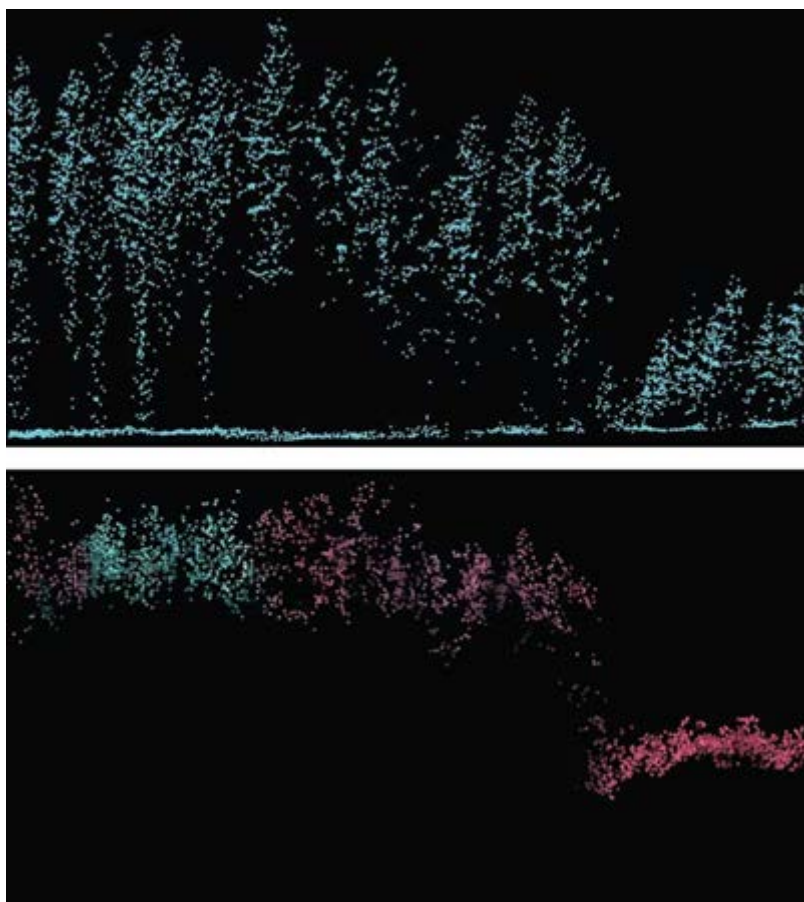




Laserskanning och digital fotogrammetri i skogsbruket

Andra upplagan

Karin Nordkvist, Emma Sandström, Heather Reese och Håkan Olsson



Arbetsrapport 407 2013

Laserskanning och digital fotogrammetri i skogsbruket

Andra upplagan

Karin Nordkvist, Emma Sandström, Heather Reese och Håkan Olsson

Förord till första upplagan

Flygburen laserskanning och automatiserad användning av digital fotogrammetri håller för närvarande på att revolutionera metoderna för skogsinventering. Detta kompendium är en första ansats att utarbeta ett läromedel på svenska inom området. Kompendiet är framtaget för distanskursen *Laserskanning och digital fotogrammetri i skogsbruket* vid Institutionen för skoglig resurshushållning, SLU, hösten 2012. Vi hoppas att det även kan få en vidare spridning i det svenska skogsbruket, samt att det kan vidareutvecklas inför kommande kurstillfällen allteftersom nya erfarenheter inom detta snabbt expanderande område vinnns. Författarna är därför också tacksamma för synpunkter på kompendiets innehåll och förslag till kompletteringar. Dessa kan skickas till huvudförfattaren Karin Nordkvist (e-post karin.nordkvist@slu.se).

Kompendiet har finansierats av Erik Johan Ljungbergs Utbildningsfond och SLU.

Vid SLU har Jonas Bohlin, Magnus Ekström, Johan Holmgren, Eva Lindberg och Jörgen Wallerman bidragit till kompendiet.

Värdefulla bidrag till innehållet har även lämnats av: Helen Rost och Martin Sjödin, Blom Sweden AB; Fredrik Walter, Dianthus AB; Ulf Söderman, FORAN Remote Sensing AB; Thomas Brethvad, COWI; Andreas Rönnberg, Lantmäteriet; Anders Boberg, Tyréns.

Där inget annat anges är figurerna gjorda av Karin Nordkvist. Dessa figurer samt kompendiets text är licensierade under [Creative Commons Erkännande-DelaLika 2.5](#).

Umeå i september 2012

Karin Nordkvist och Håkan Olsson

Förord till andra upplagan

I den andra upplagan hösten 2013 så har kompendiet kompletterats med övningsuppgifter för bearbetning av laserskannerdata för skogliga ändamål, vilka återfinns i ett appendix. Övningsuppgifterna och främst anpassade för kurser vid SLU. Uppgifterna är utformade av Karin Nordkvist, Emma Sandström och Heather Reese vid SLU. Även övningsuppgifterna är finansierade av Erik Johan Ljungbergs Utbildningsfond och SLU. Endast mindre ändringar har gjorts av texten i övrigt.

Umeå i december 2013,

Karin Nordkvist, Emma Sandström, Heather Reese och Håkan Olsson

Omslagsbild: Jonas Bohlin, SLU. Bilden bygger på data från Lantmäteriet, och visar en punktmolnsprofil från laserskanning (överst) och digital fotogrammetri.

Innehållsförteckning

1	Laserskanning för skogliga tillämpningar	7
1.1	Dagsläget för laserskanning i skogsbruket.....	7
1.2	Laserskanningens utveckling	8
2	Laserskanningens grunder.....	9
2.1	Avståndsmätning	9
2.2	Profilerande laser.....	9
2.3	Skanning.....	10
2.4	Positionsbestämning av mätpunkter.....	11
2.5	Mätnoggrannhet	11
2.6	Olika typer av skannrar	12
2.7	Laserpulsens interaktion med mark och vegetation	13
3	Förbearbetning av laserdata.....	15
3.1	Stråkutjämning och inpassning mot kontrollpunkter	15
3.2	Klassificering av markpunkter	16
3.3	Framtagning av markmodell	18
4	Databearbetning för skogliga tillämpningar.....	19
4.1	Beräkning av kronhöjdsfördelningar.....	19
5	Visuell tolkning och segmentering av laserdata.....	21
5.1	Markmodellens användningsområden.....	21
5.2	Laserdata i rasterform.....	22
5.3	Automatisk segmentering.....	23
6	Areabaserade skattningar	24
6.1	Insamling och bearbetning av fältdata	24
6.2	Regression	25
6.2.1	Exempel på enkel linjär regression	25
6.2.2	Regressionsmetoden i praktiken	27
6.3	k-MSN.....	28
6.4	Random Forests.....	29
7	Skattning av variabler för enskilda träd	31
7.1	Detektering av träd	31
7.2	Mått kopplade till enskilda träd.....	33
7.3	Sammankoppling av fält- och fjärranalysdata.....	33
7.4	Skattning av trädvariabler	34
8	Automatiserad digital fotogrammetri.....	35
8.1	Färginformation från flygbilder	35
8.2	Centralprojektion och ortofoto	36
8.3	Linjeskanning och cylinderprojektion.....	37
8.4	Stereofotogrammetri.....	37
8.5	Punktmoln från stereobilder	39
8.6	Tillämpningar i skogsbruket.....	40
9	Lagring av 3D-punktdata.....	42
9.1	LAS-filens uppbyggnad	42
10	Referenser.....	44
	Appendix: Förkortningar.....	46

Appendix: Övningsdata och Programvaror	47
Appendix: Övningsuppgifter	48
Appendix: Översikt av filer och mappar till övningsdata.....	115

1 Laserskanning för skogliga tillämpningar

1.1 *Dagsläget för laserskanning i skogsbruket*

Laserskanning har etablerats som en ny och effektiv metod för skoglig datafångst. Flygburen laserskanning mäter med decimeternoggrannhet läget för punkter på marken och i trädkronorna. Med hjälp av särskilda datorprogram kan sedan markens höjd beräknas. Därefter kan även statistiska mått för de återstående laserpunkternas fördelning i trädkronorna beräknas. Med till exempel regressionsanalys kan sedan skogliga data från provytor överföras till rasterrutor som täcker all skogsmark inom den aktuella skanningen. Denna så kallade areabaserade metod introducerades 2002 som en kommersiell metod för skogsbruksplanering i Norge och används idag vid nästan all skogsbruksplanering där. Framst är det mått på trädens storlek inom respektive rasterruta (stamvolym, höjd, grundyta etc.) som skattas på detta sätt. Beståndsindelning och uppskattning av trädslagsfördelning görs i Norge i regel med manuell tolkning av digitala flygbilder i stereo.

Även de statliga finska skogscentralerna har gått över från en traditionell, fältintensiv skogskartering till automatiserad skattning med laserskanning. Arealen som skattas med laserbaserade metoder uppgår nu till 1 - 2 miljoner hektar per år. Till skillnad från i Norge försöker man i Finland att även skatta trädslag automatiskt, bland annat genom att kombinera laserdata med digitala flygbilder. Metoderna för att automatiskt skatta trädslag i operationell skala är dock ännu förhållandevis osäkra.

I Sverige har de flesta större skogsföretag och några skogsägarföreningar gjort försök med laserskattningar, ofta på testområden om ca 10 000 ha. Resultaten är i regel goda då det gäller variabler relaterade till trädens storlek. Enligt en sammanställning gjord av OL Skogsinventering AB är skattningsnoggrannheten på beståndsnivå med den areabaserade metoden 3-6 % för grundtevägd medelhöjd, 6-14 % för virkesförråd och 7-13 % för grundtevägd stamdiameter. Dessa noggrannheter är mycket bättre än vad som uppnås vid traditionella inventeringar i samband med skogsbruksplanläggning. Noggrannheten för skattning av stamantal är ca 12-24 %. [1]

Etableringen av laserskanning i skogsbruket har inte varit lika snabb i Sverige som i Norge och Finland. Den pågående nationella laserskanningen, NH, som utförs på beställning av Lantmäteriet, kan dock ändra på detta. Av de större skogsföretagen har Bergvik varit föregångare. Där pågår för närvarande en pixelvis skattning av hela skogsinnehavet, baserat på data från den nationella laserskanningen. Något som de funnit särskilt värdefullt är möjligheten till noggrann kartläggning av områden som sannolikt är i behov av gallring. Även Fastighetsverket har beställt skattningar för stora områden. Flera andra organisationer, däribland Sveaskog, använder dessutom laserdata för bland annat avgränsning av bestånd, utan att göra skattningar med stöd av provytor.

En bidragande orsak till den snabbare spridningen av laserskanning i Norge och Finland kan vara att skogsbruksplanläggning för privata fastigheter samordnas på kommunnivå, samtidigt som staten delfinansierar planerna. Någon sådan samordning sker inte i Sverige, vilket försvårar det gemensamma utnyttjandet av laserdata och provytor. Ett intressant framtida alternativ i Sverige, särskilt för planläggning av små privata fastigheter, är tredimensionella mätningar av krontaket som fås genom matchning av digitala flygbilder. Flygbilder kommer sannolikt även framöver att registreras av Lantmäteriet med större regelbundenhet än laserdata. Tekniken att använda punktmoln från digitala flygbilder går igenom i slutet av detta kompendium.

1.2 Laserskanningens utveckling

Laserljus är enfärgade, riktade ljusvågor som är i fas. Den första användbara lasern konstruerades 1960. Idag används laser inom en rad områden: i cd- och dvd-läsare, skrivare, laserspektroskopi, studier av gaser i atmosfären, avstånds- och hastighetsmätning, topografisk kartering, kartering av havsdjup i grunda områden och mycket mer. Vid flygburen laserskanning används laserns förmåga att mäta avstånd med hjälp av tiden från att en laserpuls sänds ut, tills laserljus som reflekterats från marken eller vegetationen kommer tillbaka till sensorn. Tekniken att mäta avstånd med laser kallas ofta LiDAR från engelskans *Light Detection and Ranging*.

De första försöken att mäta trädhöjder med LiDAR gjordes i dåvarande Sovjetunionen i slutet av 1970-talet. Mätning av trädhöjdsprofiler med flygburen laser testades även i Kanada och USA i början av 1980-talet [2]-[4]. De första kända studierna av flygburna skannande lasersystem för skogsinventering genomfördes i samverkan mellan dåvarande FOA (Försvarets Forskningsanstalt, numera Totalförsvarets Forskningsinstitut, FOI) och Sveriges Lantbruksuniversitet, SLU, 1991 [5], [6]. Dessa tidiga försök visade att trädhöjd och virkesförråd kunde mätas med laserskanning. Ett problem med det experimentella system som användes i detta tidiga test var den låga noggrannheten i positioneringen. I mitten av 1990-talet utvecklades kommersiella laserskanningssystem för flygregistrering över land. I dessa system hade GPS integrerats med tröghetsnavigering (*Inertial Navigation System, INS*) vilket gjorde det möjligt att bestämma lasermätningarnas position med 0,5 meters noggrannhet eller bättre.

Ett av de första kommersiella systemen var svenska TopEye, utvecklat av SAAB 1993. Detta helikopterburna system producerade mätdata av hög kvalitet. TopEye-systemet är designat för detaljerad mätning av mindre områden och har använts mycket för mätningar av olika typer av infrastruktur samt för försök med skattningar av skog. De tidiga försöken med helikopterburna laserskanning ledde dock inte till användning inom operationell skogsinventering eftersom kostnaderna för datainsamling upplevdes som alltför höga. I Norge gjorde Erik Næsset 1995 försök med att mäta skog med skannande laser monterad på flygplan. Försöken lyckades väl, vilket ledde till utvecklingen av den så kallade areabaserade metoden för inventering på beståndsnivå med stöd av fältmätta referensytor. I slutet av 1990-talet visade finsk och svensk forskning på möjligheten att även upptäcka och mäta enskilda träd i laserdata med hög punkttäthet. En redogörelse för den tidiga utvecklingen i de nordiska länderna finns i referens [7].

Idag är laserskanning en vedertagen metod för skoglig inventering. Exempel på länder eller regioner där omfattande laserskanning för skogliga ändamål gjorts är Norge, Finland, Österrike, Spanien, norra Italien, USA, Canada, Chile och Tasmanien utanför Australien. Enstaka projekt finns i ytterligare en lång rad länder, till exempel Nepal, Tanzania och Brasilien för att nämna några. År 2011 fick professor Erik Næsset Wallenbergpriset (även kallat det skogliga Nobelpriset) för att han lanserat den areabaserade metoden som en operationell metod för skogsinventering. År 2011 kan också räknas som genombrottsåret för storskalig operationell användning av metoden i Sverige, eftersom Bergviks projekt för att skatta hela sitt skogsinnehav med laserskanning startade då. I Sverige tog det således 20 år från de första försöken 1991, till den fullskaliga användningen.

2 Laserskanningens grunder

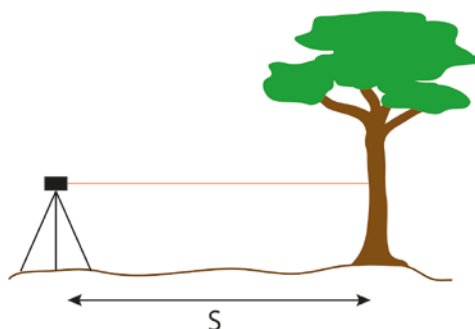
Detta kapitel beskriver de viktigaste principerna bakom laserskanning och de huvudkomponenter som används i dagens system. Framst handlar kapitlet om flygburen laserskanning. En avståndsmätande laser placeras då i ett flygplan eller en helikopter. Lasern sänder ut pulser för att mäta avståndet till punkter på marken eller i vegetationen, och en skanningsmekanism används för att sprida pulserna i ett brett stråk under planet. För positionsbestämning av mätpunkterna krävs att man med hög precision kan registrera avståndet till punkten samt laserns position och i vilken riktning pulserna sänds ut.

2.1 Avståndsmätning

Den metod som används för mätning av stora avstånd (storleksordningen 100 meter och längre) innebär att lasern sänder ut en kort (ca 4-10 ns, eller 1,2-3 m), men intensiv ljuspuls [8]. Pulsen färdas genom luften, träffar något objekt och reflekteras av objektet tillbaka till instrumentet, som registrerar hur lång tid som gått sedan pulsen sändes ut (Figur 1). Detektorn omvandlar ljuset till en elektrisk spänning där signalstyrkan är en funktion av tiden. Eftersom ljusets hastighet är känd kan avståndet mellan instrumentet och det reflekterande objektet beräknas som

$$S = v \cdot \frac{t}{2}, \quad (1)$$

där S är avståndet till objektet, v är ljushastigheten och t är den uppmätta tiden. Divisionen med 2 kommer sig av att ljuset färdas samma sträcka två gånger: först från instrumentet till objektet och sedan tillbaka. Metoden med tidtagning av returpulsen är vanlig i både markbaserade och flygburna instrument.

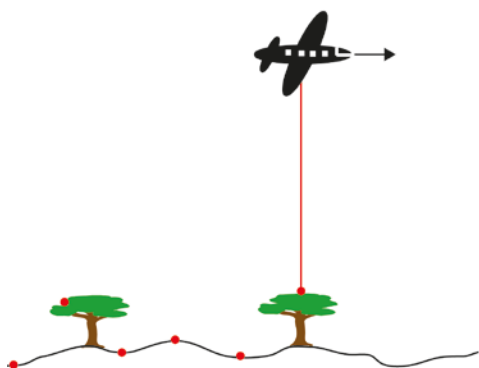


Figur 1. Avståndsmätning med laser.

Det finns även en typ av markbaserade instrument som istället mäter fasförskjutningen hos den returnerade signalen från en amplitudmodulerad laserpuls. Denna metod har mycket hög noggrannhet men lämpar sig endast för kortare avstånd.

2.2 Profilerande laser

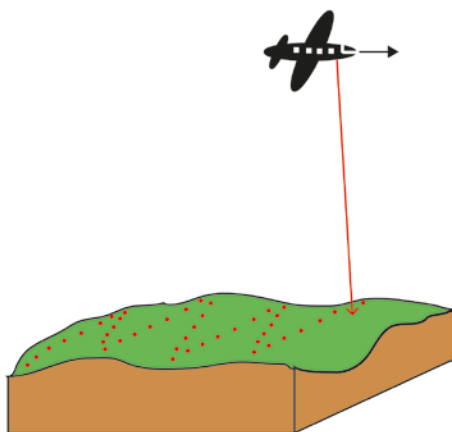
Genom att placera en avståndsmätande laser på ett flygplan kan man skapa höjdprofiler av marken och vegetationen. Under det att flygplanet rör sig framåt sänds laserpulser ut med hög frekvens och träffar marken längs en linje i planets flygriktning (figur 2). De tidiga försöken med lasermätning av skog på 1980-talet gjordes med denna teknik och potentiellt kan det vara ett billigt sätt att få skoglig statistik för stora områden.



Figur 2. Profilerande laser.

2.3 Skanning

De metoder för laserinventering av skog som nu används i det praktiska skogsbruket bygger, till skillnad från profilmätning, på skanning av hela det område som ska inventeras. Den höjdmätande lasern kompletteras därför med en skanningsmekanism som sprider mätningarna i ett stråk under flygplanet (figur 3). Skanningsmekanismen består ofta av en roterande eller oscillerande (vickande) spegel. Genom att flyga flera stråk intill varandra täcker man in större områden.



Figur 3. Flygburen laserskanning.

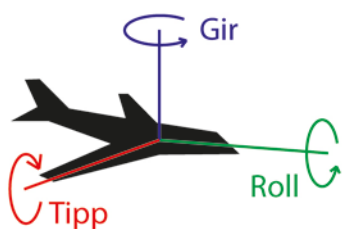
Stråkbredden bestäms av flyghöjden och öppningsvinkeln, det vill säga den maximala vinkeln mellan laserstrålen och lodlinjen. Stråkbredden och flyghastigheten avgör sedan hur stor yta som kan täckas per tidsenhet. Ju snabbare ett område kan täckas in, desto lägre blir kostnaden, och man vill därför använda en hög flyghöjd, flyghastighet och öppningsvinkel. För en given kombination av dessa parametrar bestäms punkttätheten på marken av pulsfrekvensen (antalet utsända laserpulser per sekund).

Eftersom man både vill ha en tillräckligt hög punkttäthet och hög yttäckning per tidsenhet har utvecklingen gått mot skannrar med hög pulsfrekvens. En teknisk utmaning ligger i att kombinera ökad frekvens med den större effekt som behövs vid skanning från hög höjd, eftersom lasern behöver mer tid mellan varje puls när effekten är hög. Utvecklingen har emellertid varit mycket snabb, och pulsfrekvensen hos nuvarande kommersiella system är ofta flera hundratusen utsända pulser per sekund, jämfört med ca 2 000 pulser per sekund hos de tidigaste kommersiella systemen i mitten på 1990-talet. En hög frekvens i förhållande till flyghöjd gör att en puls inte hinner tillbaka innan nästa har sänts ut, och man har därför även utvecklat system som kan hålla reda på flera pulser samtidigt i luften.

Vid höga öppningsvinklar ökar andelen laserskott som träffar vegetationen istället för marken. Det innebär dels sämre täckning av marken, dels att en del av de lasermått som används vid skogliga skattningar blir mindre representativa. Ofta begränsas därför öppningsvinkeln till ca +/- 20° vid skanning primärt för framställning av markmodeller och till ca +/- 15° vid skanning primärt för skogliga skattningar.

2.4 Positionsbestämning av mätpunkter

För att beräkna mätpunkternas position i horisontell och vertikal led måste man känna laserskannerns position och orientering vid tidpunkten för varje utsänd puls. Positionen definieras av tre koordinater i rummet (x , y och z) och orienteringen av de tre vinklar som kallas tipp, roll och gir (figur 4). Dessutom måste man känna till den utsända laserpulsens riktning i förhållande till instrumentet.



Figur 4. Tipp, roll och gir (på engelska pitch, roll, yaw).

När de första försöken med flygburen laserskanning gjordes, saknades teknik för att mäta dessa parametrar med tillräcklig noggrannhet, men i mitten av 1990-talet började man integrera GNSS (*Global Navigation Satellite Systems*) och tröghetsnavigering ([INS](#), *Inertial Navigation System*). GNSS är ett samlingsnamn för alla satellitbaserade navigationssystem. I praktiken används främst det amerikanska GPS.

GNSS ger bra mätningar av position och hastighet, men kan inte användas för att mäta tipp, roll och gir med tillräcklig noggrannhet. Med GNSS registreras positionen dessutom med relativt låg frekvens, vilket gör att flera laserpulser hinner sändas ut mellan varje mätning. För att mäta planets orientering samt fylla i luckorna mellan GNSS-mätningarna används INS. Ett INS består ofta av tre gyroskop och tre accelerometrar. Gyroskoperna mäter vinkelhastighet runt de tre axlarna, och genom att integrera med avseende på tid får man ut förändringen i orientering jämfört med ursprunglig orientering. Accelerometrarna mäter krafter och därmed acceleration. Dubbel integrering med avseende på tid ger positionsförändring jämfört med utgångspunkten. GNSS används för att korrigera för drift i tröghetsystemet.

2.5 Mätnoggrannhet

För hårdgjorda ytor beror noggrannheten och upplösningen i avståndsmätningarna i första hand på tidtagningens noggrannhet och laserpulsens längd. Eftersom avståndet till ett objekt är

$$S = v \cdot \frac{t}{2}, \quad (2)$$

så kan noggrannheten i avståndsmätningen approximeras till

$$\Delta S = v \cdot \frac{\Delta t}{2}, \quad (3)$$

där Δt är noggrannheten i tidtagningen. Ljusets hastighet är ca 300 000 km/s, vilket innebär att en mätnoggrannhet på 1 dm kräver en tidtagningsnoggrannhet på 0,67 nanosekunder ($1 \text{ ns} = 10^{-9} \text{ s}$). Ljusets hastighet är väl bestämd men påverkas något av lufttemperaturen, varför en mätnoggrannhet bättre än någon cm är svår att uppnå med flygburen laserskanning, oavsett hur väl tiden kan mätas.

Laserstrålen har låg divergens, det vill säga är väl sammanhållen – jämför den smala strålen från en laserpekare med ljuskäglan från en ficklampa. På stora avstånd är divergensen ändå tillräcklig för att pulsen ska träffa en liten yta på marken snarare än en punkt. En divergens på 0,5 mrad ($0,029^\circ$) och en flyghöjd på 1 000 meter ger en träffyta med ungefär 50 cm diameter. Större divergens och flyghöjd leder till att laserpulsen får en större träffyta, vilket påverkar mätnoggrannheten negativt. Även stor terränghöjd och öppningsvinkel kan sänka mätningens noggrannhet.

En del av laserpulsen kan reflekteras av ett objekt som inte blockerar hela pulsens strålgång, medan resten av pulsen fortsätter och reflekteras mot objekt som kommer senare i strålgången, till exempel marken. Många sensorer kan därför registrera mer än ett eko från den utsända laserstrålen. För de flesta sensorer gäller att det måste vara minst någon meter mellan ekon som registreras från samma laserpuls.

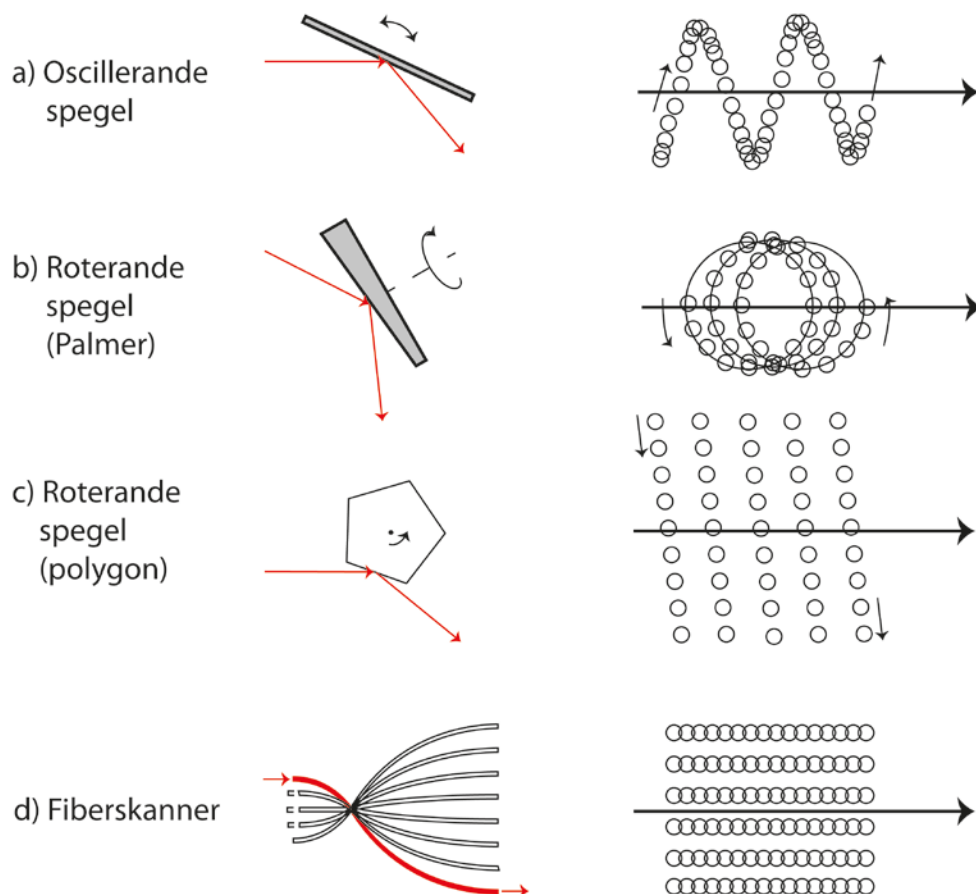
2.6 Olika typer av skannrar

Själva skanningsmekanismen består oftast av en oscillerande eller roterande spegel. En tredje typ är fiberskannern.

En oscillerande spegel svänger mellan två positioner och laserstrålarna träffar marken i ett sicksackmönster (figur 5 a). Vid vändningen saktar spegeln ner och punkttätheten vinkelrätt mot flygriktningen är därför högre i stråkets ytterkanter.

En typ av skanner med roterande spegel är Palmerskannern (figur 5 b). Spegelns yta är inte vinkelrät mot rotationsaxeln, och den utsända pulsens riktning förändras med spegelns läge. Pulserna träffar marken i ett cirkelmönster som förflyttar sig framåt i flygriktningen. Punkttätheten är högst i stråkets kanter. Polygonskannern (figur 5 c) bygger också på en roterande spegel. Spegeln behöver inte göra några accelerationer och inbromsningar, vilket ger en jämnare punkttäthet. På grund av flygplanets rörelse framåt blir mönstret något vinklat mot flygriktningen.

Fiberskannern (figur 5 d) är uppbyggd av ett knippe med optiska fibrer. En emitterad laserpuls går in i en fiber i knippets ena ände, där fibrerna är samlade i en cirkel. I knippets andra ände, där pulsen går ut, ligger fibrerna på en linje. På så sätt överförs en cirkulär skanningsrörelse till ett linjärt mönster. En roterande spegel fördelar pulserna mellan fibrerna. Spegelarna som används är mindre än i andra skannrar och kan därmed rotera snabbare. Punkttätheten blir betydligt högre i flygriktningen än över stråkbredden. Ett sätt att kompensera för det är att låta fiberknippets emitterande ände svänga fram och tillbaka så att pulserna från varje fiber bildar ett sicksackmönster. [9]



Figur 5. Olika typer av skanningsmekanismer, och resulterande skanningsmönster.

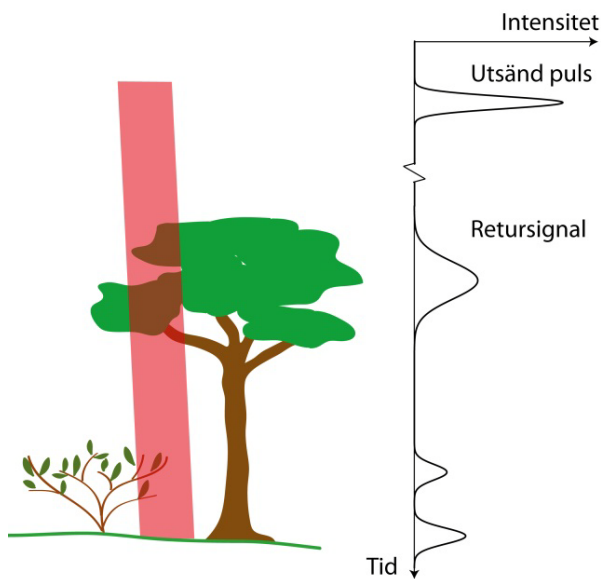
2.7 Laserpulsens interaktion med mark och vegetation

På sin väg mot marken kan den utsända laserpulsens reflekteras av marken eller av objekt över marken: trädkronor, stammar, övrig vegetation, stenar, byggnader med mera. En hård och kompakt yta resulterar i en enda distinkt returpuls, men då laserstrålen träffar till exempel en trädkrona eller kanten av ett hustak kan en del av pulsen reflekteras medan resten fortsätter. Vatten har låg reflektans och ger därför i stort sett inga returer.

Ofta registreras, förutom koordinaterna, även intensiteten för varje detekterad laserretur. Många lasersystem använder nära infrarött ljus, och en yta som har hög reflektans i detta våglängdsområde, till exempel levande vegetation, ger returer med hög intensitet. Intensiteten hos returpulserna är dock svårtolkad. Den fysikaliska innebörden av intensitetsmättet är ofta odokumenterad, och en del skannerinstrument ändrar dessutom automatiskt den mottagande sensorns känslighet beroende på om ytan som registreras är ljus eller mörk.

