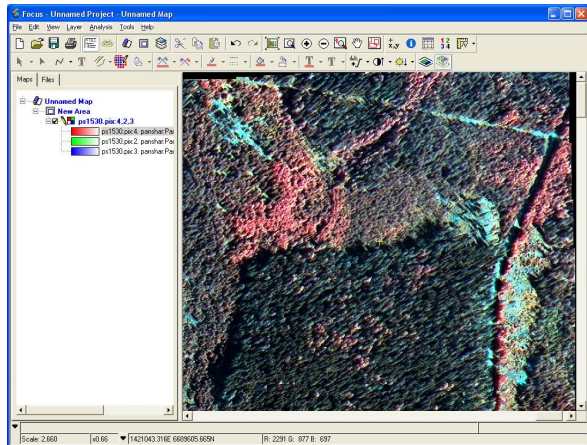


For å skille ut hvert enkelt tre, bruker vi pankromatisk eller sort/hvit data som har høyest oppløsning. Utsnittet på bildet over viser forstørrelse av et ungskogområde. Når hvert tre er identifisert, bruker vi multispektrale data for å bestemme treslag. Skog plantes med ett tre pr 2 m. Mellom plantene kommer naturlig tilvekst av gran furu og løv. Hvert pixel i LM's multispektrale data er på 2,4 m og kan inneholde 2-3 trær. Pixelstørrelsen gjør det umulig å bestemme treslag for hvert tre i ungskog. Vi benytter i slike tilfeller statistiske gjennomsnittsverdier. I Norge flyr man lavere ved kartlegging med ca 0,2 m oppløsning i Pan og 1,0 – 1,2 m oppløsning i MS. Dette forbedrer mulighetene for analyse dramatisk i ungskog. Vi får ca 6 gg fler pixler til analysen med denne oppløsningen.



Med LM sine datasett er det ikke vanskelig å telle trær i plantefelt godt og enda vanskeligere å skille mellom gran og furu. Løvandel og behov for rydding/røying er mye enklere. Vi fikk brukbare resultat for ungskog. Trærne må ha en viss størrelse for at vi skal kunne bestemme treslag med 80-90 % sikkerhet.

## Overlapp mellom bildestripene

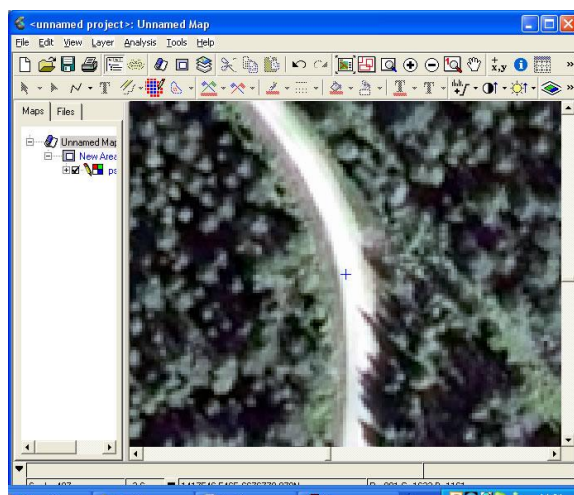
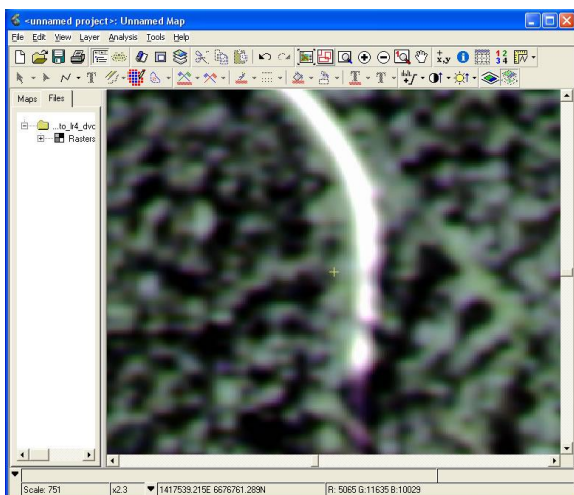


Med sentralprojeksjon vil trærne i kanten av datasettet 'helle' utover. Spesielt ille blir det i hjørnene, som på venstre bildet over. Bildet til høyre er fra midten på bildet. Beregning av kronestørrelse og antall trær blir mye vanskeligere når trærne heller.