Figur 6 illustrerar hur reflektionen från olika lager i vegetationen och slutligen från marken kan ge flera returer från en utsänd laserpuls. Antalet registrerade returer varierar mellan olika skannrar. I skog kommer ofta förstareturer från trädkronorna och sistareturer från marken. (Det finns också skannrar som med hög frekvens samplar hela laserreturen, så kallad full vågformslaser).

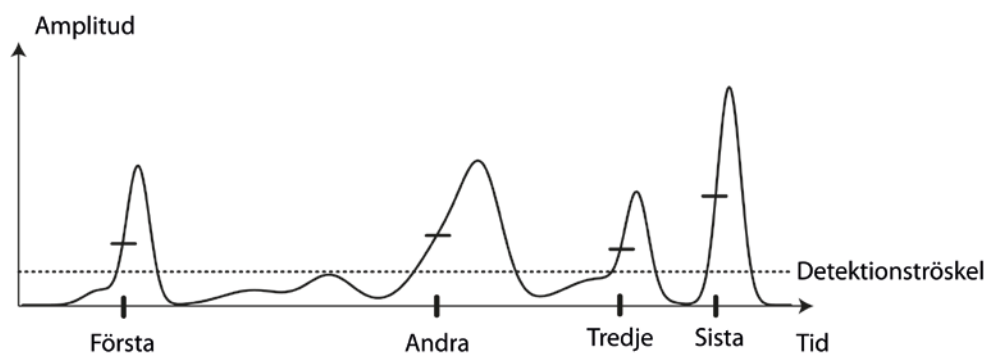
Höjdfördelningen hos laserreturerna ligger under den verkliga trädhöjdsfördelningen. Detta beror dels på att laserstrålen tränger igenom en del av trädkronan innan tillräckligt mycket energi reflekteras för att trigga en mätning i detektorn, dels på att pulserna inte bara träffar trädtopparna utan även trädkronornas sidor. Tät markvegetation kan också göra att returerna från mark och vegetation smälter samman så att markreturen delvis maskeras. Det kan leda till att returen registreras för tidigt och markhöjden överskattas.



Figur 6. Laserpulsens väg till marken. Denna situation genererar flera returur – först från trädet, sedan från busken och till sist från marken. Kurvan intill bilden representerar returpulsens intensitet som funktion av tid, med toppar orsakade av de olika objekten.

Den utsända laserpulsens längd är i storleksordningen meter, men osäkerheten i mätningarna mot en hård yta bör inte överstiga någon decimeter. Av stor vikt för mätningarna är därför den algoritm (beräkningsregel) som bestämmer vid vilken signalstyrka detektorn ska registrera en retur. En vanlig metod är att registrera en retur när en topp i retursignalen nått en viss andel, till exempel 50 %, av sitt maximala värde (figur 7). På så vis påverkas inte tidtagningen av returpulsens maximala styrka utan bara av dess bredd. Toppar som inte når upp över ett visst tröskelvärde, detektionströskeln, registreras inte alls.

Lasersystemen är kalibrerade för att ge rätt mätvärden när de träffar hårdgjorda ytor, till exempel asfalt. Vegetationen är däremot halvt genomtränglig för lasern, och det faktum att olika system kan använda olika principer för att bestämma när en returpuls ska detekteras bidrar till att mätvärden över vegetation från olika lasersystem inte är direkt jämförbara. Den dokumentation som finns tillgänglig för användaren om vilken metod som används är också ofta bristfällig.

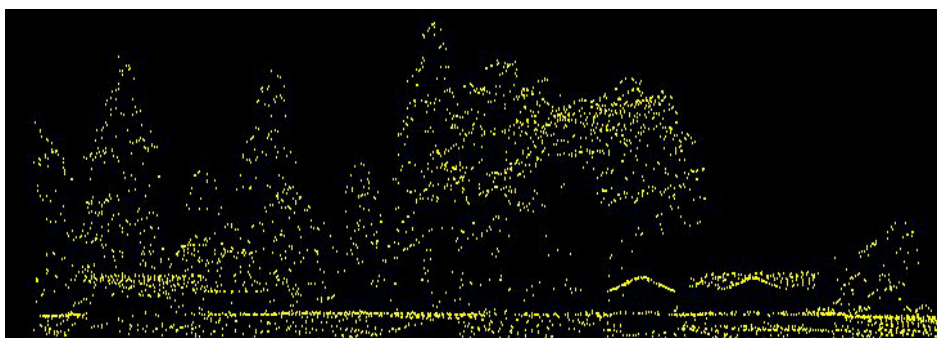


Figur 7. En vanlig metod för registrering av retursignal är att en topp i signalen triggar en registrering när den når en viss andel av sitt maximala värde, till exempel 50 % (vågräta streck). Endast toppar som når över detektionströskeln (prickad linje) registreras.

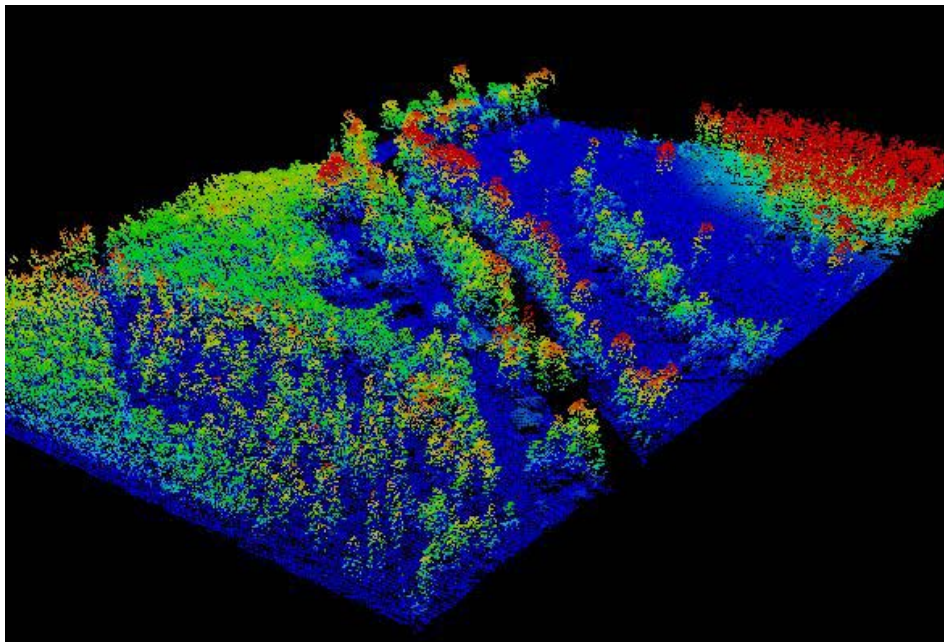
3 Förbearbetning av laserdata

De laserreturer som registreras vid skanningen bildar tillsammans ett så kallat punktmoln, en svärm av punkter koordinatsatta i tre dimensioner (figur 8 och figur 9). Innan data används för vidare analys krävs en viss förbearbetning. Denna görs i regel av dataleverantören och är inte specifik för skogliga tillämpningar. Förbearbetningen består av

- stråkutjämning och inpassning mot kontrollpunkter
- utrensning av felaktiga mätningar
- klassificering av laserreturer i klasserna mark, vatten och övrigt (inklusive vegetation)
- framtagning av markmodell.



Figur 8. Profil av laserpunktmoln där träd och byggnader framträder tydligt. Bilden bygger på data från Lantmäteriet.

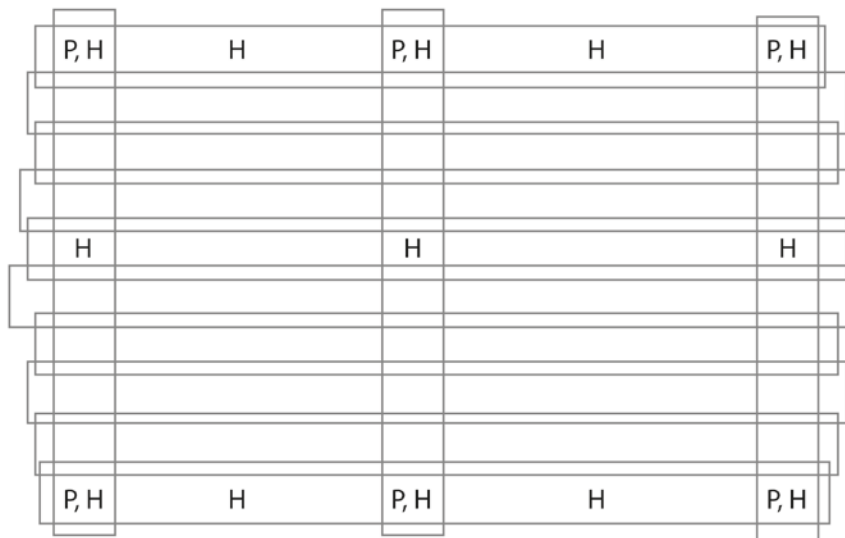


Figur 9. Punktmoln i tre dimensioner, sett snett uppifrån. Punkterna är färgade från blått till rött efter stigande höjd. Genom området löper en å. Bilden bygger på data från Lantmäteriet.

3.1 Stråkutjämning och inpassning mot kontrollpunkter

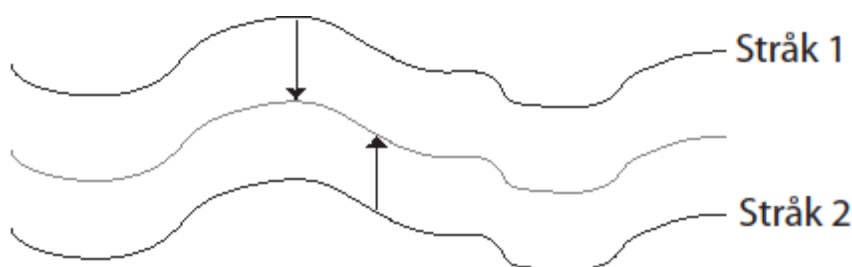
En viktig felkälla vid positioneringen av mätpunkter är navigeringssystemet med GPS och INS. När ett område skannas görs det i regel i stråk fram och tillbaka, med en viss

övertäckning i sidled mellan stråken. Man brukar också lägga tvärgående stråk i början, slutet och eventuellt mitten av stråken (figur 10). Osäkerheten i navigeringen varierar med tiden. Vid flygplaneringen anpassas stråklängden efter driften i systemet så att felet ska vara ungefär detsamma längs hela stråket. Mellan de olika stråken är skillnaden i navigeringsfel större vilket gör att stråken i allmänhet inte sammanfaller med varandra i vertikal och horisontell led.



Figur 10. Utlägg av stråk, tvärstråk, planstöd (P) och höjdstöd (H) i ett skanningsområde.

Data från olika stråk passas in i ett gemensamt skarvlöst block vid en bearbetning som kallas stråkutjämning (figur 11). Därefter passas blocket med data in mot kontrollpunkter på marken. Planstöd används för att passa in data i sidled, och består av objekt som enkelt kan identifieras i punktmolnet. Exempel på lämpliga planstöd är hustak och diken. En annan möjlighet är att använda vitmålade vägmarkeringar som syns i intensitetsdata. Därefter passas stråken in i höjddled. Till höjdstöd vill man ha öppna, jämna ytor med liten lutning. Hårda ytor som asfalt och grus är lämpliga eftersom de ger hög noggrannhet i höjdmätningarna. [10]

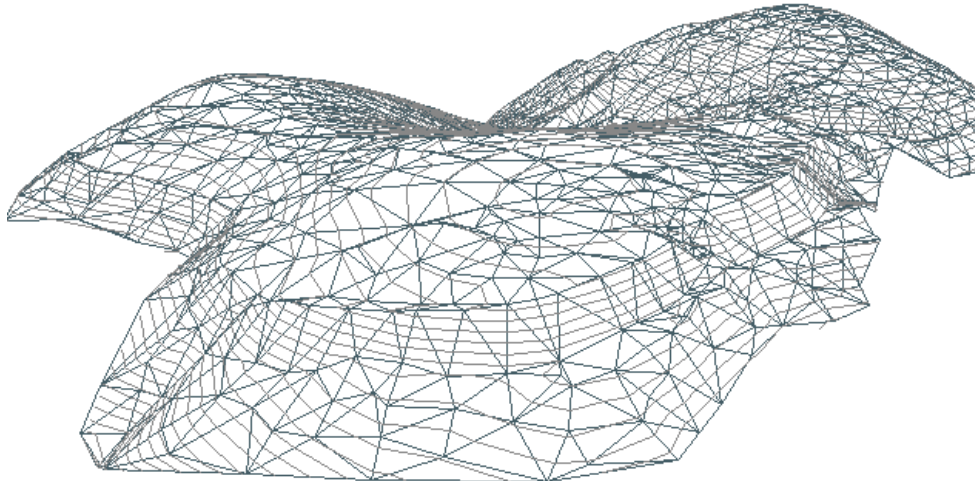


Figur 11. Höjddata i profil från två olika flygstråk. Vid stråkutjämningen passas stråken in mot varandra, antingen mot en medelyta eller mot en känd yta.

3.2 Klassificering av markpunkter

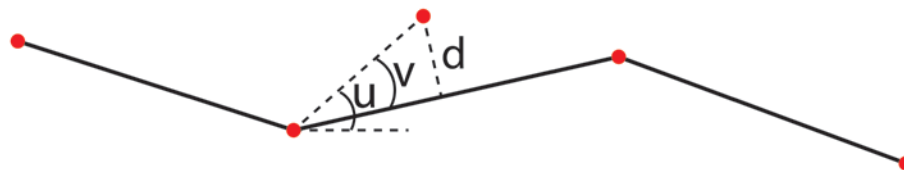
De felaktiga punkter som tas bort är sådana som ligger högt upp i luften till följd av reflektion från till exempel moln, dimma och fåglar, och sådana som tydligt ligger under marknivån. Låga punkter kan uppkomma när en puls reflekteras i flera steg så att retursignalen fördröjs, varvid punkten ser ut att ligga längre bort än den i själva verket gör. Klassificering av markpunkter kan göras på olika sätt, varav två beskrivs nedan.

En metod utvecklad på KTH av Peter Axelsson används i TerraSolids program TerraScan, som har fått mycket stor spridning [11]. Till att börja med läggs ett rektangulärt rutnät ut över punktmolnet, med en cellstorlek som bestäms av användaren. I varje cell väljs den lägsta punkten ut och klassas som en markpunkt. De utvalda punkterna binds samman i ett nät av trianglar som ger en ungefärlig representation av markytan. Denna datastruktur kallas för [TIN](#) (*Triangulated Irregular Network*) (figur 12). Nu börjar en ny process där man steg för steg lägger till nya punkter i nätverket så att det förtätas och följer markytan närmare. Punkter i



Figur 12. Markmodell i form av TIN (mörkgrå linjer). De ljusare linjerna är höjdkurvor. (Efter bild av [Robert Kropf](#) på [de.wikipedia](#). Licencierad av upphovsmannen under [GFDL](#)).

molnet undersöks en och en, och förkastas eller accepteras som nya markpunkter enligt vissa kriterier. Ett kriterium baseras på avståndet mellan punkten och den befintliga ytan, ett annat på hur mycket markytans lutning förändras om den nya punkten läggs till och ett tredje kriterium gäller den största tillåtna marklutningen (figur 13). Vilka värden som är lämpliga för dessa parametrar beror bland annat på topografin.

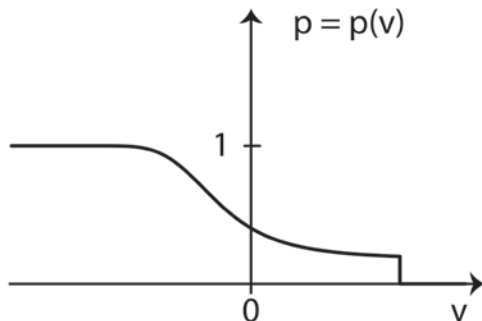


Figur 13. I Axelssons metod för markklassning accepteras en ny punkt som markpunkt om avståndet d till den befintliga ytan, vinkeln v mot ytan samt marklutningen u understiger valda gränsvärden [11].

En annan metod skapad av Karl Kraus och Norbert Pfeifer, vid Wiens tekniska universitet, går ut på att först skapa en yta med hjälp av alla punkter, där varje punkt ges lika stor vikt [12]. Ytan hamnar någonstans mellan marken och krontaket, och markpunkter återfinns med större sannolikhet under ytan än över. Varje punkt ges nu en vikt som beror av dess avstånd och riktning från ytan (figur 14). Ju lägre en punkt ligger i förhållande till ytan, desto högre vikt får den, ner till en gräns under vilken alla punkter får den maximala vikten 1.

Man sätter också ett avstånd över ytan där alla punkter får vikten 0 och alltså inte påverkar ytan. De nya vikterna används för att skapa en ny yta, genom att punkter med högre vikt ”drar” ytan till sig. Vikterna uppdateras igen, och ytterligare en ny yta skapas. Detta fortsätter ett antal gånger – ungefär 3-5 iterationer kan vara lagom. Efter den sista iterationen används

den senast beräknade ytan för att bestämma vilka punkter som ska klassas som markträffar. Denna metod används bland annat i FUSION (http://forsys.cfr.washington.edu/JFSP06/lidar_&_ifsar_tools.htm), en programvara för bearbetning av laserdata för skogliga tillämpningar.



Figur 14. Vikt p som funktion av avstånd v från ytan, i Kraus och Pfeifers metod [12]. Över ytan ($v > 0$) avtar vikten för att över ett visst värde bli 0. Punkter under ytan ($v < 0$) får högre vikt ju lägre de ligger, dock aldrig högre än 1.

3.3 Framtagning av markmodell

Efter klassningen av markpunkterna görs en digital markmodell (*Digital Elevation Model*, DEM), i form av ett TIN eller ett raster. Om markklassningen gjorts med Axelssons metod, som beskrevs i föregående avsnitt, finns redan ett färdigt TIN. Har man istället använt en metod som enbart klassar punkterna i mark och övrigt, till exempel Kraus och Pfeifers metod ovan, har man ett antal markklassade punkter som kan kopplas samman till ett TIN. Till fördelarna med ett TIN hör att alla ingående markpunkter behåller sina koordinater i planet, så att noggrannheten i punktdata bibehålls. Det är dock en komplex datastruktur som tar längre tid att skapa och bearbeta än ett raster.

Ett raster kan genereras utifrån ett TIN genom att en rastercell tilldelas den höjd som TIN-datastrukturen ger i cellens mitt. Man kan även skapa ett raster direkt från en samling markklassade punkter, antingen genom att använda markpunkterna inom rasterrutan eller genom en interpolering där även punkter strax utanför rasterrutan ingår. I ett raster behålls inte punkternas exakta position och man utnyttjar därmed inte mätningarnas fulla noggrannhet. Raster är ändå den vanligaste formen av markmodell för användning i skogliga tillämpningar, eftersom de är mycket praktiska att arbeta med.

Laserdatats punkttäthet på marken är mycket viktig för både detaljeringsgraden och noggrannheten hos den slutliga markmodellen. Punkttätheten påverkas bland annat av pulsfrekvensen, flyghöjden och vegetationen. Tät vegetation gör att färre pulser träffar marken. Markmodellens noggrannhet påverkas även av noggrannheten i avståndsmätningarna, se avsnitt 2.5.

4 Databearbetning för skogliga tillämpningar

Vid tätheter om ca 1 returnerad laserpuls/m², till exempel i Lantmäteriets NH-data, får man inte detaljerad information om de enskilda träden, men sett över en lite större yta beskriver punkternas höjdfördelning skogens höjd och täthet även då glesare data används. I detta kapitel beskrivs hur man tar fram olika mått på kronhöjdsfördelningen. Denna del av databearbetningen kan i många fall göras av dataleverantören, men kräver en viss kunskap om skogliga tillämpningar. Måtten från laserdata kan användas för visuell tolkning och tillsammans med fältdata även för skattningar av skogliga data för rasterrutor, som i sin tur kan aggregeras till bestånd.

Tätare skanning ger, åtminstone för de dominerande träden, flera returerna från varje träd. För att kunna identifiera enskilda träd och även mäta deras kronform används ofta laserdata med en täthet om ca 10 pulser/m², men även lägre tätheter har visat sig fungera [13]-[15]. Mått som används vid skattning av enskilda träd beskrivs i avsnitt 7.2.

4.1 Beräkning av kronhöjdsfördelningar

När en markmodell tagits fram (avsnitt 3.3) kan höjden över marken för varje punkt beräknas som

$$h = z_p - z_m, \quad (4)$$

där z_p och z_m är punktens respektive markmodellens höjd. Därefter kan statistiska mått på punkternas höjdfördelning beräknas. Tabell 1 visar några vanligt förekommande mått på kronhöjdsfördelning för rasterrutor och provytor. Observera att tabellen endast innehåller ett urval av mått och att flera olika beteckningar och definitioner förekommer i litteraturen, se till exempel [16], [17] och [18]. Måtten används både för manuell tolkning och för automatiserade metoder, till exempel skattning på beståndsnivå. För samtliga mått i tabellen gäller att man använder höjden h över markmodellen. Vid beräkning av höjdmått använder man i regel enbart vegetationsträffar över en tröskel på ca 2 m över marken, för att undvika att returerna från mark och låg vegetation påverkar måtten. Det kan också finnas anledning att beräkna måtten separat för till exempel första- och sista-returerna.

Med höjdpercentil menas den höjd inom ett område, till exempel en rastercell, under vilken en viss andel av punkterna i trädsiktet återfinns; 10 % av punkterna finns under den tionde percentilen, h_{10} , 20 % ligger under h_{20} och så vidare upp till h_{100} som motsvarar den högsta punkten. Det starkaste sambandet med skogens höjd finns i regel kring h_{90} . h_{100} undviks ofta då den är känslig för enstaka höga träd och felaktiga laserpunkter, och därmed även är mer beroende av punkttätheten än de lägre percentilerna.

Andelen av alla returnerade laserpulser som reflekteras från träden är ofta starkt korrelerad med kronslutenheten. Denna andel kallas för vegetationskvot, här betecknad V , och brukar beräknas som antal returerna från vegetationen dividerat med totala antalet returerna.

Så kallade krondensiteter beräknas genom att dela in höjdskillnaden mellan den lägsta och högsta vegetationsträffen i ett antal lika höga fraktioner, ofta 10 stycken. Sedan beräknas andelen returerna över respektive fraktion. Även vegetationsreturernas maxhöjd, medelhöjd, standardavvikelse och variationskoefficient är mått som ibland används.

Tabell 1. Vanligt förekommande mått på kronhöjdsfördelning

Mått	Beteckning	Definition ^a
Höjdpercentil	h_x	den höjd över marken under vilken x % av vegetationsträffarna finns
Vegetationskvot	V	n_{veg}/n_{tot}
Krondensitet	d_y	n_y/n_{tot}
Medelhöjd	h_{mean}	$\frac{\sum_{i=1}^{n_{veg}} h_i}{n_{veg}}$
Standardavvikelse	h_{sd}	$\sqrt{\frac{1}{n_{veg} - 1} \sum_{i=1}^{n_{veg}} (h_i - h_{mean})^2}$
Variationskoefficient	h_{cv}	$\frac{h_{sd}}{h_{mean}}$

^a n_{veg} är antalet returer från vegetation, n_{tot} det totala antalet returer och n_y antalet returer över höjdfraktion y om avståndet mellan den lägsta och högsta vegetationsträffen delas upp i ett antal lika stora fraktioner. h_i är höjden över markmodellen för punkt i .

5 Visuellt tolkning och segmentering av laserdata

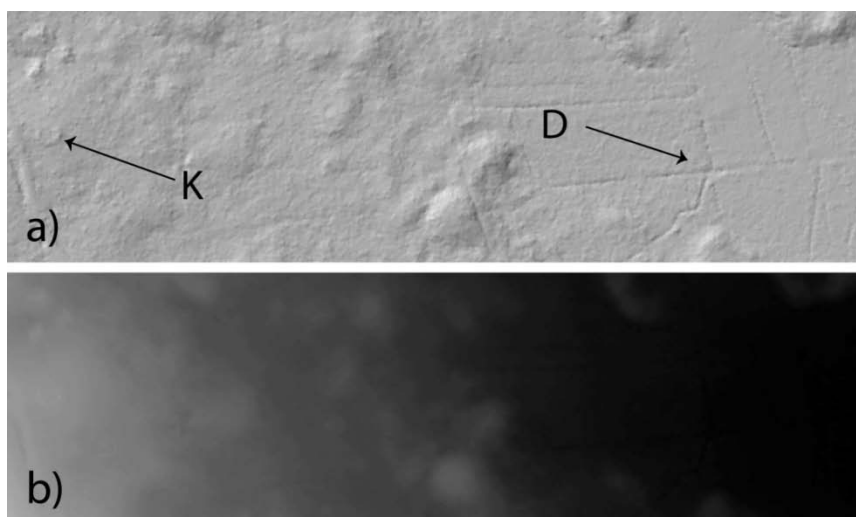
Analys av laserdata kan göras automatiskt eller manuellt (visuellt), och med eller utan stöd av fältmätta skogliga data. Detta kapitel behandlar tillämpningar som inte kräver användning av provytor med fältdata. De flesta av exemplen nedan handlar om visuell tolkning. Om skogliga mätvärden inte behövs kan laserdata också användas direkt för automatiserad identifiering av exempelvis de områden där skogen är högst eller tätast, utan stöd av fältreferensytor. En annan tillämpning som tas upp är automatisk segmentering (beståndsavgränsning).

5.1 Markmodellens användningsområden

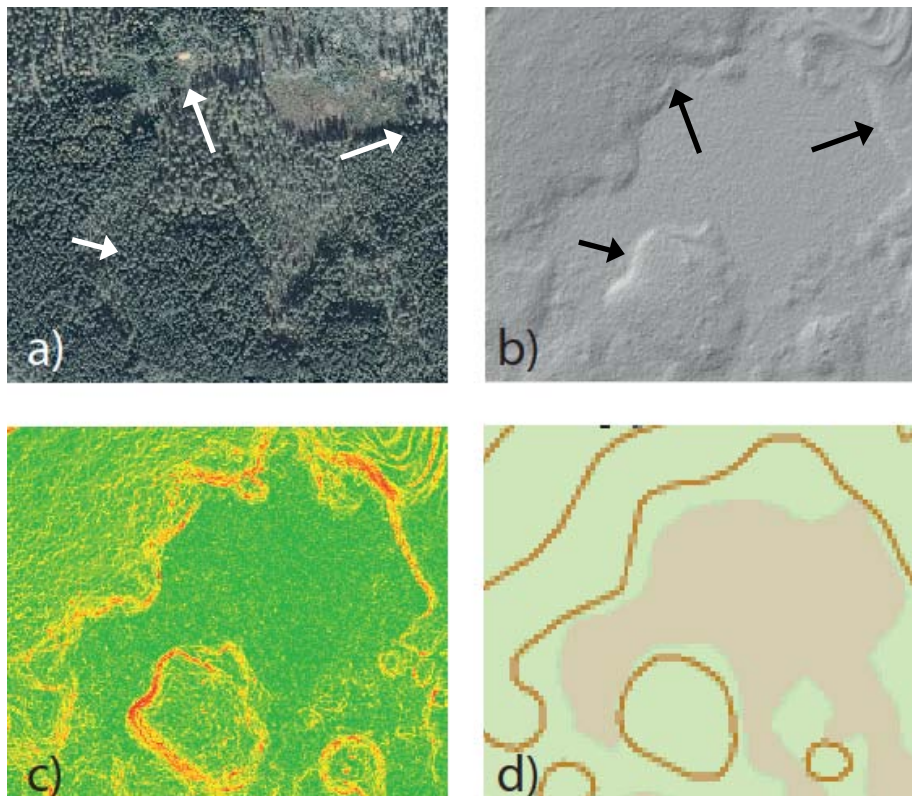
Av en markmodell i rasterform kan man göra en så kallad terrängskuggning. Den digitala markmodellen räknas om så att den återger skuggningen om den belyses av en artificiell ljuskälla. De skuggor som skapas gör det möjligt att upptäcka strukturer som inte syns lika tydligt i en vanlig rasterbild där grånivåerna används för att visa höjd. Om ljuskällan placeras i något av de övre hörnen förstår vi instinktivt modellen rätt – höjder ser ut som höjder och sänkor som sänkor. Genom att belysa modellen från olika vinklar framträder fler strukturer, vilket kan vara till hjälp när man letar efter vissa objekt.

En terrängskuggad markmodell kan användas för att hitta till exempel kolbottnar och diken. Kolbottnar är rester från kolframställning i mila eller kolningsgrop. Rester efter kolmilor kan vara runda (resmila) eller fyrkantiga (liggmila) och består ofta av en vall, ett dike eller ett antal gropar runt mittpartiet. Dessa kulturlämningar får inte markberedas eller planteras [19]. Figur 15 visar en trolig kolbotten i en terrängskuggad markmodell. Även diken och vattendrag framträder i en terrängskuggad bild. Vid till exempel skogsgödsling är det viktigt att känna till var det finns diken så att de kan undvikas.

I ett ortofoto kan det vara svårt att avgöra var gränsen går mellan myr och fast mark. Figur 16 a) och b) visar ortofoto och terrängskuggning för samma område. Pilarna pekar på myrkanter som kan vara svåra att se i fotot men är tydliga i terrängskuggningen. Markmodellen kan också användas för att beräkna branthet, sluttningsriktning, samt sannolikt fuktig mark, vilket kan vara till nytta vid planering av avverkningar.



Figur 15. a) Terrängskuggad markmodell och b) markmodell där grånivån anger höjd. Terrängskuggningen gör att ojämnheter i terrängen framträder tydligt, och avslöjar här en trolig kolbotten (K) samt ett antal diken (D). Markmodellen har här beräknats med 1 meters upplösning. Bilden bygger på data från Lantmäteriet.

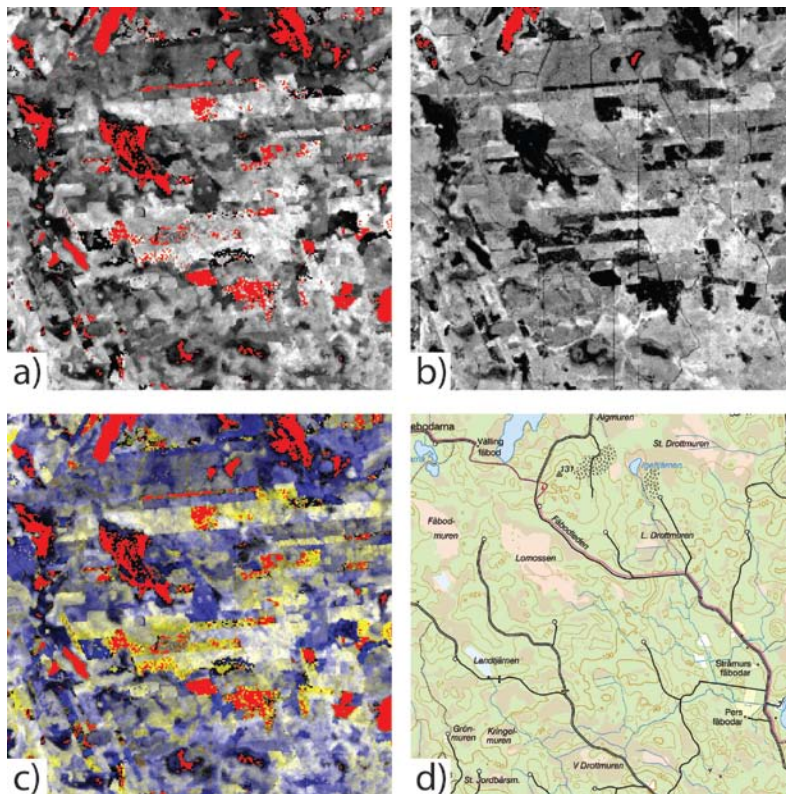


Figur 16. a) Ortofoto, b) terrängskuggning, c) branthet och d) terrängkarta för samma område. Den plana ytan i b) och det ljusbruna i d) är en myr. Pilarna markerar myrens kanter. Skalan i c) går från grönt till rött mot brantare lutning. Terrängskuggning och branthetsbild bygger på laserdata från Lantmäteriet. Ortofoto och karta: © Lantmäteriet, i2012/901.

5.2 Laserdata i rasterform

De höjd- och täthetsmått för laserreturer från trädsiktet som beskrivs i avsnitt 4.1 kan beräknas för rasterceller och visas i någon GIS-programvara. En hög höjdpercentil, till exempel h_{95} , och vegetationskvot, V , är två mycket användbara mått när man snabbt vill få en översikt av skogens relativa höjd respektive täthet.

Ett första steg kan vara att titta på rastren ett och ett (Figur 17 a och b). Man kan också kombinera de två rastren till en färgbild (Figur 17 c). Här används h_{95} i röd och grön kanal vilket ger gult, och vegetationskvot i blå kanal. Hög skog med låg kronslutenhet framträder då som gul, låg men tät skog blir blå och områden där både h_{95} och V är höga blir ljusgrå till vita. Figur 17 d visar terrängkartan för jämförelse. I detta fall beräknades måtten enbart för rasterceller med minst 10 returer. Celler med färre returer, framför allt vattenytor och myrar, har färgats röda. Vid beräkning av h_{95} användes bara punkter över 2 meter, medan alla punkter användes vid beräkning av vegetationskvoten. Fler celler saknar därför data i h_{95} -rastret. Det går även att automatiskt skapa polygonskikt utifrån de områden där h_{95} och V ligger inom valda intervall. Viktigt att komma ihåg är att h_{95} och V inte är detsamma som trädhöjd och kronslutenhet, utan bara mått som korrelerar med dessa.



Figur 17. a) h_{95} , b) V, c) kombination av h_{95} (gult) och V (blått) samt d) terrängkartan. Rött indikerar områden där lasermått inte har beräknats, på grund av för få returerna. Rastren bygger på laserdata från Lanatmäteriet. Karta: © Lantmäteriet, i2012/901.

5.3 Automatisk segmentering

Segmentering av en bild innebär att den delas upp i mindre delområden eller segment, där varje område är förhållandevis homogent med avseende på till exempel färg eller textur. På motsvarande sätt kan även raster med lasermått, till exempel höjdpercentiler, utgöra en grund för automatisk segmentering. Inom skogsbruket kan automatisk segmentering användas för beståndsavgränsning. Oftast behövs då även en viss manuell efterreditering. Segmentering används även för detektering av enskilda träd, vilket beskrivs i avsnitt 7.1.

Det finns flera strategier för segmentering. En är så kallad *region growing*, där idén är att två angränsande områden slås ihop om de är tillräckligt lika. Hopslagningen fortsätter tills något kriterium har uppfyllts, till exempel att *i)* segmenten uppnått en viss medelstorlek eller *ii)* ett visst antal iterationer har gjorts utan att någon hopslagning skett. Man kan också sätta en minsta och största tillåten segmentstorlek.

Ofta kan segmenteringsprogrammet använda data från flera olika dataskikt. Ett raster med höjdmått från laser är användbart för att skilja mellan olika höga bestånd och på motsvarande sätt kan vegetationskvoten användas för att skilja på olika täta bestånd. Flygbilder och satellitbilder innehåller information om färg som är användbar för att skilja mellan olika trädslag och marktyper. I en flygbild avbildas dock träden från olika vinklar i olika delar av bilden, varför segmentering direkt i ett ortofoto kräver program som tar hänsyn till detta för att resultatet ska bli homogent. Segmenteringar är svåra att utvärdera objektivt och ofta är det därför upp till slutanvändaren att avgöra om resultatet är tillräckligt bra.

6 Areabaserade skattningar

Med så kallade areabaserade metoder skattas skogliga variabler för grupper av träd inom ytenheter om ca 100-500 m². Det laserskannade området delas in i ett rutnät (raster) av kvadratiska celler, eller pixlar. Ett antal provytor, oftast cirkulära, med ungefär samma area som rastercellerna, inventeras i fält för mätning och beräkning av skogliga data. Vanliga variabler som beräknas är grundytbevåg medelhöjd, stamvolym, grundyta, stamantal och medeldiameter. För både rasterceller och provytor beräknas lasermått enligt beskrivning i kapitel 4. Laser- och fältdata från provytorna kopplas sedan ihop så att man utifrån laserdata kan skatta skogliga variabler för rastercellerna inom hela det skannade området. Resultatet kan sedan aggregeras till beståndsnivå genom att man beräknar medelvärdet av alla rasterceller som ingår i ett bestånd.

De idag vanligaste skattningsmetoderna är regression och k-MSN. En tredje användbar metod kallas Random Forests. Detta kapitel ger en översiktlig beskrivning av de tre metoderna och hur de brukar användas i praktiken.

Den ursprungliga ”areabaserade metoden” bygger på att sambandet mellan laserdata och skogliga data skattas med regressionsanalys. Skattningarna brukar baseras på enbart laserdata och man delar i regel inte upp de skattade variablerna per trädslag. Istället kan bestånd med olika dominerande trädslag skattas i olika strata. Regression ger väntevärdesriktiga skattningar, det vill säga att värdet på de skattade variablerna i genomsnitt blir rätt. Eftersom regressionsmetoden är modellbaserad och tillåter interpolering (och inom rimliga gränser även extrapolering) fungerar den med relativt lite fältdata. Provytorna måste dock vara någorlunda representativa – man bör till exempel inte ha med en övervikt av provytor som är ovanligt täta för sin höjd. En begränsning med regression är att varje variabel skattas separat, vilket gör att icke representativa kombinationer av skattade variabler kan uppstå.

k-MSN-metoden har blivit vanlig då man vill kombinera laser- och flygbildsdata för att göra trädslagsspecifika skattningar. Metoden bygger på imputering, vilket innebär att man till respektive rastercell ”lyfter in” skogliga data från provytor som har liknande egenskaper enligt laserdata. k-MSN kräver därför betydligt mer fältdata än regression, jämnt spritt över hela variationsvidden. En fördel med k-MSN är att den skattar många variabler simultant vilket ger mer naturliga relationer mellan skattade variabelvärden.

De areabaserade metoderna bygger på att statistiska mått (till exempel höjdpercentiler och vegetationskvot) för de laserpulser som träffar en provyta korreleras med skogliga data för provytan. Således behövs ”många laserpulser per provyta” för att till exempel höjdpercentiler ska kunna beräknas. Hur tät laserskanningen behöver vara beror på flera saker, bland annat provytestorleken. I praktiken har areabaserad skattning visat sig fungera bra även med relativt glesa data, ned mot 0,5 pulser/m², vilket gör att till exempel Lantmäteriets NH-data kan användas. En nackdel med NH-data är dock att den maximala öppningsvinkeln är upp till 20°. Vid skanning för skogliga ändamål rekommenderas ofta att öppningsvinkeln begränsas till 15° eftersom bland annat vegetationskvoten påverkas något vid högre öppningsvinklar.

Beträffande årstid så har skattningsresultaten i barrdominerade nordiska skogar blivit lika bra före lövsprickning och efter lövfällning, som under sommaren. Data från perioden med pågående lövsprickning eller lövfällning bör dock undvikas, eftersom mängden löv då kan variera inom projektområdet och mellan olika arter. Man bör även undvika att skanna när marken är snötäckt eftersom markhöjderna då blir felaktiga.

6.1 Insamling och bearbetning av fältdata

Provytor för fältdatainsamling läggs ut så att den skogliga variationen i det område som ska skattas finns representerad i fältmaterialet. En rent systematisk sampling, till exempel i form

av ett regelbundet rutnät med provytor över hela området, ger ofta för få ytor i bestånd av ovanlig karaktär. Ett sätt att få tillräckligt många provytor från alla typer av skog, utan att kostnaderna blir för höga, kan istället vara att först stratifiera området enligt till exempel trädslag och ålder. Sedan läggs ett visst antal ytor ut i varje stratum enligt någon objektiv metod. Fält- och fjärranalysdata (laser och flygbilder) bör samlas in nära varandra i tiden, gärna samma år. Det finns dock flera fördelar med att vänta med insamlingen av fältdata tills efter skanningen är gjord. Först då vet man vilka områden som verkligen blev täckta med användbara data. Laserdata kan även vara ett stöd vid valet av fältreferenser.

Om provytornas fältmätta läge inte överensstämmer med laserdata kan skattningarnas kvalitet försämrats. Koordinatsättning av provytorna bör därför göras med DGPS (differentiell GPS). Vid DGPS används två GPS-mottagare: en som tas med ut i fält och en som placeras på en väl inmätt punkt i närheten. Den senare används för att mäta skillnaden mellan den position som ges av satelliterna och den kända positionen för punkten. Information om avvikelser skickas till den GPS som används i fält så att den uppmätta positionen för provytan kan korrigeras. Det är också vanligt att istället för referensmottagaren använda data från Lantmäteriets SWEPOS-nät, som är ett rikstäckande nät av referensstationer för GNSS.

I flertalet rapporterade försök med laserskanning av skog har man koordinatsatt provytorna med 1 meters noggrannhet eller bättre, se till exempel [16] och [18]. Positioneringsnoggrannheten är viktigare ju mer heterogen skogen är och ju mindre provytorna är.

Separat fältdata kan samlas in för utvärdering av skattningsnoggrannheten på beståndsnivå, men eftersom detta medför extra kostnader brukar man främst använda sig av korsvalidering på provytenivå, vilket beskrivs nedan.

6.2 Regression

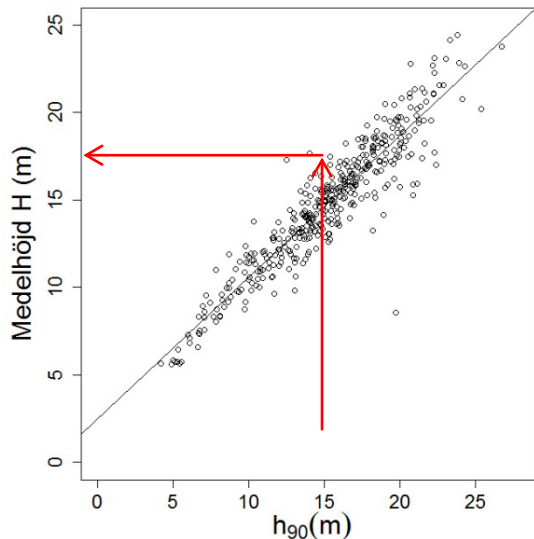
Regression är en statistisk metod för att finna ett samband mellan två eller flera slumpvariabler (stokastiska variabler). Metoden är modellbaserad, det vill säga den utgår ifrån att det finns en speciell matematisk modell som relaterar de två variablerna till varandra. Även om metoden bygger på en hel del teori är grundidén enkel. Att arbeta med regressionsanalys är något av ett hantverk, men det går att komma långt med ganska enkla modeller. Detta avsnitt inleds med ett exempel och en något förenklad beskrivning av regression, och beskriver sedan hur metoden används i praktiken.

6.2.1 Exempel på enkel linjär regression

Figur 18 visar ett spridningsdiagram med fältmätt grundyttevägd medelhöjd, H , på y-axeln och 90:e höjdpercentilen, h_{90} , på x-axeln. Varje punkt representerar en provyta där man både har fältmätningar och laserdata. I avsnitt 4.1 nämndes att det brukar finnas ett starkt samband mellan skogens höjd och h_{90} , vilket blir tydligt i spridningsdiagrammet där punkterna ligger väl samlade längs en linje. Genom att undersöka sambandet mellan medelhöjd och h_{90} kan man sedan skatta H i områden där fältdata saknas. Observera att h_{90} inte är detsamma som skogens medelhöjd, utan bara korrelerad med denna! Korrelation är ett mått på styrkan i ett linjärt samband mellan två variabler, och korrelationskoefficienten kan anta värden mellan -1 och 1. Om y ökar med x och punkterna ligger exakt på en linje är korrelationen 1 vilket innebär ett perfekt, positivt linjärt samband. Om y istället minskar med x , och punkterna ligger exakt på en linje, är sambandet perfekt negativt och korrelationen -1. Ju närmare 0 korrelationen är, desto svagare samband. I det dataset som använts i figur 18 är korrelationen mellan H och h_{90} 0,93 vilket innebär ett mycket starkt, positivt samband.

I diagrammet finns en provyta som tydligt avviker från de andra, ungefär vid $h_{90} = 20$ m och $H = 8$ m. Sådana uppenbara så kallade uteliggare bör i regel tas bort innan man gör

regressionsanalysen, eftersom de annars kommer att försämra modellen. Man bör dock verifiera att det finns sakskäl att ta bort uteliggaren. Orsaker till uteliggare kan vara exempelvis geometriska fel, avverkningar som skett mellan laserskanning och fältdatansamling, eller ren förväxling av provytor.



Figur 18. Spridningsdiagram med grundtyevägd medelhöjd, H , för provytor plottad mot 90:e höjpercentilen, h_{90} , i laserdata. Pilarna visar hur regressionsfunktionen (heldragen svart linje) kan användas för höjdsfattning.

Till att börja med måste man hitta den modell som bäst passar observerade data. I fallet med H och h_{90} ser den ut att vara en rät linje. Rätta linjens ekvation skrivs som

$$y = k \cdot x + m, \quad (5)$$

där k är linjens lutning och m är värdet på y i den punkt där linjen skär y -axeln. Punkterna i diagrammet ligger inte exakt på en linje; det finns en viss slumpmässig avvikelse eller spridning kring linjen. Vi skriver därför sambandet mellan H och h_{90} som

$$H_i = \alpha + \beta \cdot h_{90,i} + e_i, \quad (6)$$

där index i betecknar en viss provyta och e_i är avvikelsen från linjen för just den provytan. Med hjälp av regressionsanalys kan vi bestämma de värden på α och β som minimerar kvadratsumman av avstånden e_i , de så kallade residualerna, mellan mätpunkterna och linjen, under bivillkoret att medelvärdet av alla avvikelser är lika med noll. Ekvation 6 beskriver H som en funktion av h_{90} , och de två variablerna kallas för beroende respektive oberoende. H kan också kallas för responsvariabel och h_{90} för förklarande variabel.

Regressionsanalysen ger dessutom uppskattningar av hur starkt beroende de två variablerna är av varandra. Ett vanligt mått är förklaringsgraden. Denna storhet betecknas ofta R^2 och anger hur stor del av variansen i y -variabeln som regressionsmodellen beskriver. Förklaringsgraden är ett av de verktyg som brukar användas för att välja ut vilka variabler som ska vara med i modellen. Ett annat viktigt mått på modellens riktighet är spridningen kring regressionslinjen. När en tillförlitlig funktion har skattats kan den sedan användas för att skatta medelhöjden i områden där man har laserdata men saknar fältmätningar. Man kan dock inte bestämma H_i exakt – felet e_i förblir okänt.

Exemplet ovan tar upp enkel linjär regression. I många fall behövs flera oberoende variabler för att beskriva den beroende, och man talar då om multipel linjär regression. Man kan också behöva transformera de oberoende variablerna: kvadrera, multiplicera två oberoende med varandra och så vidare. Vissa samband är icke-linjära och kan lineariseras med hjälp av logaritmisk transformation innan regressionen görs. Detta tas inte upp här, och den som vill gå in djupare på teorin hänvisas till statistisk litteratur.

6.2.2 Regressionsmetoden i praktiken

Eftersom skattningsresultaten vanligtvis aggregeras till beståndsmedelvärden är det viktigt med en bra beståndsindelning. I Norge görs i regel beståndsindelningen och trädslagsbedömningen med manuell flygbildstolkning i digitala fotgrammetriska arbetsstationer, varefter övriga skogliga data skattas med stöd av laserdata. Beståndsindelningen kan också göras med stöd av segmentering av laserdata (avsnitt 5.3). Manuell efterreditering kan behövas och flygbilder är naturligtvis också användbara i detta arbete. Beståndsgränserna lagras i vektorformat i ett GIS.

Bestånden och provytorna delas ibland in i grupper efter till exempel trädslag, ståndortsindex eller ålder, så kallad stratifiering. Dessa egenskaper påverkar trädskronornas form, vilken i sin tur påverkar sambanden mellan lasermått och skogliga variabler. Skattningen kan därför få högre noggrannhet om den görs för mer homogena grupper, men det finns också exempel på försök där stratifiering inte har förbättrat resultatet.

När fältdatainsamling, eventuell stratifiering, och bearbetning av laserdata är klara tar regressionsanalysen vid. Det finns en mängd statistikprogramvaror som kan användas för att utföra beräkningarna, men för att resultatet ska bli bra krävs både statistik- och sakkunskaper. En regressionsfunktion tas fram för varje skoglig variabel som ska skattas, eventuellt uppdelat per stratum. Ett första steg är att bestämma vilken typ av modell som ska konstrueras. Residualerna ska vid regression vara jämnt fördelade. Om de ökar eller minskar med den oberoende variabeln så kan data behöva lineariseras, till exempel genom logaritmering.

Sedan väljs de oberoende variablerna ut. Man kan antingen börja med en och successivt lägga till nya, eller börja med många och stryka dem som inte är av betydelse. Den förstnämnda metoden är mer lätthanterlig, då man bör sträva efter en funktion med relativt få oberoende variabler. Man måste också fundera på vilka lasermått som rimligtvis kan ha ett samband med de variabler som ska skattas. Till exempel är det troligt att virkesvolym hänger samman med både höjd- och täthetsmått. Automatiska metoder, exempelvis *stepwise regression* eller *best subset regression*, kan ge vägledning i valet av variabler. Man bör dock inte lita blint på resultatet eftersom metoderna är rent matematiska och inte nödvändigtvis säger något om verkliga orsakssamband.

De ingående lasermåtten bör inte vara för starkt korrelerade med varandra, till exempel två närliggande höjdpercentiler. För varje variabel som lagts till eller tagits bort görs en ny, temporär, regressionsfunktion. Genom att analysera sambandet mellan residualerna från den temporära regressionsfunktionen och varje kandidat till ny variabel avgörs vilken variabel som ska läggas till närmast. Tabell 2 visar exempel på regressionsfunktioner som har använts för att skatta olika skogliga variabler.

De färdiga regressionsfunktionerna (en för varje skattad variabel och stratum) appliceras på rastret med laserdata för skattning av de skogliga variablerna. De resulterande rastren i kombination med beståndsgränserna används för att beräkna beståndsmedelvärden. Resultat från några svenska försök med regressionsmetoden visas i Tabell 3 [1].

Som nämndes i avsnitt 6.1 brukar man ofta använda så kallad korsvalidering för att utvärdera resultatet. Det innebär att man utelämnar en (eller ibland några) provytor åt gången, skattar funktionsparametrarna och använder dessa för att skatta ett värde för den eller de

provytor som utlämnats. Detta görs för alla provytor och skattningsresultatet jämförs sedan med fältmätningarna på ytorna.

Tabell 2. Exempel på regressionsfunktioner för skattning på beståndsnivå, från studier av Næsset [16] och Holmgren [17]. Från Næssets studie redovisas funktioner framtagna för mogen skog på goda boniteter

Studie	Variabel	Funktion ^a
Næsset	Medelhöjd, H	$\ln(H) = 0,35 + 0,529 * \ln(h_{90f}) + 0,355 * \ln(h_{maxf})$
Holmgren	Medelhöjd, H	$H = 1,46 + 0,95 * h_{95}$
Næsset	Stamvolym, V	$\ln(V) = 3,151 + 3,027 * \ln(h_{80l}) - 1,66 * \ln(h_{maxf}) + 1,223 * \ln(d_{50f})$
Holmgren	Stamvolym, V	$\ln(V) = -2,50 + 0,87 * \ln(D_v) + 1,49 * \ln(h_{90}) - 2,44 * relstd + 0,44 * D_p$
Næsset	Medeldiameter, D _g	$\ln(d_g) = 0,406 + 0,892 * \ln(h_{90f}) - 0,374 * \ln(d_{1f})$
Næsset	Stamantal, N	$\ln(N) = 10,33 - 0,487 * \ln(h_{0l}) - 0,667 * \ln(h_{cvf}) + 1,187 * \ln(d_{50f})$

^aSubskript _f och _l indikerar att måttet är beräknat för enbart första- respektive sistareturer. h_{max} är maxhöjden hos vegetationsträffarna, d_{50f} är antal förstareturer från trädkronor över h_{50} dividerat med totalt antal returer. D_v är antal förstareturer över 3 m dividerat med totalt antal returer. $relstd$ är standardavvikelsen i höjd dividerad med h_{95} . $d_p = (n_1+n_3)/(n_1+n_2)$ där n_1 är antal endareturer, n_2 är antal förstareturer och n_3 är antal förstareturer där andrareturen är över 3 m.

Tabell 3. Resultat från några svenska försök med regressionsmetoden, utvärderade på beståndsnivå och kompenserade för samplingsfel [1]

Skattad variabel	Medelfel (%)
Volym/ha	6-14
Grundtytevägd medeldiameter	7-13
Stamantal/ha	12-24
Grundtytevägd medelhöjd	3-6

6.3 k-MSN

Vid regressionsanalys antar man att det finns ett matematiskt samband som beskriver förhållandet mellan beroende och oberoende variabler. I så kallade icke-parametriska metoder görs inte det antagandet och man tar inte fram någon modell. Ett exempel på en sådan grupp metoder är imputering. En rastercell tilldelas den trädhöjd, volym etc. som uppmätts på den provyta som mest liknar rastercellen. Likhet mäts här i de oberoende variablerna, det vill säga lasermått, färg från flygbilder och så vidare.

k-MSN (*k Most Similar Neighbors*) bygger på imputering. I k-MSN, och andra liknande metoder, väljer man ut ett antal provytor som liknar rastercellen och beräknar ett viktat medelvärde av dessa. k står för det antal provytor man använder, och värdet kan variera mellan 1 och ca 20 beroende på syftet [18], [20]. Ett högt k ger bättre genomsnittliga skattningsresultat, men kan ge mindre realistiska relationer mellan de skattade variablerna på en given yta. För att välja ut grannarna används kanonisk korrelationsanalys. Denna del är i sig en parametrisk metod, som bygger på att den linjära kombination av beroende variabler (här skogliga data) som är mest korrelerad med en kombination av oberoende data (till exempel lasermått) räknas ut. Detta ger vikter för de olika lasermåtternas betydelse vid imputeringen. Den som vill veta mer hänvisas till vetenskaplig litteratur på området [21].

Alla beroende variabler skattas samtidigt vilket gör att de naturliga förhållandena mellan skattade variabler bevaras bättre än vid regression. Vid skattning med k-MSN används ofta mått från både laserskanning och flygbilder. Flygbilderna kompletterar laserdata genom att färgen innehåller mer information om trädslag än vad kronhöjdsfördelningen gör.

Flygbilderna har ofta en pixelstorlek på några decimeter, vilket innebär att varje provyta och rastercell innehåller många flygbildspixlar. Det finns olika sätt att använda informationen i flygbilderna. En metod är att för en provyta eller rastercell beräkna medelvärdet av de ingående flygbildspixlarna. Det görs för varje spektralband (färg) separat. Man kan dessutom använda kvoter mellan olika färger, till exempel kvoten mellan nära infrarött och rött ljus, och olika texturmått. Ett problem vid automatisk bearbetning av flygbilder är dock att trädens solsida dominerar i den ena delen av bilden och skuggsidan i den andra, och att större delar av trädkronorna projiceras i bildens ytterkanter än i dess centrum. Automatisk användning av flygbilder kan därför fordra avancerad bildbearbetning för att bli bra. Ett första steg är att använda metoder som skiljer på den spektrala informationen som kommer från den solbelysta respektive den skuggade delen av trädkronan. Valda mått från flygbilden används tillsammans med lasermått som oberoende variabler i skattningen.

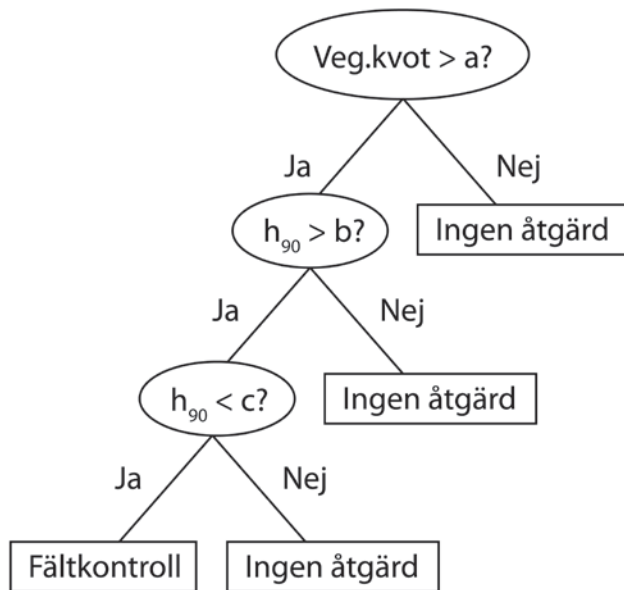
Ett annat alternativ är att för varje laserpunkt beräkna var den skulle ha avbildats i flygbilden, och färga laserpunkten efter den pixeln. Punkter som ligger i trädslagsrena ytor tillhör med stor sannolikhet ett bestämt trädslag. Sådana punkter används för att klassificera även de punkter som återfinns i blandade ytor. Eftersom det inte är meningsfullt att klassa markträffar efter trädslag kan man välja att enbart klassa förstareturer över en viss höjdröskel.

Det är viktigt med ett stort antal provytor jämnt spridda över hela variationsvidden, eftersom det för varje rastercell måste finnas flera nära grannar bland provytorna. Ju större variationen är i skogen, desto fler provytor behövs. Extrapolering fungerar dåligt eftersom rasterceller som ligger utanför provytornas variationsvidd bara får grannar på "en sida".

6.4 *Random Forests*

[Random Forests](#) är en icke-parametrisk metod som kan användas för både skattning och klassning. Den utvecklades av Leo Breiman vid University of California, Berkeley och Adele Cutler vid Utah State University [22]. Notera att *forests* inte har med skog att göra, utan syftar på att metoden använder sig av ett stort antal klassnings- och regressionsträd (*Classification And Regression Tree, CART*). Varje träd ger ett skattnings- eller klassningsresultat, och slutresultatet beräknas som ett medelvärde (i skattningsfallet) eller genom "röstning" (vid klassning). Till fördelarna med Random Forests hör att man kan använda ett stort antal oberoende variabler, även sådana som är korrelerade med varandra. Random Forests har i praktiken visat sig ge nästan lika bra resultat som regression, utan att man manuellt behöver bygga någon modell av sambandet mellan oberoende och beroende data. Metoden har därför visst intresse för framtida automatiserade tillämpningar.

figur 19 visar ett exempel på ett enkelt klassningsträd. Exemplet är inte helt realistiskt (det finns bättre sätt att identifiera gallringsskogar), utan är främst konstruerat för att beskriva hur klassningsträd fungerar. Säg att vi är intresserade av att hitta gallringsskog. Vi tror att den är tät och ligger inom ett visst höjdintervall, och vi vill avgränsa sådana områden i ett GIS för att planera en kontroll i fält. Vi skapar klassningsträdet i figur 19 och använder det tillsammans med vegetationskvot, V , och den 90:e höjdpercentilen, h_{90} , i form av raster. Varje fråga, eller nod, i trädet gäller en variabel och kan besvaras med "ja" eller "nej". Trädet mynnar ut i ett antal slutnoder som var och en motsvarar en klass. I ovanstående exempel finns två klasser: "ingen åtgärd" och "fältkontroll". En klass kan återfinnas i flera slutnoder.



Figur 19. Exempel på ett enkelt klassningsträd som kan användas för att identifiera områden med tät skog ($V > a$) inom ett visst höjdintervall ($b < h_{90} < c$).

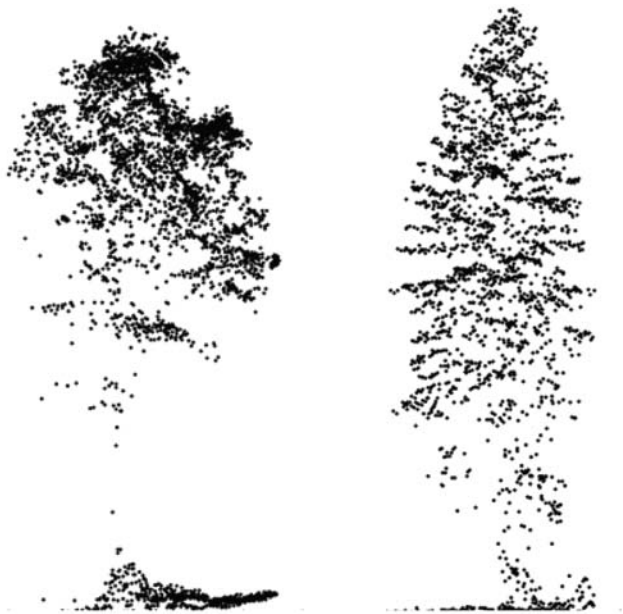
Vid skattning av skogliga variabler med Random Forests används inte klassning, utan regression. I regressionsfallet kan inte datamaterialet delas in i kategorier på det sätt som illustreras ovan. Istället konstrueras trädet så att varje slutnod innehåller en liten, relativt homogen grupp av observationer. Träningssdata, till exempel i form av fältinventerade provtytor, används för att bygga trädet. Skogliga data för en rastercell skattas sedan som medelvärdet av de träningsytor som tillhör samma slutnod som rastercellen. Random Forests använder flera klassnings- eller regressionsträd. Varje område som ska skattas körs genom samtliga träd, så att man får flera värden på virkesvolym etc. Slutresultatet beräknas som medelvärdet av resultaten från alla träd (i klassningsfallet är medelvärden meningslösa och resultatet avgörs istället genom "röstning").

Ett träd konstrueras genom att man tar ett så kallat *bootstrap sample* av provtytor, genom [dragning med återläggning](#). För att bygga första noden slumpas man fram m oberoende variabler att testa. Den variabel som bäst delar in samplet i två grupper väljs. Varje gren som går ut från noden leder till en ny nod, som konstrueras på samma sätt som den första. Trädet anses klart när antalet provtytor i varje slutnod nått under ett av användaren bestämt värde.

Skattningsresultatet kan utvärderas genom det så kallade *OOB Error Estimate* (OOB för *out-of-bag*). Det beräknas genom att de provtytor som *inte* ingått som träningsdata för att skapa ett visst träd körs en och en genom trädet för utvärdering.