Når flyet tar opp data, lages det datasett med 60% overlapp i flyets fartsretning. Mellom stripene flyet flyr, er det normalt 20% overlapp. Vår erfaring er at man bør fly med 30% overlapp mellom stripene for å fjerne de største problemene med hellende trær. I lengderetningen er overlapp god.

Dette problemet unngår man helt med satellitt.

## Pan Sharpening



Bildet til venstre er data modifisert av LM. Bildet til høyre er data bearbeidet av oss i PCI Geomatica. Dette viser dette hvor viktig det er å ha full kontroll på alle prosessene som brukes ved bearbeiding av digitale bildedata. Prosesser som kan være egnet for visuell presentasjon, kan være fullstendig ødeleggende for maskinell analyse.

De fleste sensorer (kamera) har en oppløsning på Pan sensoren som er ca 4 gg bedre enn de Multispektrale sensorene. Lysmengden er vesentlig mindre når man bare skal registrere i en del av spekteret slik som IR eller R, G og B slik at flaten må være større for å få nok lys til å registrere.



Pan Sharpening i PCI Geomatica er utviklet av dr Chang ved University of New Brunswick. Hans metode tar sikte på å bevare mest mulig av fargeverdiene ved Pan Sharpening prosessen slik at de er tilnærmet lik verdiene før prosessen.

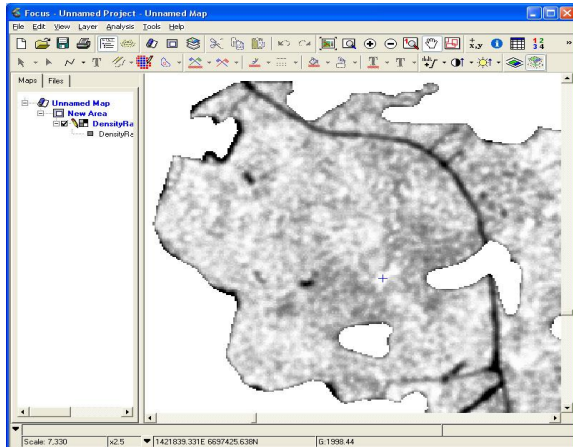
Til analysen brukte vi MS og PAN data bare minimalt justert for å få best resultat. Bildedata levert til Bergvik skog er Ortorektifisert, Pan sharpened, mosaikkert og fargebalansert. I motsetning til LM, leverte vi data i 16 bit (12 bit) istedet for 8 bit. Brukerne hos Bergvik / Stora Enso melder at vår behandling av data gir bedre visuelt inntrykk av data enn data fra LM. Vi tror dette skyldes tre forhold: 12 bit data gir flere detaljer, vår pansharpening rutine tar vare på opprinnelige farger og tonutjevningen ved mosaikking er justert for vann og sammensetting av hele bildet.

## Analyse, manuell / maskinell

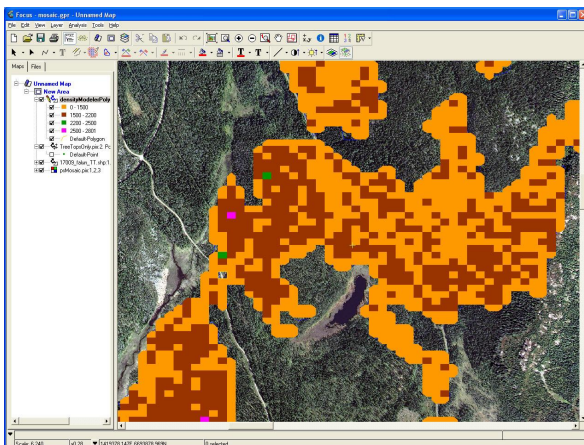
Resultatene av analysen ble som forventet relativt gode. De problemene vi antok på forhånd, stemte med resultatet. Til generell skogkartlegging er LM's datasett gode. Våre resultat er mer detaljerte og gir et mer eksakt bilde av skogens tilstand enn tilsvarende manuelle inndelinger utført på flybilder.

Kartlegging av død og syk skog gjøres svært effektivt og godt.

De største utfordringene hadde vi i ungskog. Med en oppløsning på 2,4 m i Multispektrale data, kan man vanskelig skille gran og tall/furu.

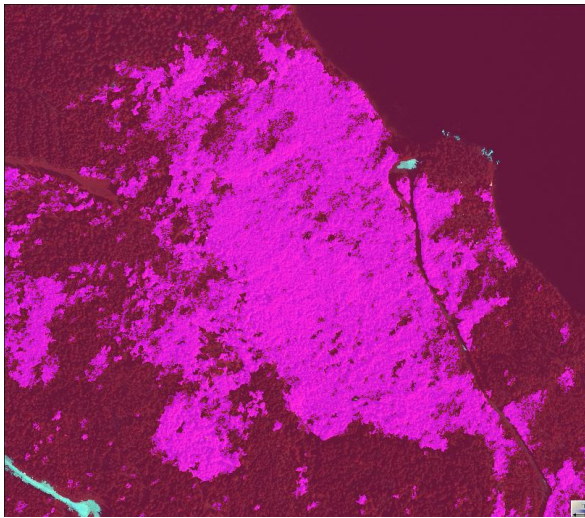


Bildet viser et analysert område. Lyse partier er tette områder og mørke partier er spredte områder. Man kan tydelig se veier gjennom området i denne analysen.

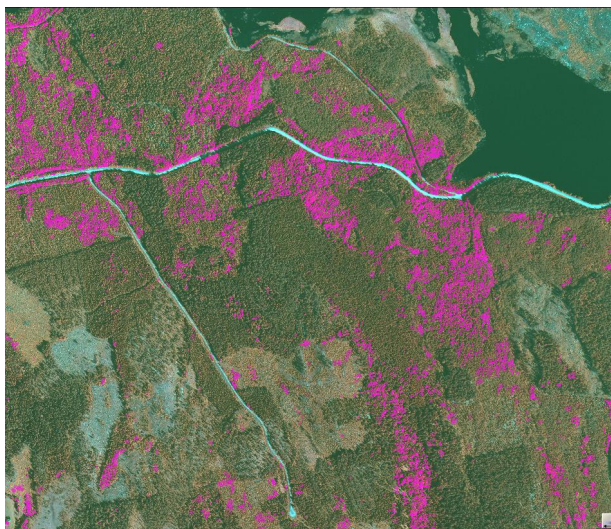


Tretelling data ble levert både som raster og vektor. Oppdragsgiver kan selv lage temakart over tetthets klasser. Skogens tetthet er registrert godt, men at tallverdiene ligger jevnt litt under det reelle.





IR data gjør det mulig å gjøre gode tolkninger manuelt. Det er likevel ved maskinell analyse at man får eksakte tall og mengder presentert. Analysen til venstre viser maskinell analyse av løv.



Bildet viser fordelene med en maskinell analyse og grafisk presentasjon av løv. Det blir lettere å se og vurdere andel løv i ulike deler av skogen. Resultat av analysen foreligger som raster og vektor kart. Dette kan brukes som overlay mot bestand for å få ut arealoversikt over andel lauv i hvert bestand.

## Oppsummering

Ved automatiske / semiautomatiske analyser må man vurdere følgende krav til bilde-data:

Data bør være i det beste dataformat som sensoren kan gi, fra DMC 30 betyr det 16 bit data.

Data bør være minst mulig bearbeidet. Alle rutiner for fargejustering, pansharpning mm påvirker sensorens registreringer og kan redusere kvalitet på analyse resultat.

Flyhøyde bør tilpasses behov for kartlegging. Detaljer i plantefelt og ungskog forsvinner med for stor flyhøyde. 4 800 m høyde kan synes å være for høyt.