7 Skattning av variabler för enskilda träd

Om pulstätheten i laserdata är tillräckligt hög får man flera returerna per trädkrona. Man kan då detektera enskilda träd och skatta variabler för dessa. Det finns naturligtvis ingen skarp gräns för hur många punkter som behövs, men ofta används omkring 10 pulser/m², även om lägre tätheter också kan fungera [13]-[15]. Figur 20 visar punktmolnsprofiler från tall och gran, skannade med ca 50 pulser/m².



Figur 20. Punktmolnsprofiler från tall och gran, pulstäthet ca 50 m². Bild: Johan Holmgren, SLU.

Förbehandlingen av data går till på det sätt som beskrivs i kapitel 3. Därefter kan arbetet grovt delas upp i följande steg:

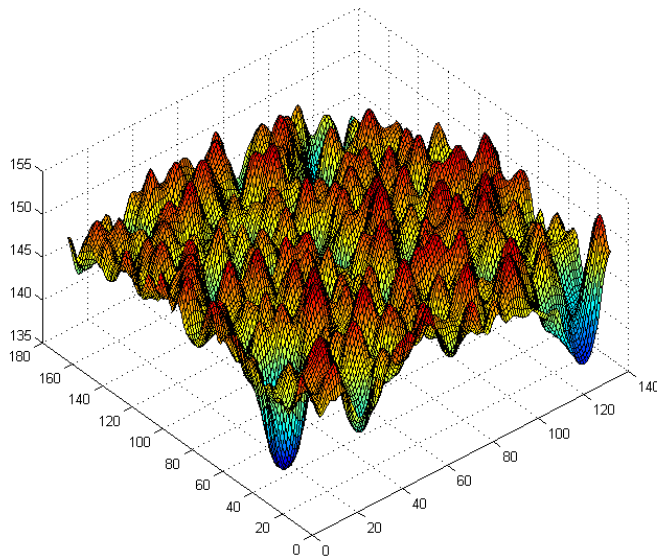
- Höjdsätt laserreturer över markmodellen.
- Detektera enskilda träd i laserdata och avgränsa trädkronor med hjälp av segmentering.
- Ta ut mått ur laserdata som beskriver de enskilda träden.
- Koppla samman träd i laserdata med träd som inventerats och koordinatsatts i fält.
- Ta fram regressionsfunktioner för variabler som ska skattas.
- Tillämpa funktionerna på alla detekterade träd.

Flera av bearbetningsstegen ovan kräver speciell programvara. Analys av enskilda träd i laserdata görs därför främst av forskare, samt av ett fåtal specialiserade företag.

7.1 Detektering av träd

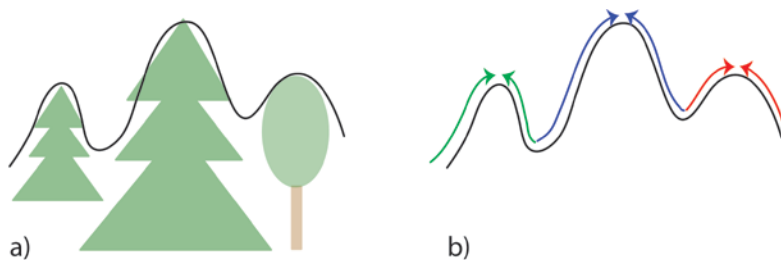
En vanlig metod för att detektera träden är att först skapa en kronhöjdsmodell (*Digital Canopy Model*, DCM), se figur 21. Lokala maxima (toppar) i kronhöjdsmodellen används som utgångspunkt i en segmentering. Målet är att varje trädkrona ska representeras av ett segment (för att inte missa några träd), och att ingen trädkrona ska ha fler än ett segment (för

att undvika ”falska” träd). Detta ställer krav på kronhöjdsmodellen, som i idealfallet ska ha exakt ett lokalt maximum per träd. En enkelt framställd kronhöjdsmodell som enbart följer de högsta punkterna i respektive rastercell uppfyller i regel inte detta villkor, eftersom den innehåller lokala höjdvariationer inom trädkronorna. Man använder sig därför av någon form av utjämnande filter. Filtret anpassas för att i möjligaste mån ta bort höjdvariationer inom trädkronorna utan att ta bort hela träd. De träd som ändå missas är ofta små som döljs av större. Två träd som står tätt ihop kan också tas för ett enda, medan träd med yvig krona och flera toppar kan tolkas som flera träd.



Figur 21. Kronhöjdsmodell med synliga trädtoppar. Bild: Åsa Persson, FOA.

En vanlig metod för segmentering av enskilda träd kallas *watershed segmentation* (watershed betyder vattendelare) och illustreras i figur 22. I varje rastercell över en viss höjd placeras ett ”frö”. Fröet får sedan klättra uppåt i den riktning som har brantast lutning, ända tills det når en punkt där alla omgivande celler har ett lägre värde. Denna punkt är ett lokalt maximum och tolkas som en trädtopp. Alla frön som klättrar till samma punkt kommer från samma segment, som antas motsvara utbredningen hos en trädkrona. Alternativt kan segmenteringen göras i tre dimensioner, genom någon form av klustring i punktmolnet eller genom att man använder så kallade voxlar, tredimensionella pixlar. Det finns ett par studier som jämför olika metoder för att detektera enskilda träd [14], [15].



Figur 22. Avgränsning av trädkronor med hjälp av watershed-segmentering. a) Kronhöjdsmodell filtrerad för att ge ett maximum per träd. b) I varje pixel i kronhöjdsmodellen sätts ett ”frö” som får klättra den brantaste vägen till en topp. Ett segment, eller träd, definieras som de pixlar vars frön klättrar till samma topp.

7.2 Mått kopplade till enskilda träd

Varje segment får ett ID-nummer, som också ges till de laserreturer som finns i segmentet. Man kan då koppla ihop alla laserreturer som hör till ett träd och beräkna olika mått för det trädet. Detta görs för alla träd. Om segmenteringen gjorts i två dimensioner, till exempel den ovan beskrivna *watershed*-segmenteringen, måste man dessutom hitta kronans nedre gräns för att inte få med punkter från undervegetationen. Bland fördelarna med laserskattning av enskilda träd märks att trädslag kan skattas från trädkronornas form, samt att bättre information om trädens storleksfördelning kan erhållas. Några mått som visat sig användbara för skattning av enskilda träd är

- trädhöjd och kronarea från segmenten
- höjdfördelningsmått i Tabell 1, beräknade för enskilda träd, samt andelen returer från trädkronan
- andel av olika returtyper, till exempel enda-, första- eller andrareturer från trädkronan
- texturmått från DCM och flygbilder
- intensitet (har dock ingen entydig fysikalisk betydelse om ingen kalibrering kan göras, eftersom den beror av många olika faktorer inklusive systemegenskaper)
- geometriska mått på trädkronans form, till exempel parametrar för en parabolisk yta som anpassas till en trädkrona.

Måtten används i senare steg för skattning av stamdiameter, stamvolym, trädhöjd och för trädslagsklassning. Andra mått som kan vara till nytta vid trädslagsklassning är färginformation från flygbilder [23]. Ett sätt är att projicera segmenten för trädkronorna på flygbilderna.

7.3 Sammankoppling av fält- och fjärranalysdata

Vid skattning med enskilda träd-metoder räcker det inte att känna till provytans position; man måste även veta de enskilda trädens placering relativt provytecentrum. Den kan till exempel mätas med ultraljudsutrustning kopplad till klaven (Haglöfs PosTex) eller med syftkompass. Positionen för provytans centrum mäts med DGPS.

För att mönstret med lasermätta träd och fältmätta träd ska stämma geometriskt kan speciella program användas som matchar fältytorna till laserdata på trädnivå. En metod går ut på att passa ihop två syntetiska bilder med varandra, och sammanfattas i följande punkter [24]:

- Skapa en syntetisk bild av trädmönstret enligt laserdata från området kring provytan.
- Skapa en syntetisk bild av trädmönstret på provytan utifrån fältmätningarna.
- Flytta och vrid ”fältbilden” så att den passar så bra som möjligt med ”laserbilden”.
- Koppla ihop lasersegment med fältmätta träd på så sätt att trädtoppsavstånden minimeras.

De syntetiska bilderna består av en mörk bakgrund där träden representeras av ljusare fläckar. Ljusheten bestäms i laserbilden av segmentens höjd, och i fältbilden av trädens uppmätta diameter eftersom man vanligtvis inte mäter höjden för alla träd i fält.

7.4 Skattning av trädvariabler

Skattning av höjd, diameter, volym etc. för enskilda träd kan göras med hjälp av exempelvis regression eller k-MSN, på liknande sätt som vid areabaserad skattning (kapitel 6). Regressionsmodeller beskriver sambandet mellan beroende (skogliga) och oberoende variabler (från laser eller flygbilder). Sambanden tillämpas sedan på alla detekterade träd inom det skannade området.

Bland fördelarna med laserskattning av enskilda träd märks att trädslag kan skattas från trädkronans form, samt att bättre information om trädens storleksfördelning kan erhållas. För att få väntevärdesriktiga resultat, det vill säga att virkesvolym, stamantal etc. i genomsnitt ska bli rätt skattade vid summering på exempelvis beståndsnivå, så måste man dock kompensera för de träd som inte upptäcks i laserdata. En metod som ibland kallas *semi-ITC* åstadkommer detta genom att alla fältmätta träd kopplas till ett segment, även om det i vissa fall innebär att ett segment får flera fältmätta träd knutna till sig [25]. Resultat från några svenska försök med skattning av enskilda träd redovisas i **Error! Reference source not found.** [1].

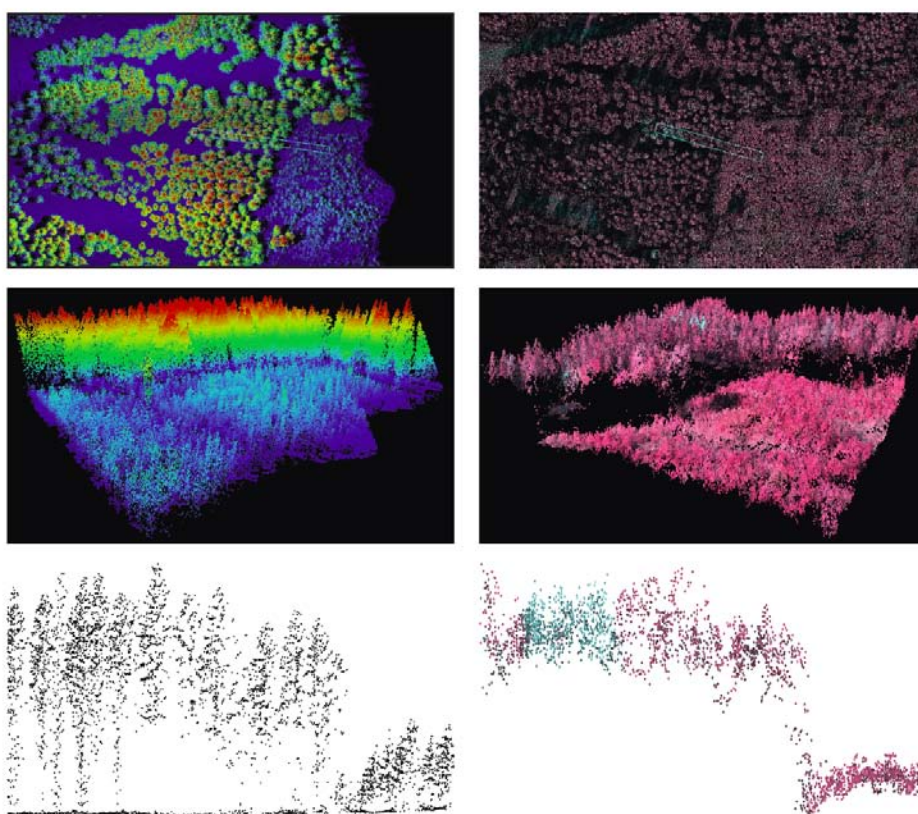
Tabell 4. Resultat från några svenska försök med skattning av enskilda träd, utvärderade på beståndsnivå och kompenserade för samplingsfel [1]

Skattad variabel	Medelfel (%)
Volym/ha	9-24
Grundytvägd medeldiameter	8-9
Stamantal/ha	6-33
Grundytvägd medelhöjd	1-2

8 Automatiserad digital fotogrammetri

Även flygbilder kan användas för att ta fram tredimensionell information om skogen. Bilder tagna i stereo kan användas för att skapa punktmoln liknande de från laserskanning, med två viktiga skillnader: *i)* punkterna från flygbilder innehåller även information om färg, vanligen blått, grönt, rött och nära infrarött, *ii)* eftersom kameran inte ”ser igenom” trädkronorna saknas i stort sett punkter från undervegetation och skogsbevuxen mark (figur 23). I Sverige finns god tillgång på flygbilder tack vare Lantmäteriets regelbundna fotografering, och med hjälp av den nya, nationella höjdmodellen (NH) kan punkternas höjd över marken beräknas trots att markpunkter saknas. Tekniken har ännu inte fått samma spridning som laserskanning i skogsbruket, men forskningsresultaten är lovande och intresset stort.

Inspiration till figurerna i avsnitt 8.2-8.4 har hämtats i Anders Bobergs *Introduktion till fotogrammetrin* [26], som även är en utmärkt källa för den som vill veta mer om fotogrammetriens grunder.



Figur 23. 3D-punktmoln från laserdata (till vänster) och från matchning av digitala flygbilder (till höger). Bilderna visar data från samma område. I profilerna (nedersta raden) är det tydligt att laserdata ger mer information om marken och trädskiktets vertikala struktur. Den digitala fotogrammetrin ger dock bra information om trädskiktets höjd och dessutom färginformation som i detta fall bl a visar döda träd som blå punkter. Bild av Jonas Bohlin, SLU, baserad på data från Lantmäteriet.

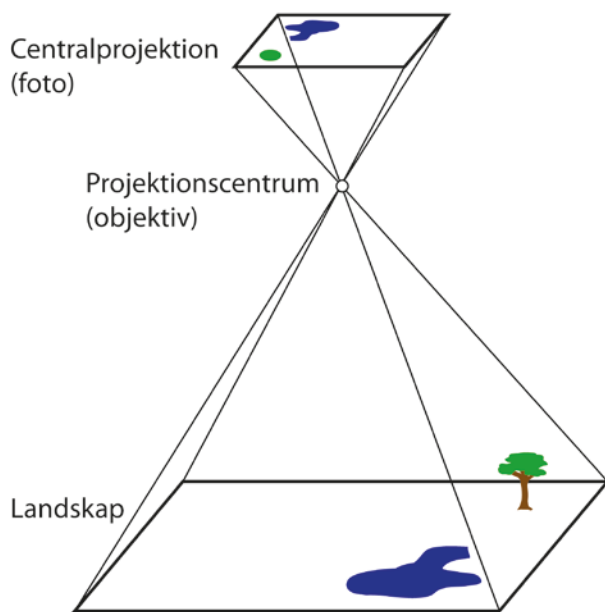
8.1 Färginformation från flygbilder

Digitala flygkameror registrerar oftast de fyra färgerna blått, grönt, rött och nära infrarött. Dessutom registreras ofta en pankromatisk kanal där ljus från en stor del av det synliga spektrumet används. Det pankromatiska bandet brukar ha mindre pixelstorlek och används för att förbättra den rumsliga upplösningen i bilden. Med färgerna blått, grönt och rött kan ”vanliga” färgbilder framställas där vegetationen framträder i grönt. En nackdel med dessa

bilder är att det blå bandet kan störas av dis i atmosfären, särskilt vid flyghöjder över 3000 m. Ofta används istället de gröna, röda och nära infraröda banden, men visade som blått, grönt respektive rött i de framställda bilderna. I dessa så kallade falskfärgsbilder framträder vegetationen i rött. En fördel med dessa bilder är att olika typer av vegetation lättare kan urskiljas i det nära infraröda våglängdsbandet.

8.2 Centralprojektion och ortofoto

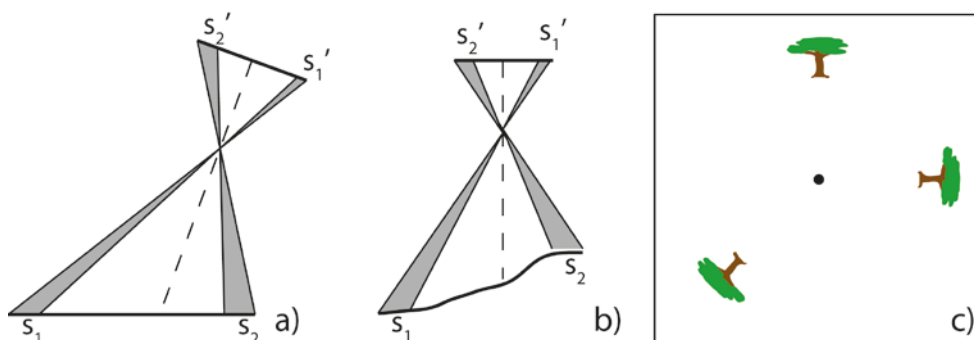
Den bild som kameran skapar av landskapet är en så kallad centralprojektion. Ljusstrålar från varje punkt i landskapet passerar genom kamerans objektiv (projektionscentrum) och fortsätter till detektorn. På så vis skapas en spegelvänd bild (figur 24).



Figur 24. Centralprojektion.

Om landskapet är plant och horisontellt, och kameran är riktad lodrätt mot marken, kommer centralprojektion (fotot) att vara en kartriktig bild av landskapet. I praktiken inträffar aldrig detta idealfall – kameran lutar, marken lutar och det finns höjder, sänkor, träd och byggnader i landskapet. Detta får till följd att bildens geometri inte stämmer överens med kartans. Figur 25 a) och b) visar hur skalan i bilden varierar på grund av höjdskillnader i terrängen och lutande kameraaxel. Höjder som träd, hus och berg ser ut att luta ut från bildens mitt, så kallad radiell förskjutning (figur 25 c). Förskjutningen är proportionell mot objektets höjd. På motsvarande sätt förskjuts sänkor in mot mitten.

Det finns flera olika kommersiella programvaror för digital framställning av ortofoton (kartriktiga bilder). Programmen räknar ut bildens läge pixel för pixel. Man kan tänka sig ortofotot som ett rutnät av pixlar som till att börja med saknar innehåll. Pixlarna i den ursprungliga flygbilden ska ges nya positioner i ortofotot. Det innebär att varje pixel i ortofotot ska fyllas med färg från någon eller några pixlar i flygbilden. Ortofotots pixlar har kända koordinater i planet. Höjden för motsvarande punkt på marken fås från en höjdmödel. Om kamerans position och vinkel är kända, kan man beräkna var i flygbilden denna del av landskapet har avbildats. Färgen från den delen av flygbilden överförs nu till den aktuella pixeln i det blivande ortofotot. Förutsättningarna för att göra ett digitalt ortofoto är alltså att man känner till kamerans orientering och har tillgång till en höjdmödel, alternativt kan skapa en ur flygfotona.



Figur 25. a) Kameraaxelns lutning och b) höjdskillnader i terrängen orsakar skalvariationer i bilden, så att de lika långa sträckorna s_1 och s_2 avbildas olika ($s_1' < s_2'$). c) Uppstickande objekt som höjder och träd förskjuts radiellt från bildens centrum.

8.3 Linjeskanning och cylinderprojektion

I den typ av digitalkamera som bland andra Lantmäteriet hittills har använt, liksom i analoga kameror, blir avbildningen en centralprojektion. Ett alternativ till den traditionella kameran är linjeskannern. Istället för att ta en sekvens av tvådimensionella bilder skannar den landskapet rad för rad. För att få stereobilder används ofta skannrar som tittar i två, eller oftare i tre riktningar: snett bakåt, rakt nedåt och snett framåt (Figur 26 a). Därmed fotograferas varje punkt på marken från tre vinklar. Skanningslinjerna är vinkelräta mot flygriktningen. Även här orsakar höjdskillnader och lutande kameraaxel skalskillnader. Liksom i fallet med centralprojektion förskjuts höjder och sänkor. Den lodrätt riktade skannern ger radiell förskjutning vinkelrätt mot flygriktningen (Figur 26 b). Framåt- och bakåtriktad skanner ger istället förskjutning snett ut mot sidorna. Förskjutningen är proportionell mot objektets höjd.

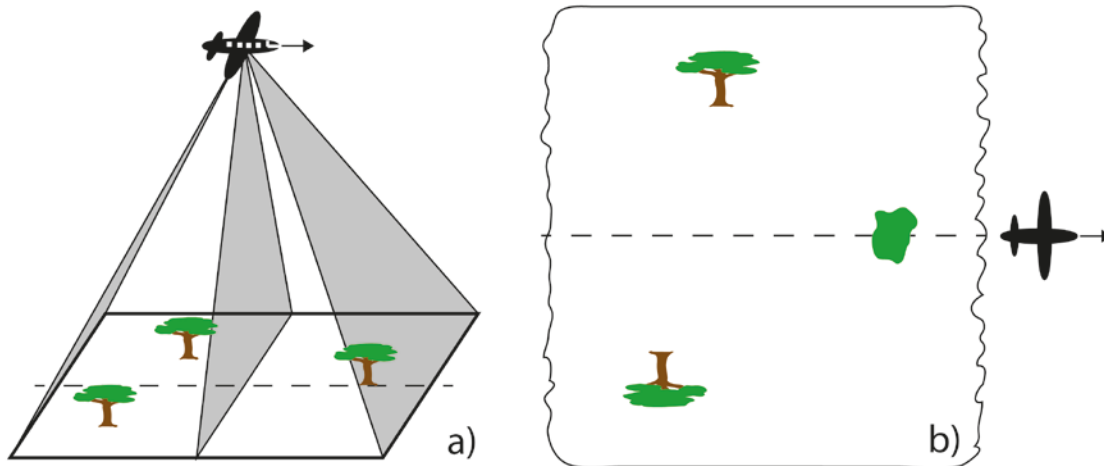
8.4 Stereofotogrammetri

Principen för stereofotogrammetri är densamma som för det mänskliga djupseendet: två bilder, från olika vinklar, kombineras till en tredimensionell modell med information om avstånd till olika objekt. Vanligen används minst 60 % övertäckning mellan bilderna, vilket med viss marginal ger stereotäckning i hela bildstråket. I de tvådimensionella bilderna förskjuts objekt ut från mitten, se figur 25. Förskjutningen är större ju närmare ögat eller kameran objektet befinner sig.

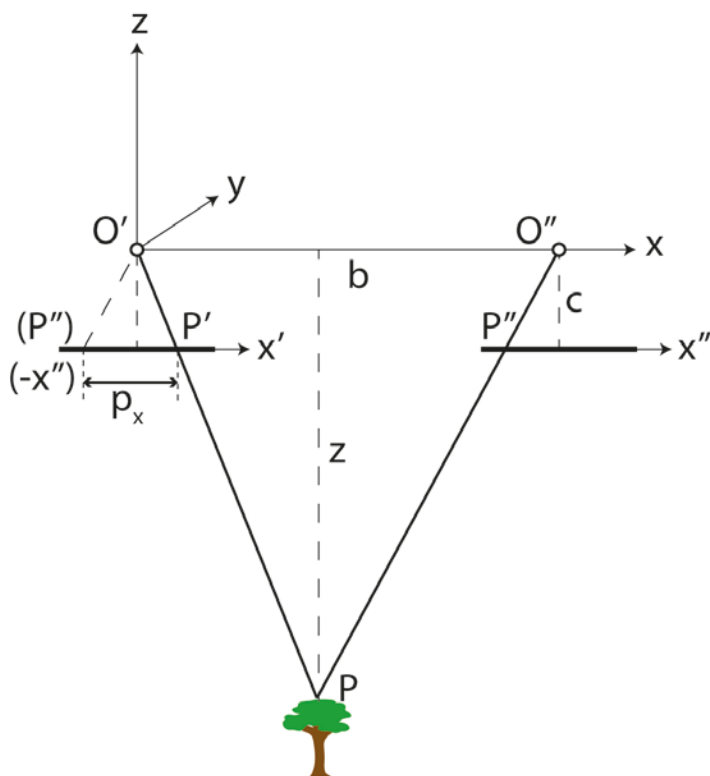
Ett objekt förskjuts på olika sätt i de två bilderna, eftersom de har olika centrum. Skillnaden i förskjutning kallas parallax och möjliggör stereoseende och avståndsbedömning. Flygbilder som används för stereotolkning överlappar varandra i flygriktningen. Ortofoton kan inte användas eftersom parallaxer orsakade av topografin har tagits bort. Numera finns speciella datorskärmar som i kombination med särskilda glasögon möjliggör stereotolkning av digitala flygbilder. Digitala fotogrammetriska arbetsstationer med sådana skärmar har ersatt tidigare analoga instrument för kartriktig tolkning av flygbilder.

Läget i plan och höjd kan beräknas analytiskt för objekt som syns i båda bilderna i paret. Det görs med hjälp av de så kallade parallaxformlerna. Figur 27 illustrerar hur parallaxformlerna härleds om kameraaxlarna är parallella med varandra och vinkelräta mot fotograferingsbasen (sträckan mellan de två punkter där kameran befann sig vid exponering). Det kallas för fotogrammetriens normalfall men är, namnet till trots, ett idealfall som i praktiken endast inträffar approximativt. I figuren har bildplanet ritats i så kallat positivläge, det vill säga på samma sida av projektionscentrum som det avbildade objektet (jämför med

figur 24 där bildplanet och det avbildade landskapet har ritats på varsin sida om projektionscentrum).



Figur 26. a) Linjeskanning i tre riktningar, där b) den lodrätt riktade skannern ger radiell förskjutning vinkelrätt mot flygriktningen. Framåt- och bakåtriktad skanner ger istället förskjutning snett ut mot kanten (visas ej).



Figur 27. En illustration av hur parallaxformlerna härleds. Koordinatsystemet xyz har origo i den vänstra bildens projektionscentrum, O' . Objektunkten P har koordinaterna x, y, z . Koordinatsystemen markerade med ' och '' har origo i respektive bildcentrum. Avbildningen av punkten P har koordinaterna x', y', z' i den vänstra bilden och x'', y'', z'' i den högra. Kamerakonstanten c är bildens z -koordinat och är negativ. Fotograferingsbasen betecknas b .

Likformiga trianglar ger att

$$\frac{x}{x'} = \frac{z}{-c}, \quad (7)$$

vilket i sin tur ger

$$x = x' \cdot \frac{z}{-c}. \quad (8)$$

p_x , x-parallaxerna, är definierade som $x' - x''$. Återigen ger likformiga trianglar

$$\frac{z}{-c} = \frac{b}{p_x}. \quad (9)$$

Ekvation 9 i 8 ger slutligen

$$x = x' \cdot \frac{b}{p_x}. \quad (10)$$

På liknande sätt kan man visa att

$$y = y' \cdot \frac{b}{p_x}, \quad (11)$$

och (eftersom y-parallaxerna i normalfallet är noll och $y' = y''$)

$$y = y'' \cdot \frac{b}{p_x}, \quad (12)$$

samt att

$$z = -c \cdot \frac{b}{p_x}. \quad (13)$$

Höjden z är objektets höjd över havet. För att bestämma dess höjd över marken måste man antingen mäta markhöjden på samma sätt som ovan, eller ha tillgång till en höjdmodell.

8.5 Punktmoln från stereobilder

Förutsättningarna för att punktmoln ska kunna genereras ur flygbilder är att *i*) bilderna överlappar varandra, *ii*) kamerans inre orientering är känd och *iii*) kamerans position och vinkel är kända med stor noggrannhet. Kamerans inre orientering mäts upp i laboratorium. Den externa orienteringen (position och vinkel) registreras med GNSS och INS (avsnitt 2.4), men dessa värden är oftast inte tillräckligt noggranna för att en bra bildmatchning ska kunna göras. Finjustering görs med hjälp av blocktriangulering. En vanlig metod för detta är *bundle adjustment*. Strålkärvarna (*bundles*) från objekt som automatiskt detekteras i båda bilderna används för att länka ihop bilderna till bildblock. Väl synliga och noggrant inmätta punkter i terrängen, till exempel målade, vita kvadrater eller brunnslock, används för att finjustera (*adjust*) orienteringen.

När förutsättningarna ovan är uppfyllda kan bilderna inom blocket matchas till punktmoln. Det finns några metoder för matchning och de ger lite olika resultat. Objektsbaserad matchning innebär att så många objekt som möjligt identifieras och matchas. Pixelvis matchning (*semi-global matching*) är en nyare metod som innebär att varje pixel matchas, vilket ger mycket täta punktmoln. Efter matchningen beräknas objektens x-, y- och z-koordinater med hjälp av parallaxformlerna (avsnitt 8.4). Resultatet blir ett punktmoln mycket likt det som fås vid laserskanning, men en viktig skillnad är att punkterna kan kompletteras med färginformation från flygbilderna. Enligt vilken princip detta görs varierar mellan olika programvaror för bildmatchning.

Eftersom antalet markpunkter är mycket lågt vid matchning av flygbilder, framför allt i skog, krävs i praktiken att man har tillgång till en befintlig höjdm modell.

8.6 Tillämpningar i skogsbruket

Ortofoton används allmänt som bakgrund till kartor i skogsbruket, liksom för indelning och ajourföring av skogsbruksplaner. Även kartering i fotogrammetriska instrument har använts, särskilt vid skogbruksplanering av större innehav. I den så kallade LMV-metoden används fotogrammetriska instrument för att göra indelningen av bestånd och mäta trädhöjden, samt för att uppskatta massaslutenhet och trädslagsfördelning. Höjden och massaslutenheten används sedan för att skatta virkesförrådet. Med införandet av digitala bilder och digitala fotogrammetriska instrument har detta arbetssätt blivit enklare. Det är dock inte så utbrett inom skogsföretagen utan används mest av konsultföretag.

Det finns vissa möjligheter att automatiskt bearbeta ortofoton, till exempel för segmentering av bestånd, men olika betraktningvinklar i olika delar av bilden fordrar att speciella program används. Dessutom innehåller en tvådimensionell bild endast begränsad information om skogens storlek.

Tekniken att genom matchning av flygbilder skapa 3D-punktmoln som avbildar krontaket på liknande sätt som laserdata innebär nya möjligheter att automatiskt få skoglig information från flygbilder. Detta är särskilt intressant för mindre fastigheter eftersom Lantmäteriet regelbundet fotograferar hela landet och distribuerar de digitala bilderna till en marginalkostnad.

Om punkterna görs om till rasterdata kan de användas för segmentering av bestånd med olika höjd och textur. Likaså bör det gå att använda 3D-modeller av detta slag för att upptäcka förändringar. Punktmoln från fotogrammetri kan även användas för skattningar av höjd, virkesförråd och grundyta på motsvarande sätt som med den areabaserade metoden. Ett nyligen rapporterat försök gav goda skattningsresultat då bilder från Lantmäteriets DMC-kamera med flyghöjden 4800 m användes (Tabell 5).

Tabell 5. Noggrannheter på beståndsnivå vid skattning av trädhöjd och stamvolym med punktmoln från Lantmäteriets digitala flygbilder, från Bohlin m fl 2012 [27]

Variabel	Använda mått i punktdata	Noggrannhet på beståndsnivå (RMSE)
Grundytevägd medelhöjd	h_{80}	8,8 %
Stamvolym	h_{80} och densitetsmått	14,6 %

Det försök som återges i Tabell 5 är mycket lovande eftersom det indikerar att skogliga skattningar kan göras automatiskt med relativt billig teknik, och med minst lika bra resultat som för manuella metoder. Förutsättningen är dock, precis som för laserbaserade skattningar, att det finns provytor med skogliga data. Skogen på den aktuella fastigheten har höga virkesförråd och är välkött. Försök av detta slag måste därför upprepas på fler fastigheter innan säkra slutsatser kan dras. Det torde dock stå klart att punktmoln från fotogrammetri är en mycket värdefull framtida datakälla för skogsbruket.