Kanten på bildet/datasettet får dårlig kvalitet. Med 30% overlapp mellom flystripene blir andelen sentrale områder på bildet større og analyseresultatet bedre.

Sensoren må ha IR kanal for kartlegging av skog, landbruk og til annen vegetasjon.

I Norden er det mer enn 4 flyselskap som konkurrerer om å levere data, slik at det bør være mulig å få utført flyvning slik man selv ønsker det.

Høyoppløselige satellittbilder gir bedre resultat, lavere analysekostnader og samme eller bedre kvalitet på ortofoto som dagens produksjon fra LM.

Vi har sammenlignet arbeidsmengde og produksjonskostnader ved ulike alternativ og satt opp følgende tabell for analysekostnader. Bildekostnader ikke inkludert.

| Laser data combination                                     | RGB scanned | FCIR scanned | DMC       | ADS-40 2 <sup>nd</sup> | Quickbird |
|------------------------------------------------------------|-------------|--------------|-----------|------------------------|-----------|
| + 0,50 €/ha                                                | 4,25 €/ha   | 3,00 €       | 3,00 €    | 2,75 €                 | 1,50 €    |
| Antall klasser som kan registreres med god presisjon > 80% | 4 klasser   | 6 klasser    | 8 klasser | 8 klasser              | 8 klasser |

Kostnadene vil variere med krav til inndeling i skogen. Store homogene datasett favoriserer satellitt. Man trenger vesentlig færre prøveflater pr arealenhet.

ADS-40 2nd edition sensoren fra Leica tilsvarer eller er noe bedre enn DMC 30. Den største årsaken til forskjellen i vår sammenligning er at ADS-40 data har 0,2 m oppløsning istedet for 0,48 m. Videre er det 30% overlapp mellom flystripene. Det er Canadiske myndigheter som har bestilt denne flyvningen.

Før man kjøper bildedata må man vite hva de skal brukes til og hvilke muligheter man har. Da får man best resultat til lavest kostnad.

#### Programvare brukt i vår test:

All ortorektifisering, pansharpning, mosaikking og tonutjevning er utført med PCI Geomatica.

Analysene av data er foretatt med egenutviklet programvare, PCI Geomatica og Definiens.

Alle store kjøpere av digitale data bør skaffe seg tilsvarende programvare for bearbeiding av data selv. Det er enkelt å lære og raskt å utføre. Da har man full kontroll på hele produksjonen.

#### Takk til:

Produktsjef Kristina Kallur-Jäderkvist og Daniel Åkerman fra LantMäteriet har gitt god bistand ved leveranse av ulike datasett. Vi sendte Kristina et forslag for hvordan vi ønsker at LM's leveranser av digitale data bør se ut. LM melder at de nå har foretatt en endring i sitt produktspekter som det skal bli interessant å studere.

Ulf Thernström fra Stora Enso har vært en god støttespiller ved å definere krav og gjennomføring av de praktiske analysene av bildene. Hans erfaring med bruk av digitale bilder til skogkartlegging har vært svært viktig.

Vi har også hatt et godt samarbeid med Tobias Jonmeister fra Foran i prosjektet om 'åtgjerder i ungskog'.

**TerraNor**

Nils Erik Jørgensen



Vedlegg 1: Forslag til Lantmäteriets produksjonslinje, brev fra TerraNor til LM:

2 april 2007

Til Lantmäteriet  
Division Informationsförsörjning, Bild  
Lantmäterigatan 2, SE-801 82 Gävle  
Att: Kristina Kallur-Jäderkvist

### Forslag til produkt for digitale data fra Lantmäteriet

Viser til møte 6 februar i Gävle og telefonsamtal i dag med Kristina Kallur-Jäderkvist og Daniel Åkerman.

### Bakgrunn:

TerraNor samarbeider med Foran Sverige om å ta frem rimeligere og bedre metoder for skogkartlegging. Metoden stiller store krav til kvalitet på digitale data. LM sin nye DMC sensor har høy kvalitet og bør kunne leverere gode data.

### Forslag til produkter från DMC 30:

Bakgrunn for dette forslaget er fra vårt møte 6. februar. Referat fra dette møtet ble gjennomgått av Daniel Åkerman og følger vedlagt.

| Produkt              | Bit storlek | Band      | Merknad                                                                                                                                                             |
|----------------------|-------------|-----------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1a: Rådata           | 12 bit      | P+RGB+IR  | Data er ikke korrigert hverken radiometrisk eller geokorrigert. Det er kun grunnleggende tilpasninger som man behøver Intergraph sin programvare for å kunne gjøre. |
| 1b: Rådata korrigert | 12 bit      | P+RGB+IR  | Data er ortorektifisert og radiometrisk korrigert. Korleksjoner er utført med Nearest neighbour for å sikre bruke i bildebehandling.                                |
| 2: Pan sharpened     | 12 bit      | P+RGB+IR  | 1b + pan sharpened. Farg + IR får samme oppløsning som Pan kromatisk                                                                                                |
| 3: Tonutjevning      | 12 bit      | P+RGB+IR  | 2 + Tonutjevning av data                                                                                                                                            |
| 4: Ortofoto mosaikk  | 12 bit      | P+RGB+IR  | 3 + sammensetting av bilder over et større område til sømløs ortofoton                                                                                              |
| 5: RGB bild          | 8 bit       | RGB/ IRGB | Enten RGB eller IR bilde som bakgrunn for digital karta                                                                                                             |

Det er 2 måter å prise dette på: enten som produkt slik de er definert eller som funksjoner man tar betalt for. Det siste er kanskje enklere fordi man da enklere kan sette sammen produkt som passer kunden best.

Oppløsning: Dette er avhengig av flyhøyde. I dag flyr dere med 0,48m oppløsning uansett. En ide er å fly med 0,25m oppløsning over skogsområdene varje 7de/8de år istedet.

Jag forutsetter at Daniel Åkerman korrigerer mitt forslag slik at det passer med virkeligheten i LM.

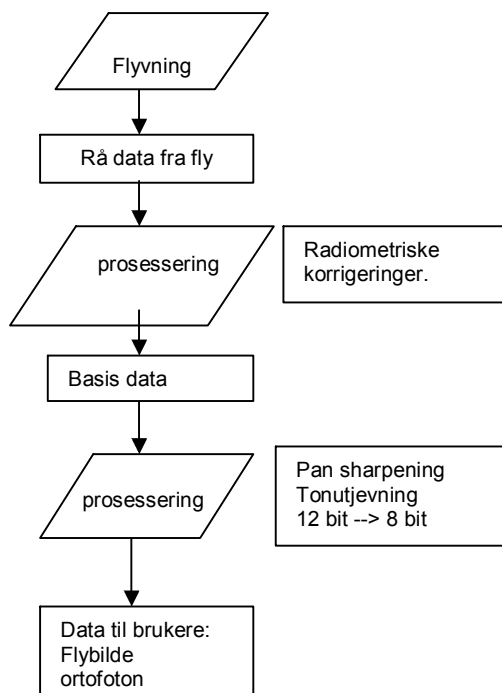
Med hilsen

**TerraNor**

Nils Erik Jörgensen

Vedlegg 2, gjeldende produksjonslinje desember 2007:

## *Lantmäteriets digitale produksjonslinje.*



### Produksjonslinje digitale data fra LM.

Basis data er i 16 bit og inneholder alle data. Bildene har lyse og mørke deler siden ene halvdel av bildet viser solsiden og den andre viser skyggesiden av trærne. Data blir registrert i 12 bit. Ved produksjon av ferdige data til brukere skjer både fargeutjevning, pansharpning og datareduksjon til 8 bit. Disse 3 prosessene reduserer datakvaliteten for maskinell behandling.

Visuell tolkning kan være bedre med de ferdige data, mens maskinell analyse av detaljer blir vanskeligere. Produksjonslinjen bør evalueres for å se om man behøver basis data eller ferdige data. Basis data er rimeligere i produksjon og klar dagen etter flyvning. Ferdige data koster mer og tar lengre tid å produsere.