9 Lagring av 3D-punktdata

Vid laserskanning och digital bildmatchning genereras stora mängder data i form av punkter. Punkternas koordinater i plan (x , y) och höjd (z) måste lagras, eventuellt tillsammans med andra attribut för punkterna. Det finns flera tänkbara format för lagring av punktdata, till exempel kan man använda en textfil (ASCII) med en kolumn för varje koordinat/attribut och en rad för varje punkt. Alternativet är att spara data i binär form, ”ettor och nollor”. Utöver de enskilda punkterna vill man också spara information som är gemensam för alla punkter i filen, till exempel vilket geografiskt referenssystem (koordinatsystem) som används.

Tidigare använde olika dataleverantörer sina egna filformat, men allteftersom tekniken har blivit mer utbredd har ett behov av standardisering uppstått. *American Society of Photogrammetry and Remote Sensing* (ASPRS) publicerade 2003 den första versionen av det så kallade LAS-formatet [28]. I skrivande stund (januari 2013) är den senaste versionen LAS 1.4. Formatet är i första hand gjort för lagring av LiDAR-data, men kan också användas för punktmoln från matchning av digitala flygbilder (digital fotogrammetri). Specifikationen är fritt tillgänglig så att vem som helst kan konstruera program för att läsa och skriva LAS-filer.

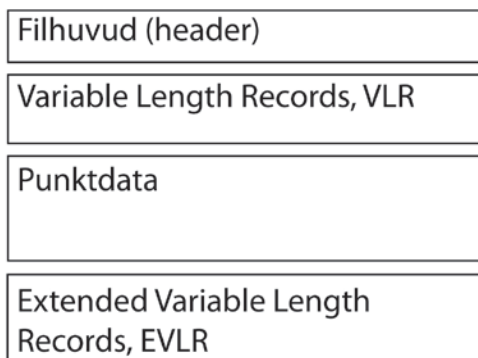
LAS-formatet är binärt. Binära filer behöver, till skillnad från textfiler, inte ”översättas” av datorn. Det går därför mycket snabbare att läsa och skriva binära filer än textfiler, vilket är av stor vikt under själva flygningen när stora mängder data snabbt måste sparas. Det är också en fördel vid databearbetning. Dessutom blir filstorleken mindre än om man hade använt text.

LASzip är ett programbibliotek för komprimering av LAS-filer till så kallat LAZ-format. Biblioteket har öppen källkod ([LGPL-licens](#)) och finns implementerat i flera kommersiella programvaror. Ofta blir en LAZ-fil bara en åttondel av storleken jämfört med om den sparats som en LAS-fil. [29]

9.1 LAS-filens uppbyggnad

Figur 28 visar schematiskt hur en LAS-fil är uppbyggd. De viktigaste delarna är filhuvudet (*header*) och den del som innehåller punktdata. Man kan lägga in ytterligare information i form av så kallade *Variable Length Records* (VLR) och *Extended VLR* (EVLR).

Filhuvudet innehåller uppgifter som gäller hela filen, bland annat geografisk utsträckning och antal punkter. Dessutom finns information som behövs för att filen ska kunna läsas korrekt. Punktdatadelen innehåller koordinater för varje punkt samt vissa andra attribut. Det finns olika format för punktdata och i princip kan man säga att skillnaden mellan dem består i vilka attribut som finns med. Några attribut, såsom intensitet, returnnummer och klass, finns alltid med. Andra attribut, till exempel färg, finns bara med i vissa punktdataformat. VLR och EVLR kan bland annat innehålla uppgifter om vilket geografiskt referenssystem som används (till exempel SWEREF99 TM).



Figur 28. Schematisk skiss över strukturen i en LAS-fil.

Eventuell klassning av punkterna, till exempel i mark- och vegetationspunkter, sker efter datainsamlingen. LAS-formatet har en uppsättning standardbeteckningar för klasser. De som har störst betydelse vid skogliga tillämpningar tas upp i Tabell 6.

Tabell 6. Några ofta använda klasser

Klass	Beskrivning
0	Aldrig klassad (har inte genomgått klassning)
1	Oklassad (har genomgått klassning men inte tilldelats någon klass, t ex punkter som inte klassats som mark i markfiltreringen)
2	Mark
9	Vatten

10 Referenser

- [1] Brethvad, T. & Iversen, H. E. (2011). Nyindelning av Bergvik Skog. [Elektronisk] I: *SkogsGIS*. Stockholm 24-25 april 2012. Tillgänglig: http://www.uli.se/images/stories/seminarier/presentationer/2012/skogsgis2012_brethvadi_iversen.pdf [2012-09-01].
- [2] Nelson, R., Krabill, W. & Maclean, G. (1984). Determining forest canopy characteristics using airborne laser data. *Remote Sensing of Environment*, vol. 15, ss. 201-212.
- [3] Aldred, A. H. & Bonnor, G. M. (1985). *Application of airborne laser to forest surveys*. Chalk River: Canadian Forest Service, Petawawa National Forestry Institute (Information Report PIX-51).
- [4] Maclean, G. A. & Krabill, W. B. (1986). Gross-merchantable timber volume estimation using an airborne LiDAR system. *Canadian Journal of Remote Sensing*, 12:7-18.
- [5] Nilsson, M. (1994). *Estimation of tree heights and stand volume using airborne lidar systems*. Umeå: Institutionen för skogstaxering, Sveriges Lantbruksuniversitet, Rapport 57.
- [6] Nilsson, M. (1996). Estimation of tree heights and stand volume using an airborne LiDAR system. *Remote Sensing of Environment*, 56:1-7.
- [7] Næsset, E., Gobakken, T., Holmgren, J., Hyypä, H., Hyypä, J., Maltamo, M., Nilsson, M., Olsson, H., Persson, Å. & Söderman, U. (2004). Laser scanning of forest resources: the Nordic experience. *Scandinavian Journal of Forest Research*, 19:482-499.
- [8] Toth, C (2010). Airborne LiDAR technology: the state-of-the art and future trends. [Elektronisk] I: ISPRS, *Latin American Remote Sensing Week 2010*. Santiago, Chile, 4-8 oktober. Tillgänglig: http://www.lars.cl/biblioteca/LARS_Thot.pdf [2012-09-01].
- [9] Wehr, A. (2009): LiDAR systems and calibration. I: Shan, J. & Toth, C. (red), *Topographic laser ranging and scanning*. Boca Raton: CRC Press, Taylor & Francis Group, ss. 129-172.
- [10] Olsson, P., Rost, H. & Reshetyuk, Y. (2011). Laserskanning. I: Lantmäteriet m fl, *Geodetisk och fotogrammetrisk mättings- och beräkningsteknik*, ss 181-204. http://www.lantmateriet.se/upload/filer/kartor/HMK/nyaHMK/pdf/Utbildning/Kompendium_20110929.pdf [2012-09-01].
- [11] Axelsson, P. (2000). DEM generation from laser scanner data using adaptive TIN models. I: *Proceedings of the XIXth ISPRS Congress, IAPRS, Vol. XXXIII*. Amsterdam, Nederländerna 16-23 juli 2000.
- [12] Kraus, K. & Pfeifer, N. (1998). Determination of terrain models in wooded areas with airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 53: 193-203.
- [13] Hyypä, J et al. (2001). HIGH-SCAN: The first European-wide attempt to derive single-tree information from laserscanner data. *The Photogrammetric Journal of Finland*, 17: 58-68.
- [14] Kaartinen, H. et al. (2012). An international comparison of individual tree detection and extraction using airborne laser scanning. *Remote Sensing*, 4:950-974.
- [15] Vauhkonen, J. et al. (2012). Comparative testing of single-tree detection algorithms under different types of forest. *Forestry*, 85:27-40.
- [16] Næsset, E. (2004). Practical large-scale forest stand inventory using a small-footprint airborne scanning laser. *Scandinavian Journal of Forest Research*, 19:164-179.
- [17] Holmgren, J. (2004). Prediction of tree height, basal area and stem volume in forest stands using airborne laser scanning. *Scandinavian Journal of Forest Research*, 19: 543-553

- [18] Packalén, P. & Maltamo, M. (2007). The k-MSN method for the prediction of species-specific stand attributes using airborne laser scanning and aerial photographs. *Remote Sensing of Environment*, 109:328-341.
- [19] Skogshistoriska sällskapet (2011-05-20). *Faktablad 05 – Kolbottnar och kolarkojor*. <http://www.skogshistoria.se/05KOLBOTTNOKOLARKOJOR.htm> [2012-09-01].
- [20] Junttila, V. (2011). *Automated, adapted methods for forest inventory*. Diss. Lappeenranta University of Technology.
- [21] Moeur, M. & och Stage, A. R. (1995). Most Similar Neighbor: an improved sampling inference procedure for natural resource planning. *Forest Science*, 41:337-359.
- [22] Breiman, L. (2001). Random Forests. *Machine Learning*, 45:5-32.
- [23] Holmgren, J. & Persson, Å. (2008). Identifying species of individual trees using airborne laser scanner. *Remote Sensing of Environment*, 90:415-423.
- [24] Olofsson, K., Lindberg, E. & Holmgren, J (2008). A method for linking field-surveyed and aerial-detected single trees using cross correlation of position images and the optimization of weighted tree list graphs. I: *Proceedings of SilviLaser 2008*. Edinburgh, UK 17-19 September.
- [25] Breidenbach, J., Næsset, E., Lien, V., Gobakken, T. & Solberg, S. (2010). Prediction of species specific forest inventory attributes using a nonparametric semi-individual tree crown approach based on fused airborne laser scanning and multispectral data. *Remote Sensing of Environment*, 114:911-924.
- [26] Boberg, A. (2001). *Introduktion till fotogrammetrin*. Stockholm: Kungliga Tekniska Högskolan, Institutionen för geodesi och fotogrammetri.
- [27] Bohlin, J., Wallerman, J. & Fransson, J. E. S. (2012). Forest variable estimation using photogrammetric matching of digital aerial images in combination with a high-resolution DEM. *Scandinavian Journal of Forest Research*, 27:692-699.
- [28] ASPRS Standards Committee (2007-05-01). *LAS specification version 2.0 proposal*. http://www.asprs.org/a/society/committees/standards/asprs_las_format_v20.pdf [2012-09-01].
- [29] Martin Isenburg (2013). LASzip – lossless compression of LiDAR data. *PE&RS*, 79:209-217.

Appendix: Förkortningar

ALS	Airborne Laser Scanning (flygburen laserskanning)
ASPRS	American Society of Photogrammetry and Remote Sensing
CART	Classification And Regression Tree (klassnings- och regressionsträd)
CHM	Canopy Height Model (kronhöjdsmodell), kallas ibland även Digital Canopy Model (digital kronhöjdsmodell, DCM)
DEM	Digital Elevation Model (digital höjdsmodell), kallas ibland Digital Terrain Model (digital terrängmodell, DTM)
DGPS	Differentiell GPS
DSM	Digital Surface Model (digital ytmodell)
FOA	Försvarets Forskningsanstalt (heter numera FOI)
FOI	Totalförsvarets Forskningsinstitut
GNSS	Global Navigation Satellite Systems (globala satellitnavigeringssystem)
GPS	Global Positioning System (globalt satellitpositioneringssystem från USA)
INS	Inertial Navigation System (tröghetsnavigeringssystem)
Laser	Light Amplification by Stimulated Emission of Radiation (ljusförstärkning genom simulerad strålningsemission)
LiDAR	Light Detection and Ranging (avståndsmätning med laser)
NIR	Nära infrarött (används om ljus i våglängdsområdet 700–1 300 nm, det vill säga just bortom synligt, rött ljus)
NH	Nationell Höjdsmodell, från Lantmäteriets pågående nationella laserskanning
SLU	Sveriges Lantbruksuniversitet

Appendix: Övningsdata och Programvaror

Övningsdata

För studenter med SLU:s användarkonto finns datamaterial till övningarna samt ytterligare instruktioner tillgängligt på servern

[\\storage-ume.slu.se\common\\$\student\geodata\Laserövningar](http://storage-ume.slu.se/common$/student/geodata/Laserövningar). Kopiera data från aktuell mapp och skapa respektive övningsmapp direkt på din hårddisk C: eller under C:\Fusion där du placerar data, ex C:\Fusion\Uppgift1. **Undvik å, ä, ö samt mellanslag i samtliga filnamn för att garantera sökvägar, gäller generellt under kursen.**

OBS: Data från webbverktyget GET samt övrig data som tillhandahålls under kursen får endast användas i studiesyfte och ej i kommersiella projekt.

Programvaror

Programvaror till övningarna är samtliga gratisprogram.

[Fusion](http://fusion.slu.se), version 3.30, instruktioner för installation finns i mappen *Installera programvaror* på servern geodata. <http://forsys.cfr.washington.edu/fusion/fusionlatest.html>

[R](http://cran.r-project.org/bin/windows/base/), version 3.01, för Windows. <http://cran.r-project.org/bin/windows/base/>

[RStudio](http://www.rstudio.com/ide/download/), version 0.97. <http://www.rstudio.com/ide/download/>

[OSGeo4W](http://trac.osgeo.org/osgeo4w/), programvara som innehåller Orfeo Toolbox med segmenteringsfunktion. Instruktioner för installation finns i mappen *Installera programvaror* på servern geodata. <http://trac.osgeo.org/osgeo4w/>

[Quantum GIS](http://www.qgis.org/en/site/) (QGIS), en alternativ programvara om du inte har tillgång till ArcMap. <http://www.qgis.org/en/site/>

[WinRar](http://www.win-rar.com/), komprimeringsverktyg för att packa upp och ner filer. Välj Download och inte Buy Now. WinRar har 40 dagars gratis testperiod. <http://www.win-rar.com/>

Diskutrymme samt operativsystemet

De flesta mjukvaror kan köras på Windows XP och Windows 7. Mjukvarorna kräver olika mycket diskutrymme; Fusion 22Mb, OSGeo4W nästan 1 Gb, QGIS 150 Mb, R 100Mb och RStudio 160 Mb. Det är även bra att känna till att laserdatafiler kan vara stora så det är bra att ha utrymme på disken som du jobbar med (ofta din C: disk), se till att du har åtminstone 1 Gb fri på aktuell disk.

Appendix: Övningsuppgifter

Uppgift 1 - välj ut ett laserskannat område som du vill arbeta med

Sedan 2009 pågår en rikstäckande laserskanning i Lantmäteriets regi. Huvudsyftet med skanningen är att skapa en ny, nationell höjdmodell (NH), men insamlade data kommer även till användning inom skogsbruket. I dagsläget finns data tillgängligt för omkring 80% av Sveriges yta. I den här övningen får du chansen att själv välja ut ett område att arbeta med i några av de kommande övningarna. För att hitta ett område som täcks in av skanningen och beställa hem laserdata använder du webbverktyget GET (Geodata Extraction Tool) från SLU. Du kan även beställa hem ett utsnitt av terrängkartan och ett ortofoto för att skaffa dig kompletterande information över området. Laserdata (.las) levereras som en rar-fil vilken du packar upp och arbetar vidare med i kommande övningar.

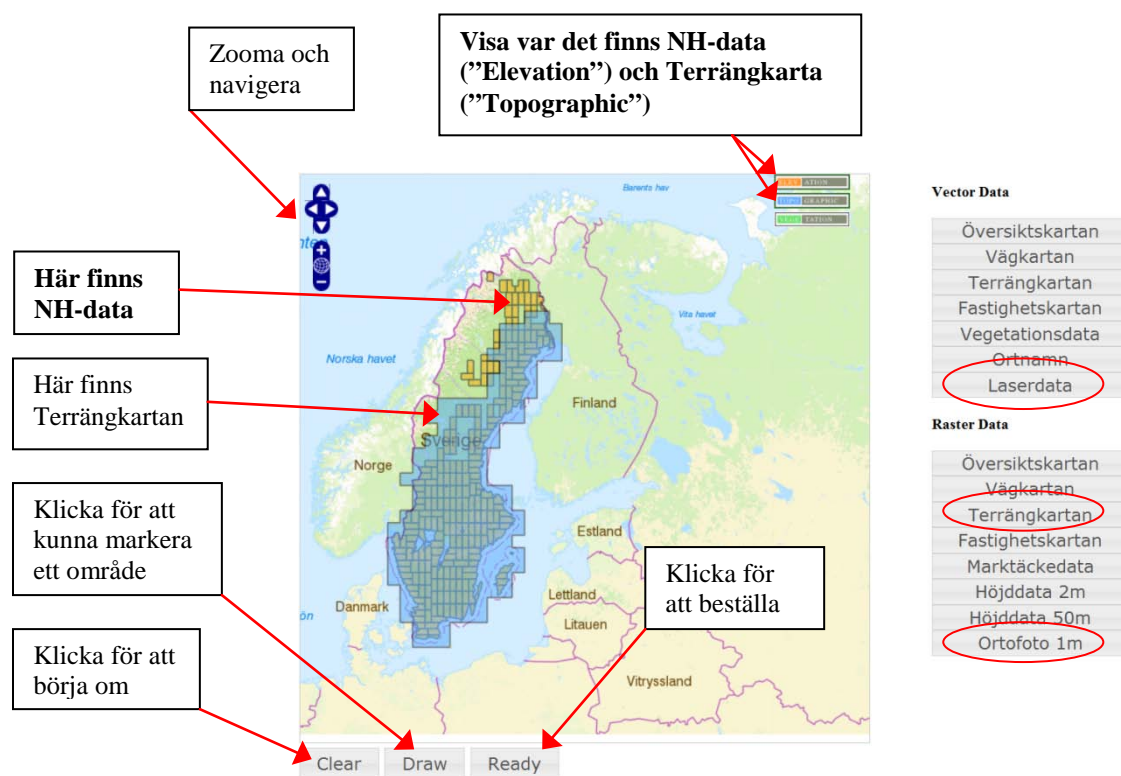
OBS: GET kan nås av studenter och anställda vid de flesta svenska universitet och högskolor. En förutsättning är att respektive universitet är anslutet till identitetsfederationen SWAMID. För att få tillgång till webbverktyget GET anger du ditt universitetskonto vid inloggning. Är du programstudent vid SLU utan tillgång till GET kan du behöva gå in på <https://idportal.slu.se> och verifiera dig. Om du inte har ett svenskt personnummer eller får problem med GET kontakta IT-stöd: it-stod@slu.se, 018-676600.

Nedan följer instruktioner hur du väljer ut ett område från GET och packar upp laserdata.

OBS: Data från webbverktyget GET samt övrig data som tillhandahålls under kursen får endast användas i studiesyfte och ej i kommersiella projekt.

Välj område

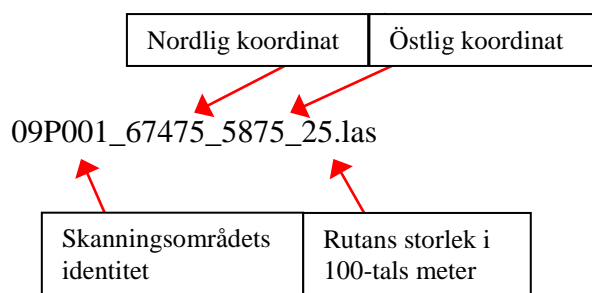
1. Gå in på <http://atlas.slu.se/get/>:



2. Använd knapparna uppe till höger för att visa var det finns både NH-data och terrängkarta. **Förväxla inte dessa – det viktiga är NH-datat/Laserdata!** Gå till ett område med NH-data som du vill arbeta med, gärna ett område som du känner till väl. Du kan flytta dig runt i kartan genom att dra och släppa, eller med hjälp av navigeringsknapparna uppe till vänster. Zoomar gör du med navigeringsknapparna eller musens scrollhjul. Zooma in på ditt område så att kartbilden blir några kilometer bred. Det finns ingen skalstock, så det får bli på ett ungefär.
3. Klicka på **Draw** under kartan och rita upp en rektangel som täcker ditt område, rektangelns sidor ska vara några kilometer långa. Rektangeln kan justeras genom att dra i hörnen. Vill du börja om klickar du på Clear och sedan Draw igen.
4. Laserdatat levereras i 2,5 km × 2,5 km block, och därför levereras alla laserdata block som finns inom ditt utvalda område (det kan vara flera .rar filer som du kommer att få).
5. Nästa steg är att välja vilken karta du vill beställa. Under Vector Data till höger om kartan klickar du på knappen **Laserdata** och under Raster Data kan du välja om du vill beställa hem **Terrängkarta** och **Ortofoto** (knapparna ska bli gröna). Om du skulle vara i ett område som har NH-data men saknar terrängkarta klickar du istället på **Översiktskartan** (mindre detaljerad topografisk karta). Ortofoton används senare i Fusion mjukvaran för att kunna visa laserdata, och det är därför praktiskt att nu ladda ner ortofoton över samma området som dina laserdata.
6. För att skicka beställningen klickar du på **Ready** under kartbilden och fyller i din mailadress i fönstret som kommer upp. Knappen **Sweref99 TM** ska vara grön, om den inte är det så måste du klicka på den. Sweref99TM är det koordinatsystem som används i

Sverige sedan 2007 (mer information finns på [Lantmäteriets sidor om geodesi](#)). Klicka sedan på **Start the process**.

7. Du bör inom kort få ett mail med en länk som du följer för att hämta hem datat. Data kommer i form av en zip-fil som du kan packa upp genom att högerklicka och välja Extract/ Packa upp. Det följer även med information om ditt nedladdade data, under mappen Docs beskrivs filernas innehåll utförligt.
8. Laserdata levereras packade i rar-format som komprimerar filstorleken avsevärt. För att packa upp rar-filerna finns ett gratisprogram, WinRAR, som du kan ladda ner från internet och sedan använda vid uppäckning.
9. I filnamnet på las-filen hittar du området identitet och koordinater för nedre vänstra hörnet samt rutans storlek i 100-tals meter.



Uppgift 2 - Visualisera punktdata

Starta Fusion. Finns ingen genväg i startmenyn eller på skrivbordet kan du söka genom att trycka på Windows startknapp och skriva Fusion. En kort presentation av Fusion, samt instruktioner för hur man installerar programmet finns på servern geodata\Laserövningar\Installera programvaror.

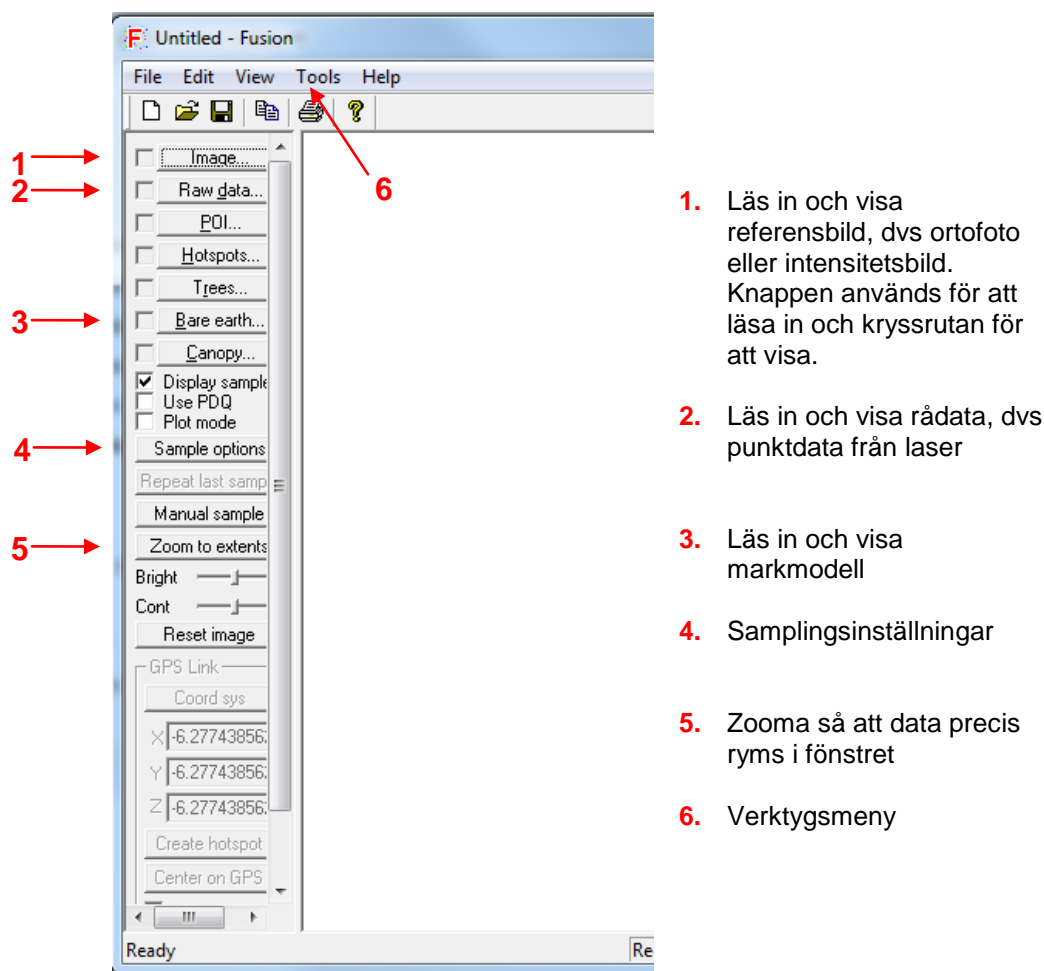
Fusion kan visa punktmoln, markmodeller, ortofoton, kronhöjdsmodeller med mera. Punktmolnen är oftast lagrade som .las-filer. För att kunna läsa in och visualisera punktmoln behöver Fusion en georefererad referensbild, till exempel ett ortofoto. Om ingen färdig bild finns tillgänglig kan man generera en ur laserdata. Bilden måste vara i samma projektion och samma enhet (oftast meter) som laserdata men behöver inte ha samma geografiska utsträckning, utan kan täcka ett större eller mindre område.

Denna övning innehåller följande moment:

- Generering av intensitetsbild
- Visualisering av punktmoln
- Inställningar för visning av punktmoln

Innan du sätter igång: Du kommer att få arbeta med laserdata från det område du valde i uppgift 1. Om du inte valde ut något område i uppgift 1 finns ett reservdataset på servern geodata i mappen Uppgift_2, (fullständig sökväg återfinns i appendix Övningsdata och Programvaror). Skapa mappen **uppgift_2** direkt under **C:** eller under mappen C:\Fusion (d.v.s. på datorns hårddisk) och kopiera din las-fil dit. **Undvik å, ä, ö samt mellanslag i samtliga filnamn för att garantera sökvägar, gäller generellt under kursen.**

Figur 1 visar hur Fusions huvudfönster ser ut.

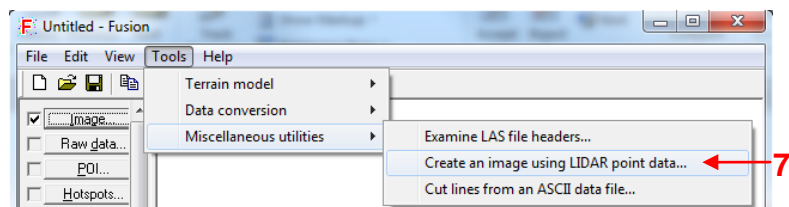


Figur 1. Fusions huvudfönster.

Generering av bild ur laserdata

Här skapar du den referensbild som behövs för att du sedan ska kunna läsa in punktdata (laserdata) i Fusion. Man kan välja att skapa bilden utifrån intensitetsvärdena eller höjdvärdena i laserdata. Här beskrivs hur man tar fram en intensitetsbild.

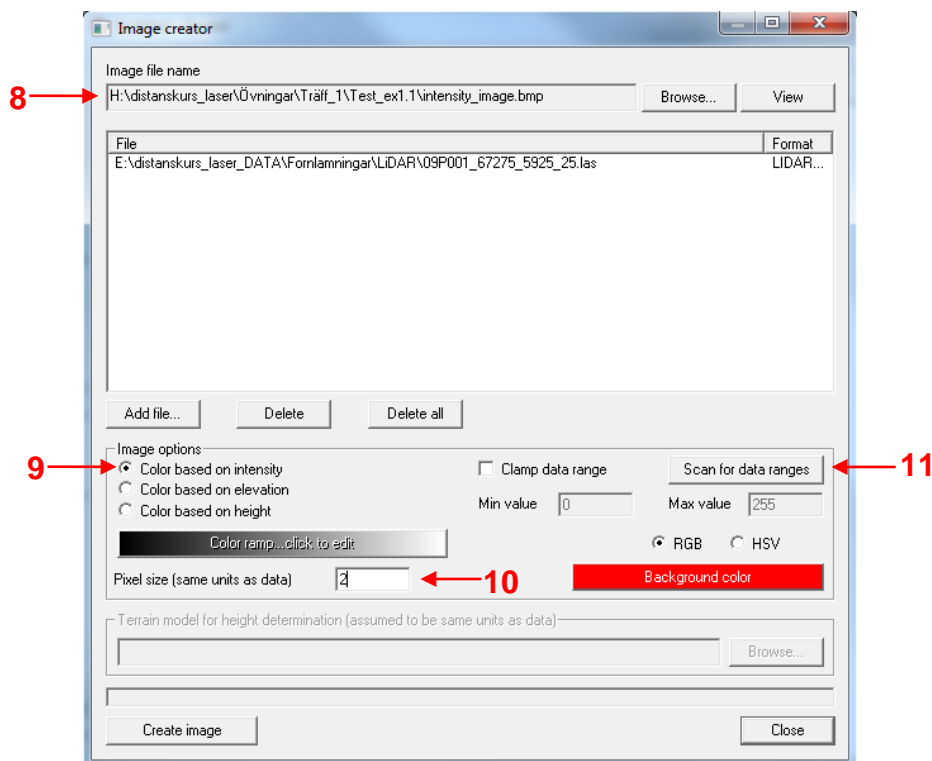
1. Gå till verktymsmenyn Tools (6) i den övre menyraden. Fortsätt till Miscellaneous utilities | Create an image using LIDAR point data (7, figur 2).



Figur 2. Välj att skapa en bild ur punktdata.

2. Navigera till C:\uppgift_2\ och öppna din .las-fil. Välj filen och klicka Open. Nu öppnas fönstret Image creator (figur 3).

- Under Image file name (8) anges var den färdiga bilden ska sparas. Klicka på Browse, navigera till C:\uppgift_2\ och skriv in filnamn, till exempel **IntImage_2.bmp** (Man ska egentligen kunna spara i andra format också, men det verkar vara någon bugg som gör att åtminstone .tif fungerar dåligt just här. Var därför noga med att ange .bmp som filändelse). Klicka Save.
- Under Image options, låt Color based on intensity (9) vara ikryssat. Bilden skapas då utifrån intensiteten hos laserreturerna.
- Skriv 2 vid Pixel size (10). Det betyder att bildens pixlar kommer att få sidlängd 2 meter. Mindre pixelstorlek ger en mer detaljerad bild, men också fler tomma pixlar eftersom fler pixlar kommer att sakna laserreturer.



Figur 3. Image creator.

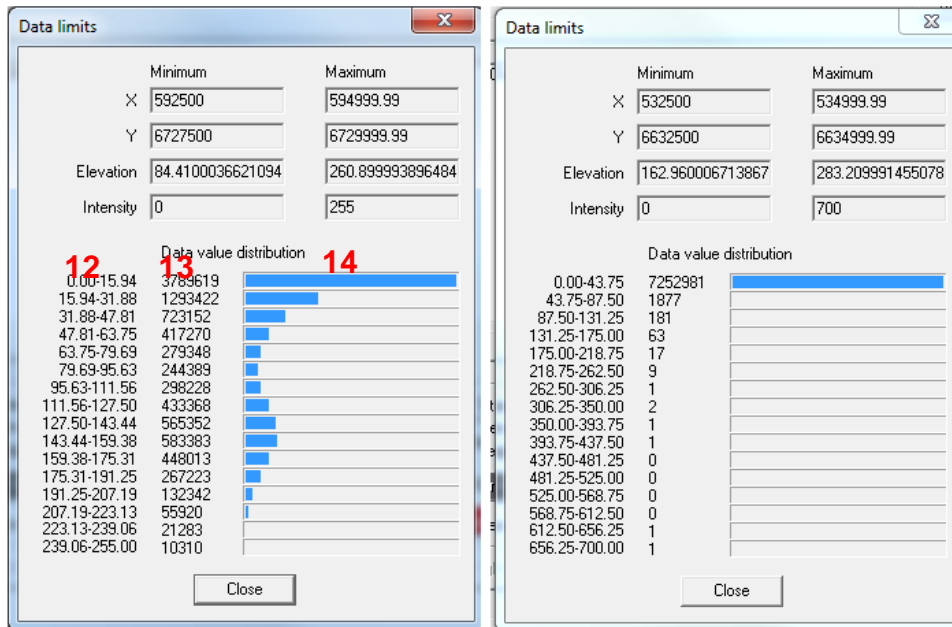
- För att hela gråskalan från svart till vitt ska utnyttjas på bästa sätt kan man behöva anpassa den till de intensitetsvärden som finns i bilden. Klicka på *Scan for data ranges* (11). Efter en stund får du upp ett histogram, där du kan se hur många returerna som faller inom olika intensitetsintervall. Figur 4 visar två exempel.

Den vänstra kolumnen (12) i figur 4 innehåller intensitetsintervallen och den högra (13) visar hur många returerna som faller inom respektive intervall. De blå stolparnas (14) längd är proportionell mot antalet returerna i intervallet. I det vänstra exemplet är intensitetsvärdena spridda mellan 0 och ungefär 223,13. I det högra har nästan alla returerna en intensitet mellan 0 och 43,75. Om man låter gråskalan representera intensiteter från 0 och hela vägen till 700 i det högra exemplet kommer hela intervallet 0-44 att hamna i den svarta änden och intensitetsbilden kommer att bli mycket mörk.

- För att anpassa bildens gråskala till de intensitetsvärden som förekommer, kryssar du nu i *Clamp data range* till vänster om (11). I rutan *Min value* skriver du sedan in den lägsta

förekommande intensiteten (0 i de två exemplen i figur 4). I *Max value* skriver du in den högsta intensiteten du vill visa. I det vänstra exemplet skulle det bli ungefär 224 och i det högra förslagsvis 44.

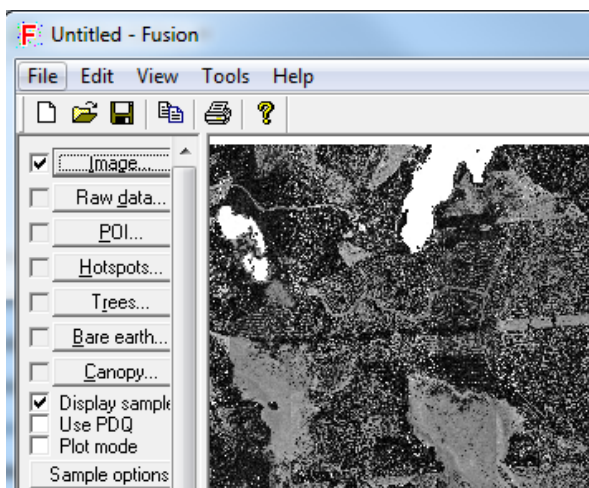
8. Klicka på Create image. När bilden är klar, klicka Close för att stänga fönstret.



Figur 4. Exempel på intensitetshistogram.

Inläsning av intensitetsbild

9. Klicka på Image-knappen uppe till vänster (1 i figur 1).
10. Navigera till den mapp där bilden ligger (C:\uppgift_2).
11. Vid Files of type: välj Image files.
12. Välj din bild (IntImage_2.bmp) och klicka Open.
13. Nu ska rutan intill Image-knappen vara ikryssad och fotot synligt (figur 5). Om inte, kryssa i rutan.



Figur 5. Intensitetsbild.

Fråga 1: Försök att se vad som är skog, hyggen och eventuellt annat i intensitetsbilden. Vad är ljusst och vad är mörkt? Förklara varför, med utgångspunkt i kompendiets avsnitt 2.7 samt stencilen NIR-ljus och Intensitet (finns på servern geodata, fullständig sökväg återfinns i appendix Övningsdata och Programvaror). Det är alltså en intensitetsbild, och laserljuset är i det nära infraröda området.

Tips: Du kan förmodligen behöva zooma in och ut, osv. Denna kurs är tänkt att vara ganska självständig, och det förväntas att du själv ska hitta många svar efter att ha letat i mjukvaran, hjälp och manualer. Sätt att göra saker och svar på frågorna är inte alltid beskrivet utan det är meningen att du ska experimentera själv också. I detta fall, för att hitta hur man zoomar in, kan du söka i FUSIONS manual (C:\FUSION\doc\FUSION_manual.pdf), där kommer du att hitta det mesta om FUSION.

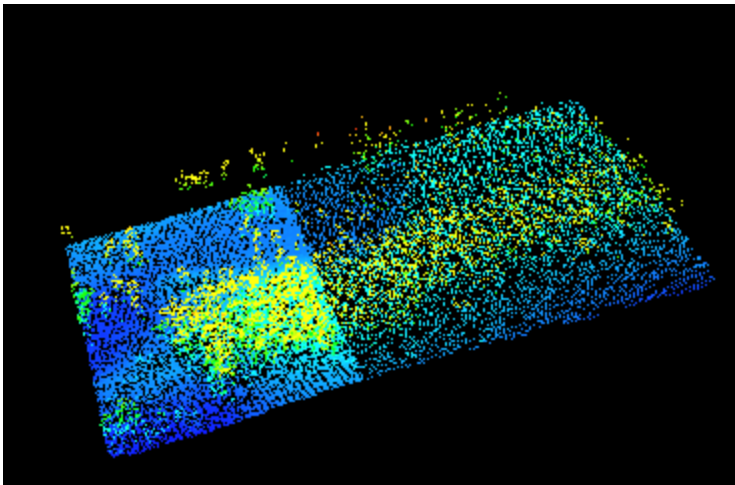
Fråga 2: Kanske ser du också några röda områden (i figur 4 råkar de vara vita på grund av de inställningar som användes när bilden gjordes). Det är områden som saknar laserreturer. Vad kan det bero på? (Tips: avsnitt 2.7 i kompendiet).

Inläsning och visualisering av punktmoln

14. Klicka på Raw data-knappen i Fusions huvudfönster (**2** i figur 1).
15. Navigera till C:**uppgift_2**\, välj .las-filen och klicka Open. (Har man flera las-filer kan man markera allihop och öppna samtidigt, men gör inte det nu).
16. I fönstret som kommer upp ska det stå "None" i Symbol-kolumnen, detta för att undvika att punkterna ritas upp i huvudfönstret. Att rita upp punkterna på detta sätt tar lång tid och är onödigt. Om det står något annat än "None", dubbelklicka där och välj None vid Symbol. Klicka på OK.
17. Spara ditt projekt nu, genom att gå till File | Save as i huvudfönstret. Spara projektet i C:**uppgift_2**\ tills vidare, inte på servern.
18. Se till att rutan vid Raw data-knappen (**2** i figur 1) är ikryssad. Håll ner vänster musknapp och dra upp en rektangel i bilden. Fönstret LIDAR Data Viewer (LDV) öppnas och visar punktmolnet. Hur lång tid det tar beror på datorn och på hur stort område som markerats. **Börja gärna med ett väldigt litet område för att testa.**
19. Genom att klicka i LDV-fönstret och hålla in vänster musknapp kan man vrida punktmolnet för att betrakta det ur olika vinklar.
20. Ta en stund för att vrida runt punktmolnet. Punkterna är färgade efter höjd. Punktmolnet finns lagrat i en .las-fil. Eftersom vi inte har läst in någon markmodell som vi kan subtrahera från punkthöjderna så är det höjd över höjdreferensytan, t ex havet, vi ser på skalan till vänster. Titta på några olika delar av området i huvudfönstret genom att dra upp nya rektanglar. Om det går väldigt snabbt att få upp punktmolnet i LDV-fönstret så kan du prova med ett större område. Titta på hyggen/myrar, enstaka träd, skogsbryn...
21. Du kan zooma med hjälp av musens skrollhjul. Om du tappar bort dig, högerklicka och välj "reset orientation". Du kommer då att se ditt område rakt uppifrån. Du kan också zooma genom att hålla in ctrl och vänster musknapp samtidigt som du flyttar muspekaren uppåt eller nedåt.

22. Du kan också välja att använda olika storlek och symbol för punkterna. I LDV-fönstret, högerklicka och välj "Marker" i menyn som kommer upp. Testa ett par olika markörtyper och storlekar. Välj en kombination som ger en tydlig bild.
23. I LDV-fönstret: högerklicka och välj "Image plate" i menyn. Du kommer nu att se ett utsnitt av intensitetsbilden under ditt punktmoln. Om din referensbild hade varit ett ortofoto skulle du istället ha sett det.

Fråga 3. I vissa delar av området kan du se en kant, där det på ena sidan är tätt mellan punkterna och på andra sidan glesare (figur 6). Vad beror det på? (Tips: avsnitt 3.1 i kompendiet).



Figur 6. Vissa områden (vänster) har högre punkttäthet. Varför? (fråga 3).

Inställningar för visning av punktmoln

Via knappen Sample options (4, figur 1) i huvudfönstret kan man ändra inställningarna för hur punktmoln visas. Klicka på den för att få upp fönstret "Sample options". Om du har ändrat en inställning och sedan vill titta på samma område som sist, klicka på knappen repeat last sample i huvudfönstret (under knappen sample options).

24. **Sample shape:** Med default-inställningen Stroked box håller man in vänsterknappen för att rita upp en rektangel (som du har gjort i ovanstående punkter). Med stroked circle ritar man istället en cirkel. Med Fixed box eller Fixed circle markeras en kvadrat respektive cirkel av bestämd storlek när man klickar i referensbilden. Storleken anges under Size. Testa att använda Fixed box och circle med olika diametrar/sidlängder. Du behöver då bara klicka en gång i bilden för att få upp punkterna i LDV-fönstret. Med Corridor drar du upp en linje i bilden. Då visas punkter i en "korridor" längs linjen. När du är klar: ändra tillbaka till Stroked box.
25. Du har kanske märkt att om du drar upp en för stor ruta så kan det ta riktig lång tid att rita upp punkterna. Med funktionen **Decimation** kan du ange hur många av alla datapunkter som ska visas. Ange siffran n för att visa var " n :te" punkt. På så sätt kan du titta på större områden än om varje punkt ska ritas upp. Testa olika "decimeringsgrad" och hur stora områden du då kan visa. När du är klar: Ändra tillbaka till 1.
26. **Returns:** Här anges vilka returter, eller rättare sagt returnnummer, som ska visas (om du är osäker på vad returnnummer betyder, titta tillbaka på avsnitt 2.7 och figurerna 6 och 7). Om enbart 1 kryssas för visas bara förstareturner, och så vidare. Testa att visa enbart första-, enbart andra- och enbart tredjere-turner (använd knappen repeat sample i huvudfönstret för att titta på samma som sist men med de nya inställningarna). Använd gärna ett område med både skog och öppen mark. Fundera över vad du ser! Kryssa sedan för allihop när du är klar.
27. **Color:** Här anges hur punkterna ska färgas. Ofta vill man färga punkterna efter höjd, och det är också standardinställningen.
28. Under **Color**, välj "Color by return number" och ta ett nytt "sample" med stroked box. Välj ett område där det finns både skog och öppen mark. På skalan till vänster i LDA-fönstret ser du att rött motsvarar 1, d. v. s. förstareturner, och gult motsvarar 2, d. v. s. andrareturner. Grönt, turkost och blått är tredje, fjärde respektive femtereturner. Om du tittar nära kan du kanske se några gröna punkter nära marken. Vrid punktmolnet och kolla från olika håll: från sidan, uppifrån och underifrån.

Fråga 4: I områden med öppen mark ser du troligen mest röda punkter. Varför?

Fråga 5: I områden med skog ser du troligen mest röda punkter i trädskronorna och gula på marken. Varför?

Inlämning

Svara på frågorna 1-5.

Extrauppgift: ortofoto som referensbild

29. Gå till <http://atlas.slu.se/get>. Beställ hem ett ortofoto från ditt område på samma sätt som du gjorde i uppgift 1. Det behöver inte ha exakt samma utsträckning, men gärna ungefär.
30. Ladda hem ortofotot och spara det i **C:\uppgift_2**.
31. Läs in ortofotot som referensbild på samma sätt som du gjorde med intensitetsbilden (punkt 7-11 ovan). Ortofotot kommer då att ersätta intensitetsbilden och bli din referensbild.
32. Ta ett sample med stroked box. I LDA-fönstret, högerklicka och välj Image plate. Du ser nu ortofotot under punktmolnet.
33. Du kan nu använda skrollhjulet för att höja och sänka fotot i förhållande till punkterna. (För att zooma får du istället använda ctrl och vänster musknapp).

Uppgift 3 - en första kontroll av datamaterialet

3D-punktdata från laserskanning, och även från digital fotogrammetri, lagras ofta i filformatet .las. De viktigaste delarna av en las-fil är headern, som innehåller allmän information om filen, och den del som innehåller punktdata där koordinater och annan information om varje punkt sparas. Mer om las-filer kan du läsa i kompendiets kapitel 9.

I Fusion finns möjlighet att titta på .las-filens header, som innehåller en sammanfattande beskrivning av vad som finns i filen (så kallad metadata). Det finns också ett verktyg som heter Catalog som ger en mer utförlig beskrivning av flera viktiga egenskaper hos datamaterialet, så som täckning och punkttäthet. Verktöget kan användas för en enkel kontroll av datamaterialets kvalitet. Catalog-verktyget körs från windows kommandoprompt.

Syftet med övningen är att

- du ska få en uppfattning om vilken information som finns i .las-filens header
- du ska lära dig att använda Fusions Catalog-funktion för en enkel kvalitetskontroll
- du ska få en introduktion till hur man använder batch-kommandon i Fusion.

Den sista punkten är viktig eftersom en stor del av Fusions potential ligger i funktioner som körs från kommandoprompten. Ingen fara om du inte har erfarenhet av detta tidigare – i övningsinstruktionen finns en kort introduktion till hur batch-filer och kommandon fungerar, följt av instruktioner steg för steg.

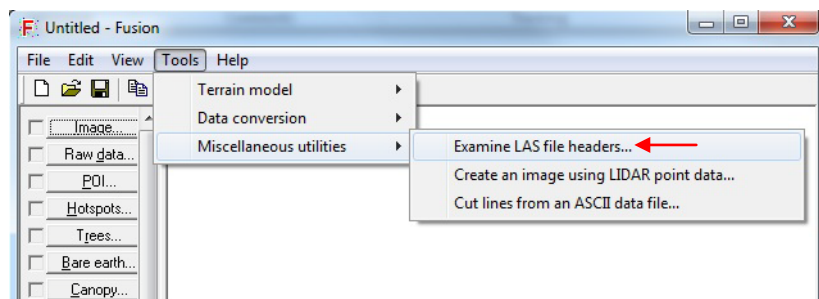
Du kommer fortsätta att arbeta med samma data som i uppgift 2, från det område du valde i uppgift 1. Om du inte valde ett område i uppgift 1 finns ett reservdataset på servern geodata (fullständig sökväg återfinns i appendix Övningsdata och Programvaror). Kopiera las-filen från mappen Uppgift_2.

Innan du börjar: Skapa en mapp för uppgift 3 på datorns hårddisk: **C:\uppgift_3** alternativt **C:\Fusion\uppgift_3**. Kopiera .las-filen som du använde i uppgift 2 dit.

Du behöver inte läsa in någon referensbild eller punktdatafil (las) för den här övningen eftersom du inte ska titta på punktdata i Fusions Lidar Data Viewer (LDV).

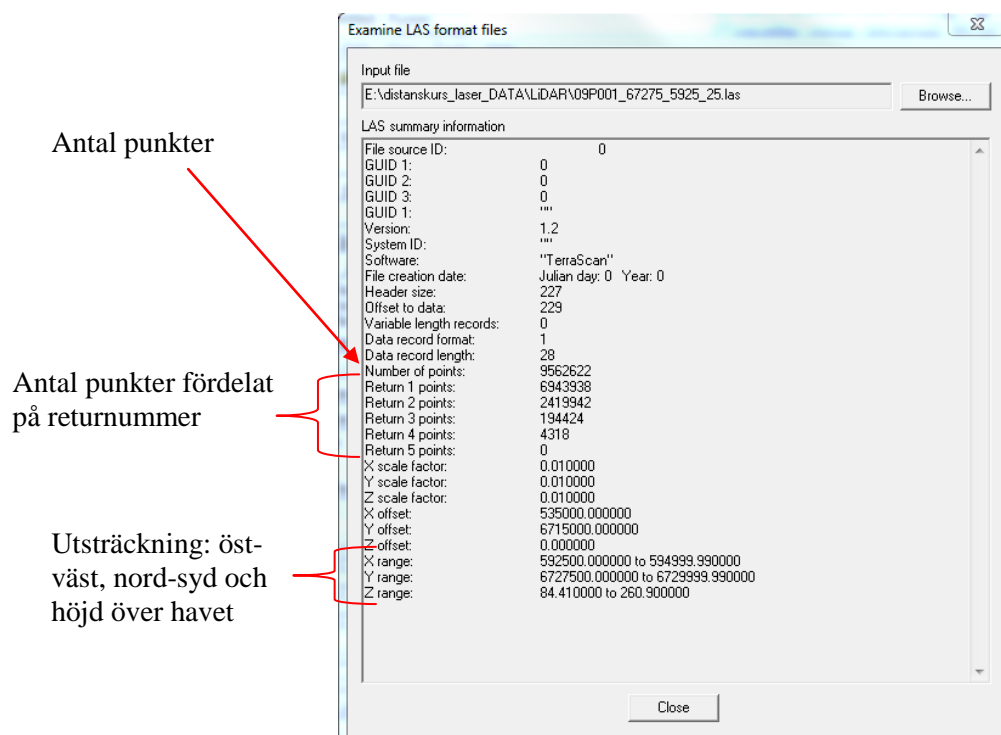
Visa header-informationen

1. Verktøyet Examine LAS file headers nås via verktøymenyn i Fusions hovedfönster (figur 1).



Figur 1. Examine LAS file headers nås via Tools i Fusions hovedfönster.

2. Navigera till mappen där din .las-fil ligger (C:**oppgift_3**) og dubbelklikka på filen. Header-informationen visar då i ett nytt fönster (figur 2)



Figur 2. Headerinformationen.

Fråga 1. En utsänd laserpuls kan ge opphov till 0, en eller flera retur (punkter). Hur många retur som maximalt kan registreras från en puls varierar mellan olika instrument. Titta på den del som visar antalet retur fördelat på returnnummer i din las fil. Hur många retur verkar det här instrumentet kunna registrera per puls?

Fråga 2. Hur stort är det skannade området (utsträckning i öst-västlig og nord-sydlig riktning, i km)? I det här fallet, med NH-data, används koordinatsystemet SWEREF99 TM, så enheten är meter.

Fråga 3. Vad är den lägsta respektive högsta höjden över havet för de returerna som registrerats i området?

Batch-filer och kommandotolken

Många av verktygen i Fusion körs via kommandon som kan skrivas in i Windows kommandotolk. Istället för att skriva in kommandon ett och ett i kommandotolken kan man också samla flera kommandon i en batch-fil och sedan köra hela filen. En batch-fil består alltså av ett eller flera kommandon. Den kan skapas och redigeras i en texteditor, exempelvis Windows "Anteckningar" (Notepad). Även om man bara ska använda ett enda kommando kan det vara bra att spara det i en batch-fil. Det gör det enklare att hålla reda på vad man har gjort, att göra ändringar och att återanvända kommandon.

För att **skapa en batchfil** skriver man in sina kommandon i texteditorn, går som vanligt till File | Save As, men döper filen till **filnamn.bat** och sätter filtyp till All Files (*.*)

För att **öppna en existerande batch-fil** för redigering måste man högerklicka på den och välja Edit.

För att **köra batch-filen dubbelklickar man på den**. Det går alltså inte att öppna filen för redigering genom att dubbelklicka!

Alternativt kan man också köra batch-filen genom att skriva in dess namn i Windows kommandotolk. Instruktioner om hur man gör det finns i slutet av övningen.

Catalog-verktyget

Fusions catalog-verktyg ger en sammanfattning liknande den i headern, men man kan också välja att få ut mer detaljer, till exempel bilder av punkttäthet och intensitet. Resultatet sparas dels i en html-fil som kan öppnas i en webbläsare, dels som separata bilder.

Verktyget kan köras direkt i windows kommandotolk (figur 3) eller via en batch-fil. **Det senare alternativet rekommenderas**, eftersom det är enklare att hålla reda på vad man har gjort och att göra eventuella ändringar om man har sina kommandon sparade i en fil.

För en fullständig beskrivning av Catalog-verktyget, se sidan 30 i Fusion-manualen (finns som pdf-fil i C:\FUSION\doc, samt på Fronters startsida). Det är starkt rekommenderat att du titta på detta medan/efter du läsa följande instruktioner.

I den här övningen ska du göra en batch-fil med ett kommando som anropar Catalog-verktyget, köra verktyget och titta på resultatet. Det finns även ett exempel på hur koden kan se ut i slutet av övningen.

3. Börja med att starta windows Notepad (Startmenyn | All Programs | Accessories | Notepad).
4. Nu ska du i texteditorn skriva in kommandot som anropar Catalog-verktyget. Syntax är `..\catalog [Switchar] [datafil] [catalog-fil]`.

5. Klammrar [] ska inte skrivas i batchfilen, de markerar bara sådant som ska fyllas i av användaren.
6. I varje batchfil börjar du kommandot med `..\` detta instruerar programmet att gå upp en mappnivå och leta efter verktyget (*catalog* i det här fallet) där.
7. *Catalog* är namnet på verktyget.
8. *[Switchar]* är instruktioner till Fusion för hur programmet ska köras. Varje switch föregås av ett snedstreck, `/`.
9. *[Datafil]* är namnet på las-filen som ska undersökas.
10. *[Catalog-fil]* är namnet man vill ge till den fil som skapas av programmet. Egentligen är *[Catalog-fil]* bara början på filnamnet eftersom flera filer skapas och automatiskt får olika ändelser.
11. I texteditorn skriver du nu in

```
..\catalog /density:4,1,4 /firstdensity:4,1,4
C:\uppgift_3\[DinLasfil] C:\uppgift_3\catalog_output
```

Allt ska stå på en enda rad i batch-filen, fast här ovan har vi visat den i två rader (för pedagogiska syften). Se även "Exempel på kod för catalogkommando" sist i uppgiften.

12. Switchen *density* används för att göra en bild av punkttätheten i området (punkter/m²). Här räknas alla returerna. Första siffran, 4, anger att punkttätheten beräknas för rasterceller om 4 m², d. v. s. 2 meters sidlängd. Andra och tredje siffran, 1 och 4, anger gränser för färgerna i den bild som skapas. Celler med färre returerna än 1 per m² färgas röda, celler med mellan 1 och 4 returerna/m² blir gröna och celler med fler än 4 returerna/m² blir blå.
13. Switchen *firstdensity*: som ovanstående men enbart för förstareturner.
14. `C:\uppgift_3\[DinLasfil]` är sökväg och namn på din lasfil. Byt ut *[DinLasfil]* mot filens namn (**inklusive** ändelsen `.las`)!
15. `C:\uppgift_3\catalog_output` är första delen av filnamnet hos de filer som kommer att skapas när du kör kommandot.
16. I texteditorn, gå till File | Save As. Navigera till din övningsmapp, **C:/FUSION**. Där skapar du en ny mapp som du döper till **batch**. Gå in i den, döp filen till **catalog_3.bat** och sätt filtyp till All Files (*.*). Klicka på Save.
17. **Alternativ 1:** I utforskaren i datorn, gå till mappen **C:/FUSION/batch** och dubbelklicka på filen **catalog_3.bat**. Nu bör kommandoprompten dyka upp och diverse text rulla (eller flasha snabbt) förbi. **Alternativ 2:** Använd kommandotolken, se instruktioner i slutet av uppgiften.
18. När det är klart, gå till **C:\uppgift_3** och dubbelklicka på filen **catalog_output.html** för att öppna den.
19. Om det inte har tillkommit några filer i din mapp (**C:\uppgift_3**) skriv in `pause` på en ny rad i batch-filen (**catalog_3.bat**) under den befintliga koden och kör filen igen. Nu ska kommandotolken stanna upp och du har möjlighet att se om du kan identifiera vad som är fel. (Kom ihåg att om du vill redigera filen, behöver du använda Notepad för att öppna den.)
20. Titta en stund på innehållet i filen och fundera över vad du ser!

Överst finns information om vilken version av Fusion och Catalog som använts, vid vilken tidpunkt filen skapats, vilket kommando som använts med mera.

Under *File Summary* visas las-filens geografiska utsträckning, antal returerna och punkttäthet.

Under *Classification Summary* visas antal returerna per klass. Mer om detta finns i kompendiets kapitel 9 (som ingår i läsanvisningarna först senare i kursen...). Kortfattat: det data vi använder kommer från Lantmäteriets NH-skanning. Vi har fått klassificerat data, där punkterna har delats in i klasserna mark, vatten och "oklassat". Klassen "oklassat" innehåller allt utom vatten och mark, det vill säga vegetation med mera. Det finns också kommersiella mjukvara (t.ex., TerraScan) som bearbeta och producera sådana klassningar.

Under *Return Density* finns information om minsta, största och genomsnittlig pulståthet i området. Bilden är uppbyggd av celler om 2*2 meter, eftersom vi satte cellarean till 4 m².

First Density visar samma sak som *Return Density* men enbart beräknat för förstareturner.

Fråga 4. Titta under *File Summary*. Hur många returerna har registrerats totalt?

Fråga 5. Titta under rubriken *Return Density*. Vad är det högsta antalet returerna per m² (square unit)?

Fråga 6. Vad är den genomsnittliga punkttätheten?

Fråga 7. Bilderna som visar antal returerna/m² har i regel band med tätare data i nord-sydlig riktning. Varför? (Ibland kan det också finnas en "kant" i öst-västlig riktning. Försök i så fall att förklara det.) Tips: Avsnitt 3.1 i kompendiet.

Fråga 8. Punkttätheten är oftast som allra högst i kanterna av dessa band. Varför? Tips: Avsnitt 2.6 i kompendiet.

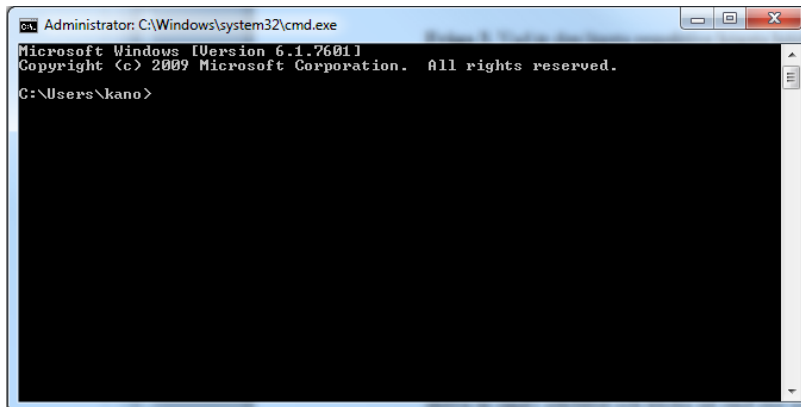
Fråga 9. Titta efter hur hyggen och skog ser ut i bilden med punkttäthet (return density). Vad har hög respektive låg täthet? Vad beror det på? Tips: Avsnitt 2.7 i kompendiet.

Inlämning

1. Svaren på frågorna 1-9.
2. Lämna också in den .html-fil som du skapade.

Windows kommandotolk - extrainformation

Istället för att dubbelklicka på batch-filen kan man välja att köra den via windows kommandotolk. I kommandotolken måste man vara i den mapp där filen ligger. Man ser i prompten var man befinner sig: i exemplet i figur 3 står jag i C:\Users\kano. För att flytta till mappen ovanför skriver man in `cd . .` och trycker enter. I exemplet skulle man då hamna i C:\Users. För att gå in i undermappen kano igen skriver man `cd kano`.

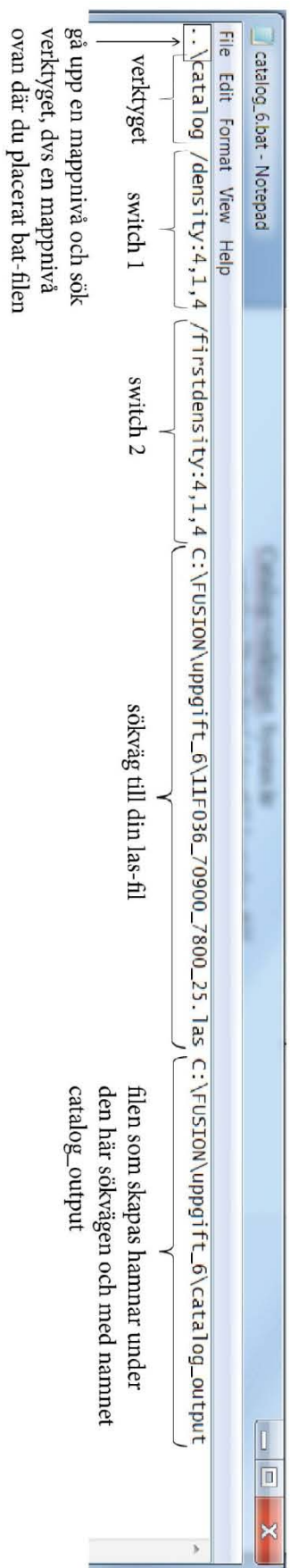


Figur 3. Windows kommandotolk, eller prompt.

För att köra batch-filen genom kommandotolken (alternativ 2):

Starta kommandotolken genom att gå till Windows startmeny och skriva `cmd`. Gå till katalogen C:\FUSION\batch (`cd . .` för att komma uppåt och `cd mapp` för att komma till mappen "mapp"). Skriv nu in namnet på din batchfil, `catalog_3.bat`, och tryck enter.

Exempel på catalogkommando i batch-fil



Uppgift 4 - Segmentering

Segmentering är ett sätt att automatiskt dela in en bild i mindre områden, så kallade segment. Områdena är mer eller mindre homogena med avseende på till exempel färg eller textur.

Vid laserskanning genereras ett stort antal punkter, koordinatsatta i plan (x,y) och höjd (z). Även intensiteten hos det reflekterade ljuset registreras för varje puls. Som vi har sett i Uppgift 2 ("Visualisera_punktdata") kan man skapa en bild av det skannade området utifrån intensitetsvärdena. Man kan även använda punkternas höjd för att skapa en bild. Området delas in i rasterceller, och varje cell tilldelas höjden hos exempelvis den högsta punkten i cellen. Bilden kan sedan visas i till exempel ArcMap genom att varje höjd motsvarar en gråton från svart (låga höjder) till vitt (höga höjder), eller någon annan färgskala. En sådan bild kan ligga till grund för en segmentering, där varje segment kommer att innehålla någorlunda jämnhögt skog.

Även ortofoton från flygfotografering kan användas vid segmentering. Färgen, särskilt i det nära infraröda området, påverkas av mark- och vegetationstyp, eftersom olika marktyper och vegetation reflekterar olika mycket av solens nära infraröda ljus. Ortofoton kan därför användas för att segmentera efter trädslag. Det är lättare i dessa bilder att skilja mellan löv och barr än att skilja mellan tall och gran eller mellan olika lövträd.

För att utnyttja både höjd- och färginformation kan man basera segmenteringen på en kombination av både laserdata och ortofoto.

Det finns idag några olika mjukvaror som kan göra segmentering, men de flesta är kommersiella produkter (t.ex., eCognition, Dianthus, Erdas). Det förekommer många olika algortimer för segmentering. I den här övningen kommer du att prova ett segmenteringsverktyg i ett gratis opensource program som kallas OrfeoToolbox, vilket körs via ett shell program som heter OSGeo4W.

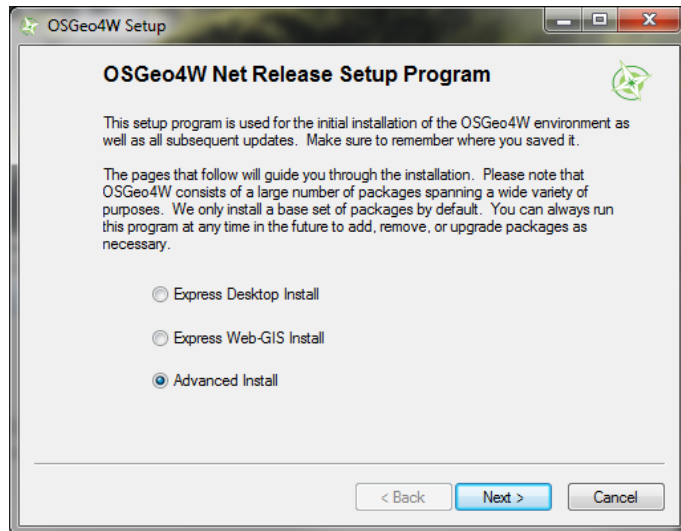
Inlämningen består av att du laddar upp shapefilerna med dina segmenteringsresultat (zippad mapp). **OBS:** varje shp-fil har tillhörande ytterligare tre filer (prj, dbf, shx) vilka samtliga måste skickas med shp-filen för att den ska gå att öppna i ArcMap.

Innan du börjar

Kopiera mappen **Uppgift_4**, (från servern geodata) till en mapp som du kallar för Uppgift_4 under **C:** alternativt **C:\Fusion\uppgift_4**.

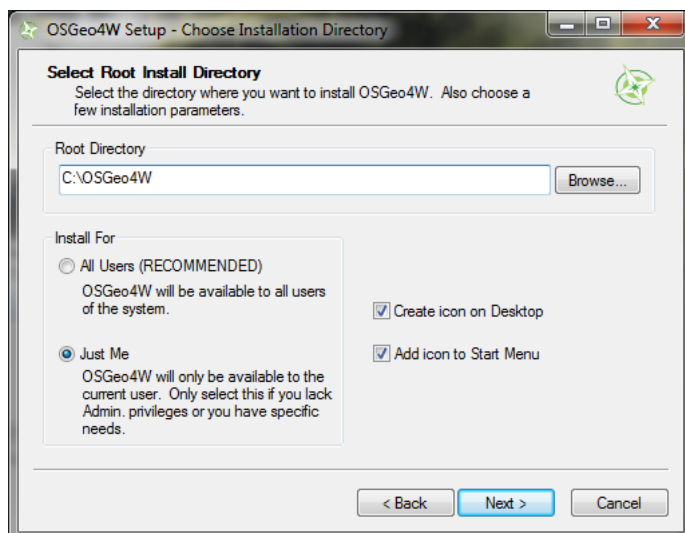
Lägg till Programmet OSGeo4W och verktyget OrfeoToolbox (OTB)

1. På <http://trac.osgeo.org/osgeo4w/> kan du titta på steg 1 under Quick Start for OSGeo4w Users. Klicka på *OSGeo4w Installer* och *Run* filen osgeo4w-setup.exe.
2. När "OSGeo4w Net Release Setup Program" fönstret poppar-upp, välj *Advanced Install* som i figur 1. Klicka sedan på Next.



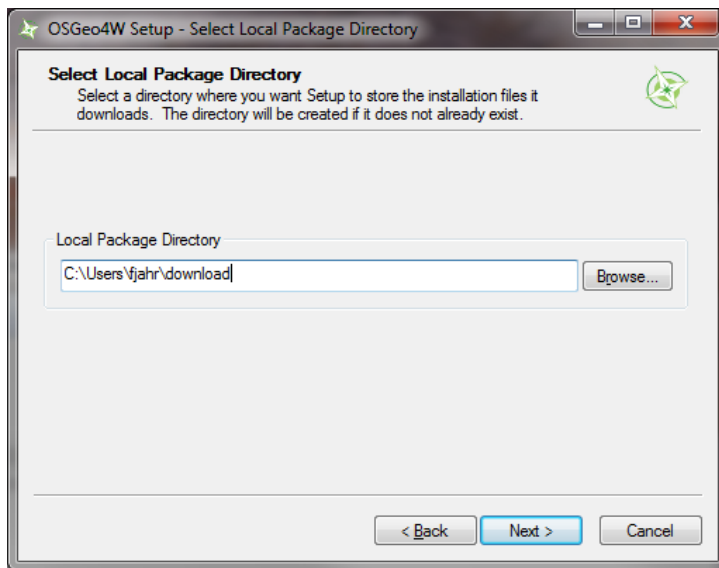
Figur 1. OSGeo4w Net Release Setup Program.

3. Välj *Install from Internet*.
4. Välj *C:\OSGeo4W* till din *Root Directory*, och *Install For Just Me* som i figur 2.



Figur 2. Select Root Install Directory.

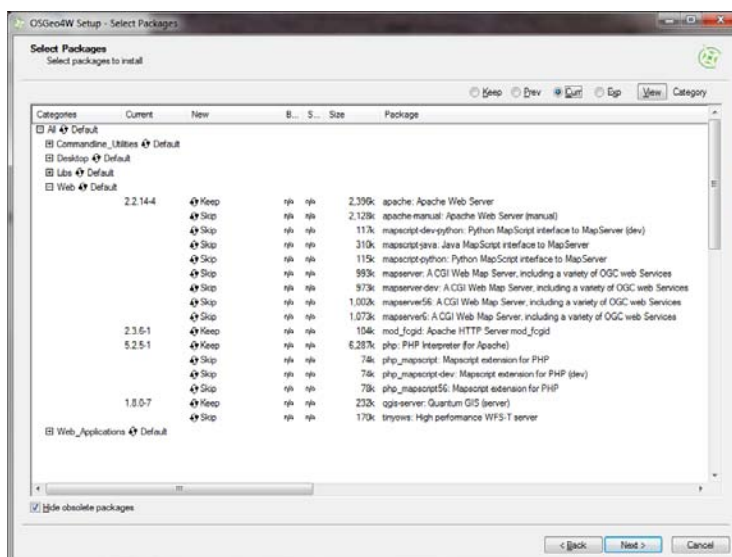
- I Select Local Package Directory kommer din hemmapp upp som en default (t.ex., C:\Users\ditt_username). Du kan lägga till en mapp (t.ex., "download"), så att det är C:\Users\ditt_username\download (figur 3). Klicka på next.



Figur 3. Select Local Package Directory.

- Använd *Direct Connection* som Internet Connection. Klicka på Next.

Nu börjar processen av att ladda ner olika Packages till OSGeo4W. Men först ska vi välj ut några särskilda extra packages vi vill ha, inklusive OrfeoToolbox. I fönstret "Select Packages" finns 5 olika rader med en plus tecken bredvid. Klicka på plus tecken som står bredvid "Web" (inte Web_Applications, men Web). Du kan sedan göra fönstret större så att du kan se namnen på Packages längst till höger (figur 4).



Figur 4. Select Packages.

- Dessa packages är en lista av alla som är tillgängliga. Vissa är redan en del av de packages som installeras via Default, och därför har ordet "Keep" i raden. Det finns också


packages som inte är installerad som default och har ordet "Skip" i raden, men kan installeras om du vill.

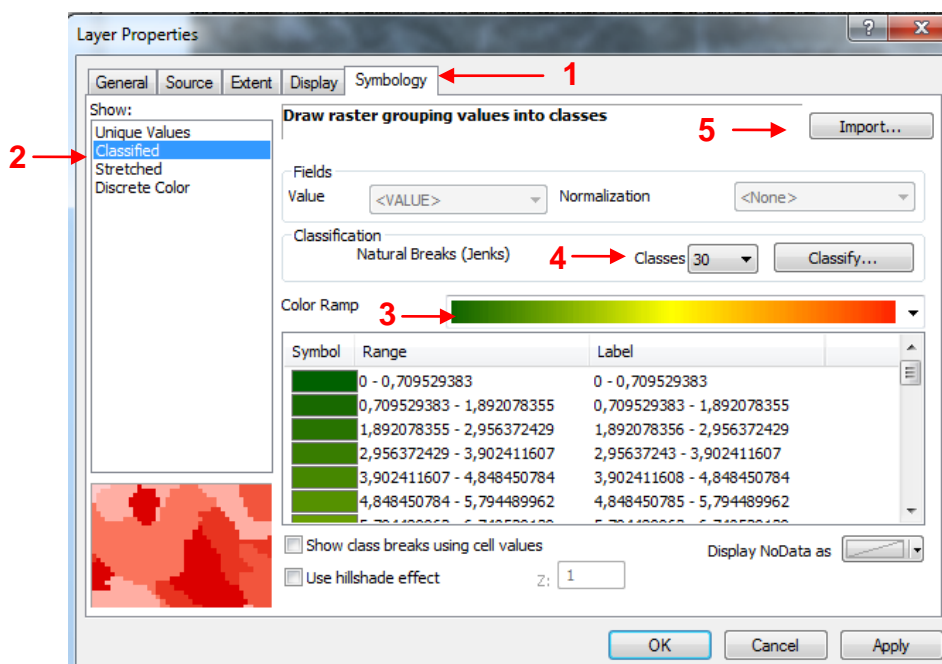
8. Vi vill installera en av dessa, som kallas "qgis-server: Quantum GIS (server). För att förbereda installation av denna, klicka på ordet *Skip* bredvid "qgis-server...". När du klicka på den, ändras ordet *Skip* till *Keep*. Om du av misstag klicka på något annat paket, kan du bara klicka igen för att välja mellan Skip/Keep.
9. Det kan vara en del packages som redan har ordet *Keep* i raden – lämna dessa som de är, under varje kategori.
10. Klicka på plus tecken bredvid Libs. Som extra packages, välj *Keep* för följande:
 - gdal: The GDAL/OGR library and commandline tools
 - gdal-python: The GDAL/OGR Python Bindings and Scripts
 - laszip: The LASzip compression library
 - libgeotiff: The Libgeotiff library...
 - libjpeg: A library for manipulating JPEG image format files...
 - liblas-python: libLAS Python bindings
 - libtiff: A library for manipulating TIFF format...
 - ossim: Ossim
 - otb-bin: OrfeoToolbox application modules
 - otb-python: Orfeo Toolbox python...
 - proj: The PROJ.4 library and commands for coordinate system...
 - qgis-common: Quantum GIS (common)
 - qgis-grass-plugin: GRASS plugin for Quantum GIS
 - shapelib: Library and utilities for working with Shapefiles
11. Klicka på plus tecken bredvid Desktop. Som extra packages, välj *Keep* för följande:
 - grass: GRASS GIS – stable release
 - otb-monteverdi: Monteverdi – Desktop application based on Orfeo...
 - qgis: Quantum GIS (desktop)
 - qgis-full: Quantum GIS full (desktop meta package...)
 - saga: SAGA System for Automated Geographical Analyses
12. Klicka på plus tecken bredvid Commandline_Uutilities. Som extra packages, välj *Keep* för följande:
 - gdal: The GDAL/OGR library and commandline tools
 - libgeotiff: the Libgeotiff library...
 - liblas: The libLAS commandline utilities
 - proj: The PROJ.4 library
 - python: Python meta package...
 - setup: OSGeo4W Installer/Updater
 - shapelib: Library and utilities for working with Shapefiles
 - shell: OSGeo4W Command Shell
13. Klicka på *Next*. OSGeo4W börjar installeras.

Titta på indata

14. Starta ArcMap. Spara ditt projekt genom att gå till File | Save As, navigera till C:\Uppgift_4, döp filen till **uppgift4.mxd** och klicka Save.

Obs! Om du inte har ArcMap kan du istället jobba med Quantum GIS (QGIS), som följer med OSGeo4W shell. Instruktioner och support för QGIS ges inte här, men det är ganska lätt att förstå och användarvänligt. För att starta QGIS kan du gå till Programs/OSGeo4W/Quantum GIS Desktop, eller bara skriva in qgis på OSGeo4Ws kommandofönstret. I QGIS borde du kunna se ett fönster till höger som kallas Sextante Toolbox. Här finns tillgång till verktyg från andra program, t.ex., Fusion! Det är fortfarande buggigt och fungerar inte smidigt just nu, men när det kommer att fungera i framtiden, är detta bra att känna till!

15. I ArcMap, klicka på *Add Data-symbolen*  i fönstrets övre del.
16. I Add Data-fönstret som kommer upp, vid *Look in*: Gå till C:\Uppgift_4 och välj orthophoto_12_epsg3006.tif-filen som ligger där (dubbelklicka inte). Klicka *Add*. På frågan om du vill bygga pyramider kan du svara ja. Nu ska du se att ortofotot har visat sig i ditt ArcMap-fönster.
17. Göra liknande för att lägga till filerna **09P001_67275_5925_25_height_segm.tif** och **09P001_67275_5950_25_height_segm.tif** som ligger under C:\Uppgift_4. Dessa bilder kommer från höjddata från laserskanningen, men har jämnats, eller "suddats", ut något för att ge en bättre segmentering. Höjderna är beräknade över marken.
18. I ArcMaps *Table of Contents* i vänstra delen av fönstret kan du se alla filer som lästs in. Dubbelklicka på en av filerna med höjddata (09P001_67275...) för att öppna *Layer Properties*.

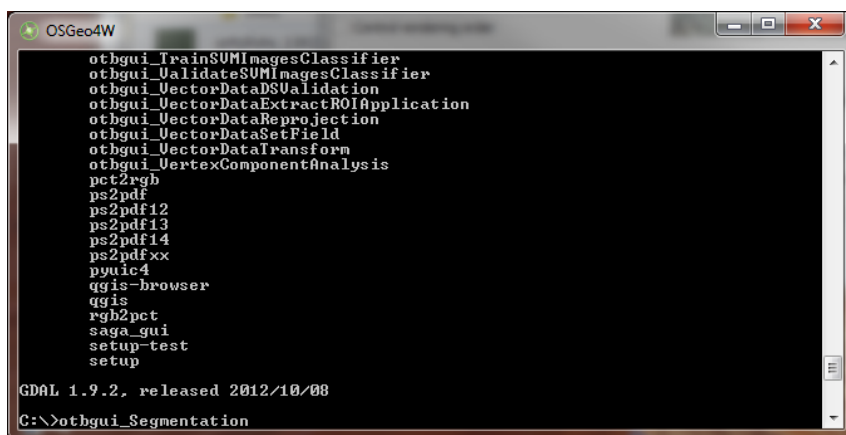


Figur 5. Layer Properties för en av bilderna med höjddata (09P001_67275_5925_25_height_segm.tif).

19. I *Layer Properties*, gå till fliken *Symbology* (1 i figur 5). I vänstra delen av fönstret under *Show*, klicka *Classified* (2) (Får du en fråga om att skapa histogram, svara ja). Detta görs för att dela in höjdvärdena i bilden i intervall, och ge varje höjdintervall en viss färg. Klicka på den svart-vita skalan vid *Color Ramp* (3) och välj en som går från grönt till rött istället.
20. Tryck på knappen vid *Classes* under *Classification* (4) och välj 30. Det betyder att avståndet från marken till maxhöjden kommer att delas in i 30 intervall. Klicka OK för att stänga fönstret. Den aktuella "höjdbilden" har nu en färgskala från grönt (låg höjd) till rött (hög höjd).
21. I ArcMaps *Table of Contents*, dubbelklicka nu på den andra "höjdbilden" för att öppna *Layer properties* för den.
22. Gå till fliken *Symbology* och klicka på *Classified*. För att höjdintervallen/klasserna i färgskalan ska bli lika för den här bilden som för den förra så klickar du nu på *Import...* (5) högst upp till höger i fönstret. Du bör då kunna se namnet på den förra filen du "färgade", 09P001_67275... Välj den och klicka OK. Klicka OK igen för att stänga *Layer Properties*

Gör segmenteringen

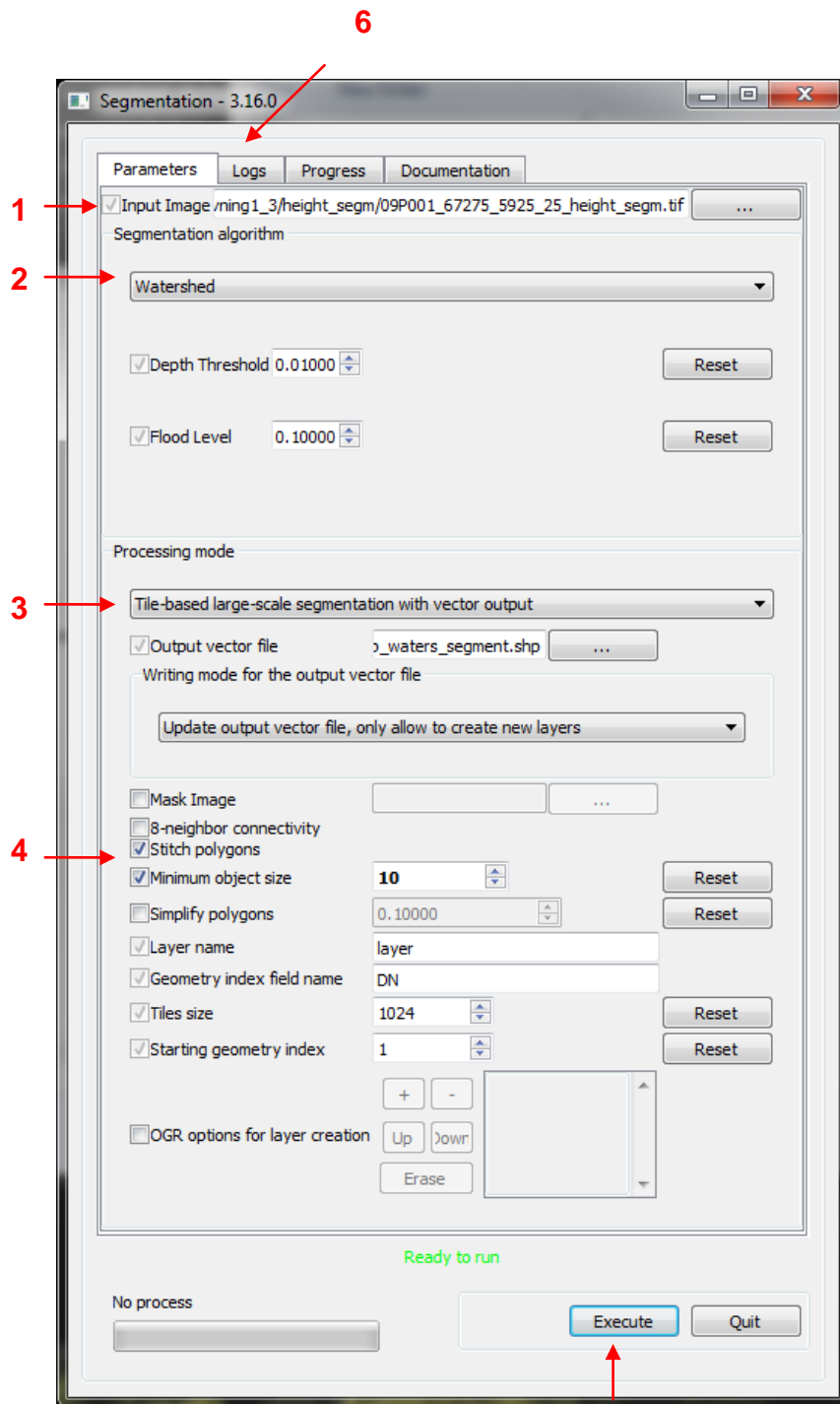
23. Öppna OSGeo4W shell genom att klicka på till Programs/OSGeo4W/OSGeo4W. Det kommer att dyka upp ett kommandofönster. Här kan du köra flera program via detta fönster. Om du någon gång vill stänga kommandofönstret är det bara att skriva "exit" vid C: prompten. (Egentligen borde man kunna köra Orfeo Toolbox från QGIS, men det finns fortfarande buggar som gör att det fungerar inte riktigt).
24. I kommandofönstret skriv `otbgui_Segmentation` (figur 6).



```
OSGeo4W
otbgui_TrainSUMImagesClassifier
otbgui_ValidateSUMImagesClassifier
otbgui_VectorDataDSValidation
otbgui_VectorDataExtractROIApplication
otbgui_VectorDataReprojection
otbgui_VectorDataSetField
otbgui_VectorDataTransform
otbgui_VertexComponentAnalysis
pct2rgb
ps2pdf
ps2pdf12
ps2pdf13
ps2pdf14
ps2pdfxx
pyuic4
qgis-browser
qgis
rgb2pct
saga_gui
setup-test
setup
GDAL 1.9.2, released 2012/10/08
C:\>otbgui_Segmentation
```

Figur 6. OSGeo4W kommandotolk.

25. Detta kommando startar Orfeo Toolbox (OTB) graphical user interface (GUI) för Segmentation (figur 7).



Figur 7. Orfeo Toolbox graphical user interface för Segmentering.

26. Välj en av de height_segm.tif filerna som en *Input Image*, 1 i figur 7.
27. Välj segmenteringsalgoritmen. Det finns flera här, men den som fungerar just nu med filer som har bara ett skikt (t.ex., height_segm filerna) är en *Watershed algoritmen*. Välj denna under 2 i figur 7.
28. Välj *Tile-based large-scale segmentation with vector output* och ge en output filnamn med en .shp extension. Till writing mode kan du lämna den med default "Update output vector file, only allow to create new layers", 3 i figur 7.

29. Bocka för *Minimum object size* och ändra den till 10. Du kan ändra här för att få olika nivåer av detalj i segmenteringen. Se till at *Stitch polygons* är också förbockade, vilket är default, 4 i figur 7.
30. *Execute* (5 i figur 7) och vänta på resultat. Visa meddelande kan dyka upp som säger någonting om "errors" eller "attempt to delete shape...". Du kan ignorera dessa. Det viktiga är att det skapas en Shape-fil med data i den. Om du har problem kan du titta på Logs (6 i figur 7).
31. Om du vill veta mer om vad du gör i Orfeo Toolbox kan du kolla på fliken "Documentation" och på deras websida.
32. Nu ska du göra segmenteringsprocessen med ortofotot. Använd då *Mean-Shift* algoritmen istället, eftersom det fungerar bättre med flerbandsfiler än Watershed. Detta kommer att ta längre tid eftersom ortofotot har tre band. När du köra "Mean-shift" kan det hända att du får ett felmeddelande som säger "Attempt to delete shape with feature id.... But it is marked deleted already". Detta felmeddelande påverkar inte resultatet och du kan bortse från det.
33. Nu kan du prova en Mean-Shift segmentering med en kombinerad fil av både trädhöjd samt ortofotot. Filen har 4 band (band 1 är trädhöjd från laser medan band 2,3,och 4 är färg information från ortofotot). Filen heter "Kombi_height_orto.tif"
34. Titta på resultaten av segmentering i ArcMap. Jämför resultat med varandra och emot din indata källa. För att göra detta behöver du göra så att bara en outline används i ArcInfo för att visa segmenterings shape-filerna (Sätt *Fill Color* till *No color* och *Outline Color* till något som syns mot bakgrunden. Du kan också öka *Outline Width* så att linjerna blir tjockare. Tryck OK.)

Notera att Orfeo Toolbox är gratis och ett ganska enkelt segmenteringsverktyg. Det är inte buggfritt och utvecklas kontinuerligt av frivilliga personer. Det finns andra kommersiella mjukvaror som gör segmentering på ett mer avancerad sätt, men just nu är det ständig utveckling av segmenterings-verktyg och under de närmsta åren kan vi förväntas se mer smidiga verktyg i både kommersiella och Open Source mjukvara.

Du behöver inte lämna in svar på några frågor i den här övningen, men fundera ändå över vad du ser: jämför segmenteringarna och titta på dem både mot höjddata (..._height_seg.tif) och ortofoto i bakgrunden. Hur ser det ut i skogskanter, skuggor, hyggen, olika delar av skogen...?

Inlämning

Shapefilerna med dina tre segmenteringsresultat från 1) höjdlaserdata, 2) ortofotot, och 3) den kombinerade höjd+orto filen. **OBS:** Varje shp-fil har tillhörande ytterligare tre filer (prj, dbf, shx) dessa måste samtliga skickas med för att shp-filen ska kunna öppnas i ArcMap.

Uppgift 5 - Introduktion till R

I skattningsövningarna längre fram i kursen kommer du att använda R, en utvecklingsmiljö/programmeringsspråk som ofta används för statistiska beräkningar. R är gratis och har öppen källkod. Det består av ett grundläggande ”paket” av funktioner för statistik och för att rita grafer, samt ett stort antal tilläggspaket skapade av användare runt om i världen.

Syftet med den här övningen är att du ska få

- en introduktion till hur R fungerar genom att använda några funktioner och färdiga skript
- ta en första titt på sambanden mellan lasermått och fältmätta skogliga variabler.

R kan, precis som Windows kommandoprompt, användas genom att man ger kommandon ett och ett i prompten, eller genom att man samlar en rad kommandon i en fil, ett så kallat skript. Den första metoden kan vara bra när man snabbt vill testa saker, men så fort uppgiften omfattar mer än ett par rader kod är det mer praktiskt att använda skript. Det går då enklare att göra ändringar i koden, att hålla reda på vad man har gjort och att återanvända kod.

Ett R-skript kan skrivas i en texteditor, till exempel Windows Notepad (anteckningar), och sparas med filändelsen .R. Det är dock praktiskt att använda en speciell texteditor som kan färga särskilda delar av koden enligt ett visst mönster, för att koden ska bli mer lättläst. I övningen kommer vi att använda RStudio, ett program med en sådan texteditor och andra funktioner speciellt anpassade för R. Även RStudio är gratis. Ladda ner dessa två program.

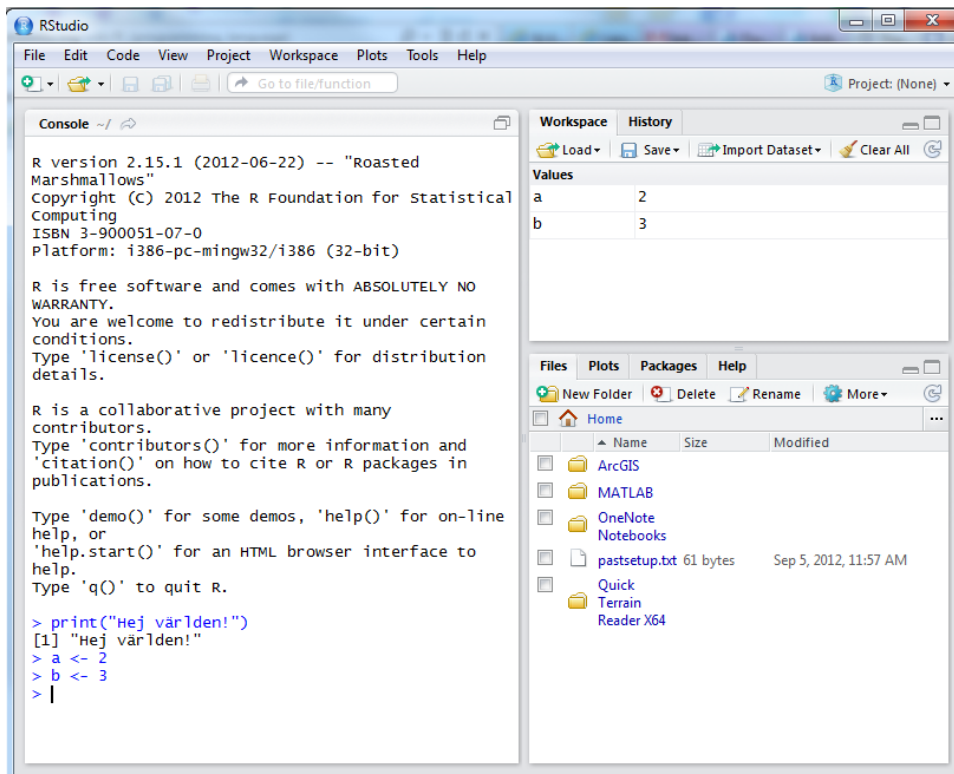
Innan du börjar: Kopiera filerna (plotData.txt, readPlotData.R, Scatterplot.R) under mappen Uppgift_5 (från servern geodata) till **C:\Uppgift_5** alternativt **C:\Fusion\Uppgift_5** (som du skapar).

Crash course

1. RStudio startas från skrivbordet eller från Windows startmeny om det finns genvägar där, annars från Startmenyn -> All Programs -> RStudio -> RStudio.

RStudio-fönstret kommer troligen att se ut ungefär som i figur 1, med fönstret indelat i tre delar. I konsolen, Console, kan du skriva in kommandon direkt och köra dem genom att trycka enter.

I den övre högra delen i figur 1 kan man välja mellan att visa Workspace (mer om det senare) och History (en lista över de kommandon du skrivit in sedan du började). I den nedre högra delen kan man bland annat se grafer man ritat upp och hjälpsidor. Storleken på de olika delarna av fönstret kan ändras genom att man drar i deras kanter.



Figur 1. RStudio.

2. I konsolen (figur 1), skriv `print("Hej världen!")` och tryck enter. Texten `[1] "Hej världen!"` följd av en ny rad med tecknet `>` bör då dyka upp på skärmen.
3. Det du har gjort är att tala om för R att på skärmen skriva ut (print) en text (inom `" "`-tecken), och att den texten ska lyda Hej världen!
4. Skriv nu in `a <- 2` och tryck enter, sedan `b <- 3` och tryck enter (figur 1). Det du nu har gjort är att skapa två objekt, `a` och `b`, som har värdena 2 respektive 3. Tänk dig att teckenkombinationen `<-` är en pil som betyder att värdet 2 ges till objektet `a`.
 Ingenting skrivs ut på skärmen den här gången, men om du tittar i övre högra delen (Workspace) så ser du att både `a` och `b` finns där med sina respektive värden. Det betyder att objekten `a` och `b` nu finns i minnet, och när du skriver in `a` i konsolen så kommer det att tolkas som siffran 2, och när du skriver in `b` tolkas det som 3.
5. Prova att skriva in till exempel `a+b`, `a-b`, `a/b` och `a*b` i konsolen. Avsluta varje kommando med enter. Resultatet bör inte komma som en överraskning!
6. När R används såhär är det som en vanlig miniräknare. Om du skulle vilja använda något av resultaten i en senare beräkning måste du spara det som ett objekt. Skriv in `c <- a+b` och tryck enter. Objektet `c` dyker då upp i ditt Workspace. Nu börjar det kanske bli tydligare vad som är poängen: du kan göra långa kedjor av beräkningar och annat, och använda resultatet i ett steg som indata till nästa.

Skript för att läsa in data

I den här delen av övningen kommer du att använda ett färdigt skript (program) för att läsa in data från laserskanning och fältinventering av provytor. Du kommer sedan att få göra spridningsdiagram med lasermått och fältmätta variabler, samt titta på hur starkt sambandet är mellan dessa. Därefter får du själv ändra i skriptet för att undersöka fler variabler. Laser- och fältdata från provytorna finns i filen **plotData.txt** i övningsmappen.

7. *a*, *b*, *c* och eventuellt andra objekt du skapat kommer inte att vara till någon nytta nu, så ta bort dem genom att klicka på *Clear All* i Workspace (övre högra hörnet av Workspace-delen). Svara Ja på frågan om du är säker på att du vill ta bort alla objekt.
8. Gå till File | Open File... högst upp till vänster. Navigera till **C:\Uppgift_5** och öppna filen `readPlotData.R`. Den kommer att öppnas i en ny del av fönstret, troligen uppe till vänster. **Läs igenom skriptet** och försök att få en uppfattning om vad det gör. Verkar det förvirrande nu så klarnar det förhoppningsvis under övningens gång!

Allt som står efter # är en kommentar. Den exekveras inte, det vill säga R struntar i att göra något med den. Kommentarer används för att beskriva och förklara koden för den person som läser den. Som du ser innehåller det här skriptet många kommentarer, men bara fyra rader kod.

9. Kolla sökvägen till filen med provytedata, och ändra om det behövs (rad 22 i koden).
10. Kör programmet genom att i övre menyraden gå till Code | Run Region | Run All, eller genom att trycka Ctrl+Alt+R.

I Workspace ska du nu kunna se `dfPlotData` och `dataFil`. Intill `dfPlotData` står det ”53 obs. of 37 variables”. En data frame består av ett antal kolumner, i det här fallet 37, där alla kolumner har samma antal rader och där alla rader i en kolumn har samma datatyp (heltal, flyttal, textsträng...). Man kan tänka sig en data frame som en tabell. Här har den 53 rader, en för varje provyta, och 37 kolumner, en för varje variabel (lasermått eller fältmätt variabel).

11. Klicka en gång på `dfPlotData` i Workspace. I den del av fönstret där skriptet finns kommer det nu upp en ny flik, `dfPlotData`. Där kan du se kolumnnamn och innehåll i data ramen (figur 2). Titta en stund för att bekanta dig med innehållet. Vad tror du att kolumnnamnen står för? Vilka enheter används?

	PlotID	RT90E	RT90N	VolumeTotal	VolumePine	VolumeSpruce	VolumeBirch	VolumeAspen
1	1001	1690996	7136999	118.59	111.93	1.71	4.94	0.00
2	1102	1691198	7137096	58.91	0.00	1.67	57.25	0.00
3	2006	1693197	7136595	174.82	5.54	141.20	24.32	0.00
4	2010	1691998	7135999	142.54	110.46	1.41	30.67	0.00
5	2102	1691694	7135898	132.21	85.73	11.43	35.05	0.00
6	2106	1694298	7137206	150.95	38.09	33.74	79.12	0.00
7	3015	1689107	7133993	95.95	30.76	59.03	6.17	0.00
8	3019	1688303	7133200	164.16	104.75	24.90	34.51	0.00

Figur 2. Här kan du titta på innehållet i data ramen dfPlotData.

Lasermåtten i filen plotData.txt, som du nu läst in till en data frame vid namn dfPlotData, har beräknats med hjälp av programmet Fusion. Man har sedan lagt ihop dessa data med de fältmätta variablerna för provytorna. För översiktens skull är den här filen rensad och innehåller bara några av alla de mått som Fusion beräknar. Tabell 1 sammanfattar de olika variablerna och lasermåtten i filen.

Tabell 1. Beskrivning av de fältmätta variabler och lasermått som finns i filen plotData.txt (och alltså även i data ramen dfPlotData).

Namn	Förklaring	Typ av mått
PlotID	Provytans ID	Fält
RT90E	Östlig koordinat, RT90	Fält
RT90N	Nordlig koordinat, RT90	Fält
VolumeTotal	Volym alla trädslag	Fält
VolumePine	Volym tall	Fält
VolumeSpruce	Volym gran	Fält
VolumeBirch	Volym björk	Fält
VolumeAspen	Volym asp	Fält
VolumeContorta	Volym contorta	Fält
BasalArea	Grundyta	Fält
HGV	Grundytvägd medelhöjd	Fält
DGV	Grundytvägd medeldiameter	Fält
Total_return_count	Totalt antal retur	Laser
Total_return_count_above_1_50	Antal retur över 1,50 m	Laser
Elev_minimum	Lägst retur	Laser
Elev_maximum	Högst retur	Laser
Elev_mean	Medelhöjd för returerna	Laser
Elev_CV	Variationskoeff. för returerna	Laser
Elev_P01	Höjdpercentil 1	Laser
Elev_P05	Höjdpercentil 5	Laser

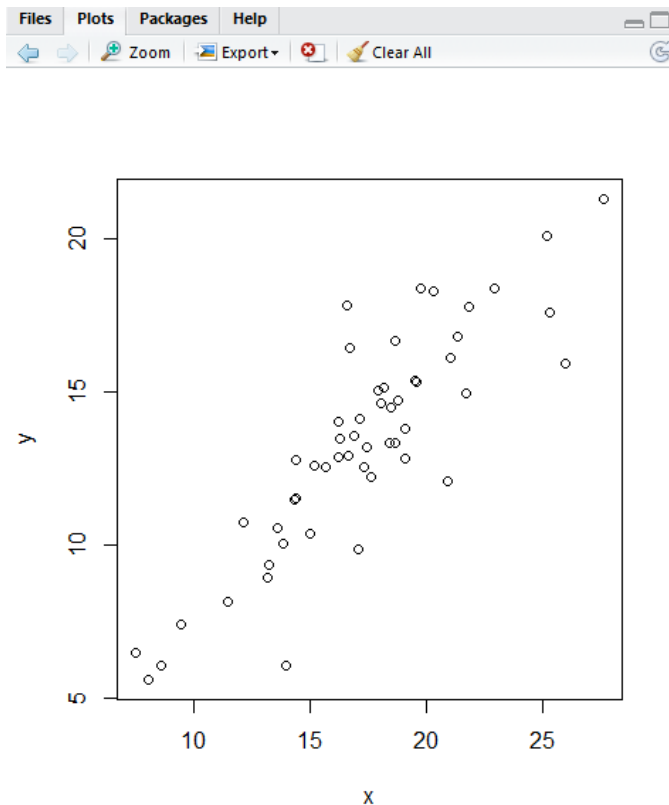
Elev_P10	Höjdpercentil 10	Laser
Elev_P20	Höjdpercentil 20	Laser
Elev_P25	Höjdpercentil 25	Laser
Elev_P30	Höjdpercentil 30	Laser
Elev_P40	Höjdpercentil 40	Laser
Elev_P50	Höjdpercentil 50	Laser
Elev_P60	Höjdpercentil 60	Laser
Elev_P70	Höjdpercentil 70	Laser
Elev_P75	Höjdpercentil 75	Laser
Elev_P80	Höjdpercentil 80	Laser
Elev_P90	Höjdpercentil 90	Laser
Elev_P95	Höjdpercentil 95	Laser
Elev_P99	Höjdpercentil 99	Laser
Percentage_first_returns_above_1_50	Andel förstareturer över 1,50 m, %	Laser
Percentage_all_returns_above_1_50	Andel alla returer över 1,50 m, %	Laser
First_returns_above_1_50	Antal förstareturer över 1,50 m	Laser
Total_first_returns	Totalt antal förstareturer	Laser

Skript för att göra spridningsdiagram

Nu ska du fortsätta med att göra spridningsdiagram (*scatter plot*) och beräkna korrelationen mellan olika mått och variabler. Vad korrelation är beskrivs längre ner. Spridningsdiagrammet ska ha ett lasermått på x-axeln och en fältmått variabel på y-axeln. Varje provyta ritas sedan in som en prick i diagrammet. Då kan man se om det verkar finnas något samband mellan lasermåttet och den fältmätta variabeln.

12. I RStudios huvudfönster, gå till File | Open File. Navigera till **C:\Uppgift_5** och öppna filen **Scatterplot.R**.
13. Läs koden och försök att förstå vad den gör. Om det känns oklart – ta det lugnt! Det lossnar nog när du får börja experimentera själv.
14. Kör koden genom att gå till Code | Run Region | Run All eller genom att trycka Ctrl+Alt+R.

Spridningsdiagrammet ska nu ritas upp i en av rutorna i RStudios huvudfönster, figur 3.



Figur 3. Spridningsdiagrammet.

Fråga 1. Vilket lasermått har plottats på x-axeln och vilken variabel på y-axeln?

Fråga 2. Punkterna följer någotsånär en linje, där större värde på x hänger ihop med större värde på y. Varför?

Nu till korrelationen: Längst ner i konsolen ska du kunna se texten

```
> cor(x, y)
[1] 0.8725024
```

Det betyder att korrelationen mellan x- och y-variabeln är 0,87. Korrelation är ett mått på styrkan i ett linjärt samband. Om punkterna hade legat exakt på en rät linje hade korrelationen varit 1 vilket innebär ett perfekt, positivt linjärt samband. Om punkterna hade legat på en linje, men linjen hade lutat åt andra hållet så att y minskar när x ökar, så hade korrelationen varit -1, ett perfekt, negativt linjärt samband. Om spridningsdiagrammet hade sett ut som en hagelsvärm utan urskiljbar trend hade korrelationen varit nära noll.

I det här fallet är korrelationen 0,87 vilket innebär ett starkt positivt samband.

15. Prova nu att plotta några andra lasermått mot grundtevägd medelhöjd, och beräkna

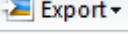
korrelationen. Detta gör du genom att på raden `x <- dfPlotData$Elev_maximum` i koden byta ut `Elev_maximum` mot kolumnnamnet på det mått du vill använda. Ta några mått som du tror har ett starkt samband med höjden, och gärna också något som du tror har ett svagt samband.

16. Gör ett spridningsdiagram av det mått som har starkast linjärt samband med höjden. Sätt även titel på diagram och axlar samt uppge korrelationsmått.

I skriptet byt ut `plot(x, y)` mot texten nedan.


```
plot(x, y, main=paste("main title, korrelation
= ",round(cor(x,y), digits=3)), xlab="x-axis label",
ylab="y-axis label")
```

Koden kommer även att avrunda korrelationsvärdet till tre siffror. Skriv in den rätta etiketten för din "x-axis label" och "y-axis label".

Spara diagrammet genom att klicka på symbolen  och gå vidare till *Save Plot As Image*. Spara bilden i C:\Upppgift_5.

Fråga 3. För vilket lasermått hittar du det starkaste linjära sambandet med grundytvägd medelhöjd?

Fråga 4. Vad är korrelationen mellan grundytvägd medelhöjd och det måttet?

Fråga 5. Varför tror du att just det måttet är "bäst" på att beskriva medelhöjden? Tips: kapitel 4 i kompendiet.

Extrauppgift

På y-axeln kan du prova andra variabler än HGV, som skulle vara intressanta att skatta. Testa vilka lasermått de korrelerar bäst med, och om det verkar finnas samband som inte är linjära (till exempel att punkterna i spridningsdiagrammet verkar följa en böjd kurva istället för en linje).

Du kan läsa mera om R på <http://www.statmethods.net/index.html>. Det finns massor på nätet från användare av R, inklusive Forum (<http://community.r-user.org/>), Blog (<http://www.r-bloggers.com/>), samt exempel av kod (<http://www.r-project.org/>)

Inlämning

Lämna in svar på frågorna 1-5 samt bild av spridningsdiagram.

Snabbkommandon RStudio

Ctrl+Alt+R - Kör samtlig kod/hela skriptet

Ctrl+Enter - Kör en rad - förutsätter att markören står i början av aktuell rad

Tabb - För automatisk ifyllnad av olika variabler

Uppgift 6 - markmodellens nytta vid drivningsplanering

Det ursprungliga syftet med Lantmäteriets NH-skanning är att skapa en ny nationell höjdmodell, det vill säga en modell med markhöjder över hela landet. En sådan markmodell kan även vara till nytta inom skogsbruket vid till exempel planering av avverkningar, eftersom man kan få ut information om var det finns branter, myrkanter och annat som kan vara svårt att se i till exempel ett ortofoto. Man kan också hitta kolbottnar (rester efter kolmilor, räknas som fornminnen och ska därför inte markberedas), diken, stengärdesgårdar och annat.

Den här uppgiften är **mer öppet formulerad än tidigare för att uppmuntra till diskussion** och funderingar, och det är inte meningen att lösningen ska kännas solklar på en gång. Ta chansen att tillsammans resonera er fram till bra lösningar!

Scenario

Föreställ dig att du arbetar med drivningsplanering. Ditt mål är en effektiv planering, där hänsyn dessutom ska tas till mark och vatten (undvik körskador) och förarnas komfort och säkerhet.

Ur tillgängliga data ska du ta fram information som kan användas för att uppnå målet om effektiv och skonsam planering. **Den ”praktiska” delen av den här uppgiften stannar vid detta**, men tänk dig att informationen sedan ska kunna användas i ett GIS för att göra en karta med förslag på lämpliga körvägar.

Rent praktiskt bör du främst fokusera på att ta fram information om markens höjd, lutning, lutningens riktning och terrängskuggning, men diskutera gärna i vidare svängar exempelvis naturvärden, kulturlämningar, grundförhållanden etc.

Läs hela instruktionen, inklusive frågorna under **Inlämning**, samt den rekommenderade läsningen innan du sätter igång.

Data och mjukvara

Du har tillgång till laserdata från Lantmäteriets NH-skanning. Använd data från det område du valde ut i uppgift 1. Om du inte valde ut ett eget område i uppgift 1 kan du kopiera las-filen från mappen uppgift_2 på servern geodata.

Du kan även ladda ner terrängkarta och ortofoto från GET. Titta igen på uppgift 1 om du behöver information och hjälp.

Vad gäller mjukvara har du tillgång till FUSION och ArcMap.

Läsning

- http://www.skogforsk.se/PageFiles/65200/Resultat_nr6_2011_low.pdf
- Avsnitt 5.1 i kurskompendiet
- Introduktion till Fusion, stencilen finns på geodata.

FUSION-tips

Två verktyg som du kommer att ha nytta av i denna uppgift är GridSurfaceCreate och DTM2ASCII. Dessa beskrivs utförligt i manualen.

Steg ett är att med hjälp av verktyg i Fusion skapa en digital höjdmodell (.dtm) utifrån laserpunkterna. Detta gör du med hjälp av verktyget GridSurfaceCreate, läs mer utförligt om verktyget i Fusion-manualen. I manualen hittar du även verktygets syntax, dvs koden som du skriver in i notepad och sedan gör en batchfil av och som i sin tur kör verktyget i Fusion. Har du glömt hur man gick till väga titta i uppgift 3.

Switchen "class" gör det möjligt att välja vilken klass den digitala höjdmodellen skapas ifrån. De LAS-filer som används i uppgiften har genomgått en markfiltrering (kompendiet avsnitt 3.2) för att identifiera troliga markpunkter. Om du inte använder switchen "class" kommer alla punkter att användas för att skapa markmodellen. Titta i Fusion-manualen och i kompendiets kapitel 9, särskilt tabell 6, för att få reda på hur du ska göra för att använda enbart markklassade punkter. Du bör också fundera över vilken cellstorlek du ska välja och i inlämningen motivera ditt val. Det finns flera aspekter att ta hänsyn till här.

Nästa steg är att använda verktyget DTM2ASCII i Fusion. Verktyget gör om ovanstående digitala höjdmodell till en ASCII-fil (.asc) som i sin tur kan läsas in i ArcMap. Syntax och utförlig beskrivning av verktyget finns i Fusion-manualen.

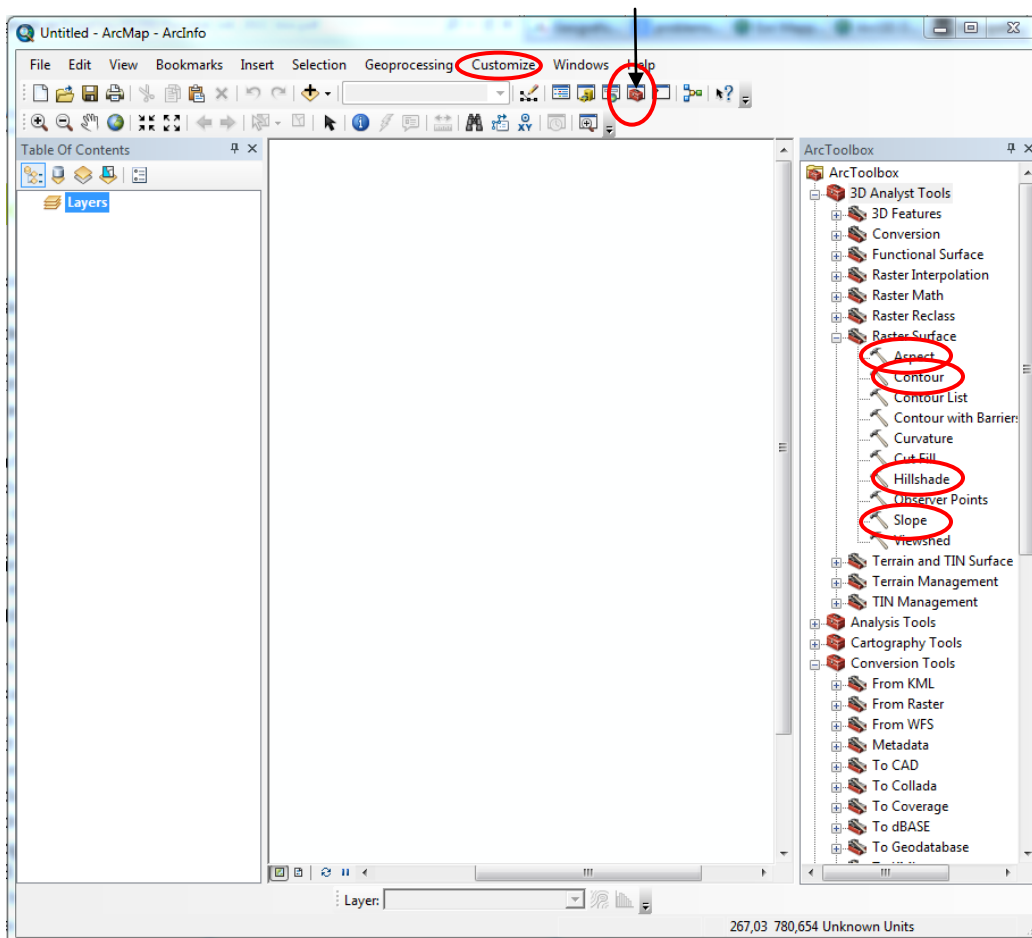
ArcMap-tips

I ArcMap kan du ha nytta av följande verktyg (figur 1):

- Hillshade (ArcToolbox -> 3D Analyst Tools -> Raster Surface)
- Slope (ArcToolbox -> 3D Analyst Tools -> Raster Surface)
- Aspect (ArcToolbox -> 3D Analyst Tools -> Raster Surface).
- Contour (ArcToolsbox -> 3D Analys Tools -> Contour.

För att du ska kunna använda verktyg i 3D Analyst Tools måste *3D Analyst* vara förbockad. Det görs under *Customize -> Extentions*.

ArcToolbox öppnar du genom att klicka på Toolbox-symbolen  (figur 1).



Figur 1. ArcMap med viktiga knappar och funktioner markerade.

Buggar: ArcMap kan bete sig lite oberäkneligt. Ibland fungerar ett verktyg och ibland inte, utan synbar orsak. Det kan då hjälpa att starta om ArcMap och/eller starta ett nytt projekt.

Inlämning

Uppgiften lämnas in i form av ett Word-dokument eller en pdf-fil med bilder eller screenshots från rater i ArcMap.

- Skriv några raders/en halv sidas introduktion om hur du har tänkt när du har löst uppgiften
- Redovisa vilka produkter du har tagit fram ur laserdata/laddat ner från GET och vilket syfte de fyller för uppgiften. Kommer du på andra datakällor du skulle vilja använda är det fritt fram att spåna kring dessa.
 - För sådant du gjort i FUSION: redovisa dina batch-kommandon. Motivera dina val av switchar och parametrar, samt de värden du använt på dessa.
 - För sådant du gjort i ArcMap: ta med en bild på produkten, beskriv vilket verktyg du använt, vilka inställningar du använde och varför. Du kan ta en screenshot eller exportera bilden till exempelvis jpeg (File -> Export map).

- För sådant som laddats ner från GET, och eventuella egna idéer om användbara data: beskriv varför du tagit med den produkten.
- Apropå markmodellen i FUSION: resonera kring vilken cellstorlek som kan vara lämplig att använda och varför. Det finns flera aspekter att ta hänsyn till!
- Skriv ca en halv sida om hur den information du tagit fram skulle kunna användas i en GIS-analys för att ta fram förslag på lämpliga vägar, föreslå gärna några verktyg för att automatgenerera körvägar och diskutera runt dessa.

Lycka till!

Om du vill fördjupa dig inom detta område kan Mattias Wählberg von Knorrings examensarbete rekommenderas; [Beslutsstöd för placering av basvägar i terrängen och beräkning av medelterrängtransportavstånd](http://stud.epsilon.slu.se/5013/1/Wahlberg_Von_Knorrning_M_121029.pdf),
http://stud.epsilon.slu.se/5013/1/Wahlberg_Von_Knorrning_M_121029.pdf.

Uppgift 7 - laserdata i rasterform

Vid gles laserskanning (under ca 5 punkter/m²) får man inte detaljerad information om de enskilda träden, men sett över en lite större yta (med en grupp av träd) beskriver punkterna skogens höjd och täthet. Ett sätt att använda informationen är att dela in det skannade området i ett raster, eller rutnät, och för varje rastercell beräkna olika höjd- och täthetsmått. Rastren ger inte absoluta värden på skogens höjd etc, men man kan ändå få en bra uppfattning om hur det ser ut bara genom att titta på rastren.

Den här uppgiften är **öppet formulerad för att uppmuntra till diskussion** och funderingar, och det är inte meningen att lösningen ska kännas solklar på en gång. Ta chansen att tillsammans resonera er fram till bra lösningar!

Scenario

Föreställ dig att du vill hitta gallringsskog och avverkningsmogen skog inom ett område. För att kunna planera besöken i fält så effektivt som möjligt vill du redan på kontoret skaffa dig en översiktlig bild av situationen.

Läs hela instruktionen och samt ta en titt på det du behöver av den rekommenderade läsningen innan du sätter igång.

Data och mjukvara

Du har tillgång till laserdata från Lantmäteriets NH-skanning. Använd data från det område du valde ut i uppgift 1. Om du inte gjorde uppgift 1 kan du kopiera las-filen från mappen uppgift_2 (geodata). Du kommer också att använda markmodellen (filändelse .dtm) som du gjorde i uppgift 6.

Du kan även ladda ner terrängkarta och ortofoto från GET, se <https://docs.google.com/file/d/0B3yuX8WpOEGESm5RNnBKV0ZfZfK/edit?pli=1> för information och hjälp.

Vad gäller mjukvara har du tillgång till FUSION och ArcMap.

(Om du vill undersöka närmare vilka lasermått och skogliga variabler som hänger ihop är det givetvis OK att använda data från uppgift 5 samt R och RStudio. Data från den övningen kommer visserligen inte från det område du arbetar med, men kan ändå säga en hel del.)

Lästips

- Kapitel 4 och 5 i kurskompendiet
- <http://www.skogsstyrelsen.se/Global/PUBLIKATIONER/Skogsskotselserien/PDF/07-Gallring.pdf>, sid 5-10
- <http://www.skogforsk.se/sv/KunskapDirekt/Gallra/Gallringsprogram-och-stamval/Nar-ar-det-dags-att-gallra/>
- http://stud.epsilon.slu.se/2257/1/Sjodin_M_110114.pdf
- http://ex-epsilon.slu.se:8080/archive/00002955/01/Arbetsrapport_235.pdf
- Introduktion till Fusion, stencilen finns på geodata.

FUSION-tips

I den här uppgiften kommer du att ha nytta av verktygen GridMetrics och CSV2Grid. Dessa beskrivs utförligt i manualen och även i Lathunden till FUSION (rekommenderas starkt att du använder denna). Kom ihåg att strukturen av en Fusion batch-fil diskuterades i uppgift 3. Denna övning är mer självständig och du kommer att skriva egna kommandon till GridMetrics och CSV2Grid och de blir lite mer komplexa. Det är bra att veta att du kan ha flera "switches" i din kommando-linje. Det kan se ut så här (kom ihåg att det ska vara i en enda linje av kod i textfilen, och att detta är bara ett exempel):

```
C:\FUSION\GridMetrics /outlier: /minht: C:\groundfile.dtm  
heightbreak cellsize C:\outputfile.csv C:\datafil.las
```

Tips: switches som du ska ta med till GridMetrics inkluderar "outlier", "minht", och en switch för att ta med ett minimum antal retur, samt att inte räkna på intensitet. Du kan självklart använda fler switches om de krävs för målet av din uppgift.

Du kan antingen skriva en batch-fil per verktyg (ett före GridMetrics och ett för CSV2Grid), eller skriva båda verktygskommandona i samma batch-fil och köra allt i ett svep. Men om du kör allt i en fil, var då extra observant på att använda rätt kolumner i CSV2Grid (om du kör ett verktyg i taget kan du öppna CSV-filen från GridMetrics i Excel och dubbelkolla kolumnnumren).

I en senare uppgift kommer du att få ta ut laserdata för provytor och i rasterform, göra egna regressionsfunktioner och tillämpa dessa på rastren för skattning av skogliga variabler. Du kommer då att använda data från ett område i Västerbotten där vi har fältinventerade provytor, men du kommer delvis att kunna återanvända batch-filer från den här uppgiften.

Excel-tips

Resultatet av GridMetrics är en CSV-fil. CSV står för *comma separated value* vilket betyder att elementen i filen skiljs åt av kommatecken. Filen är ett slags tabell, där varje rad motsvarar en rastercell och varje kolumn ett lasermått.

Om du vill titta på filen kan det göras i Excel. När du öppnar den kommer allt att visas i samma kolumn till att börja med. För att visa måtten i separata kolumner börjar du med att markera kolumnen A (där allting ligger) genom att klicka på bokstaven A. Gå sedan till fliken *Data* och klicka på *Text to Columns*. Där väljer du *Delimited*, klickar *Next* och under *Delimiters* bockar du för *Comma*. Klicka *Next* och sedan *Finish*. Det kan ta en stund eftersom filen är så stor.

Observera att du inte ska spara ovanstående kolumnindelning när du stänger filen. Kolumnindelningen i Excel gör att separeringen med kommatecken försvinner, och CSV2Grid kan då inte hantera filen och plocka ut rätt kolumn. Det går bra att öppna och utföra kolumnindelning av filen men stäng då filen *utan* att spara ändringar.

ArcMap-tips

I ArcMap kan du ha nytta av följande verktyg:

- Composite Bands (Arc Toolbox -> Data Management Tools -> Raster -> Raster Processing).

Composite Bands är ett bra verktyg för att kombinera ihop olika raster men titta även på lagren enskilt för att hitta lämpliga områden. Prova även att variera antalet färgkanaler och därmed kombinationer av färger, välj att till exempel lägga ett lasermått i röd kanal, ett annat i grön kanal och inget alls i blå kanal (ta bort krysset i *Visible*-rutan för blå kanal genom att högerklicka på blå ruta). Om kontrasten i bilderna inte är bra (svaga färg) kan man i ArcMap sträcka på bilderna i displayen.

Tips: Istället för Composite Bands kan transparensen på överliggande lager ställas in för ett liknande resultat men i en annan färgskala.

Se upp med att använda strikta gränser och sätt eventuella klasser med marginal! Datat är tänkt för att ge en överblick och indikation av fastigheten, aktuella områden bör kontrolleras med fältdata.

Inlämning

Uppgiften lämnas in i form av ett Word-dokument eller en pdf-fil. Raster/bilder från ArcMap kan för enkelhets skull klistras in som bilder eller screenshots i dokumentet.

- Skriv några raders eller upp till en halv sidas introduktion om hur du har tänkt när du löst uppgiften
- Redovisa vilka produkter du har tagit fram ur laserdata/laddat ner från GET och vilket syfte de fyller för uppgiften.
 - För sådant du gjort i FUSION: redovisa dina batch-kommandon. Motivera dina val av switchar och parametrar, samt de värden du använt på dessa.
 - För sådant du gjort i ArcMap: ta med en bild på produkten (screenshot eller t ex jpg går bra, klistra in i ditt dokument).
 - För sådant som laddats ner från GET: beskriv varför du tagit med den produkten.
- Visa i din resultat vilka områden du skulle välja i första hand om du gav dig ut i fält för att leta efter gallringsskog och avverkningsmogen skog, och förklara varför. Det går bra att göra en skärmdump från ArcMap eller att exportera bilden till exempelvis jpeg (File -> Export Map) och klistra in i din Word-fil. Du kan markera områden i bilden med något ritprogram eller genom att sätta in pilar etc direkt i Word (Insert -> Symbol).

Lycka till!

Uppgift 8 – förberedande regressionsövning

Uppgift 8 går ut på att förbereda data till kommande skattningsövningar och att börja undersöka sambanden mellan lasermått och skogliga variabler. Den skattningsmetod ni kommer att använda är i stort sett den areabaserade metod som utvecklats till operationell tillämpning i skogsbruket av professor Erik Næsset i Norge och som har fått stor operationell spridning. Ni kommer i en senare uppgift att använda data från samma område i en skattning med hjälp av Random Forests.

Uppgift 8 består av följande steg:

- ta fram en markmodell över aktuellt område
- klipp ut laserdata från provytor och höjdsätt punkterna över marken
- beräkna ett antal lasermått för provytorna
- lägg lasermått och fältdata från provytorna i en textfil
- läs in textfilen i R och undersök sambanden mellan lasermått och fältvariabler, med hjälp av statistiska mått och grafer

Ni kommer att få arbeta mer på egen hand med att ta fram regressionsfunktioner som beskriver en fältvariabel som en funktion av ett eller flera lasermått.

När ni har fått fram lämpliga regressionsfunktioner ska ni tillämpa dessa på lasermått i form av raster som täcker hela det aktuella området, detta är en del av uppgift 11. Ni kommer då inte att behöva beräkna lasermåtten i rasterform själva, utan får förberedda data.

Eftersom det är mycket viktigt att lasermåtten tas fram på samma sätt för provytorna och rastret får ni i den här uppgiften detaljerade instruktioner för hur ni i FUSION ska ta fram mått på provytenivå. En del av uppgiften är att läsa i FUSION-manualen och förklara vad de inställningar (switchar och parametrar) som används innebär.

Lästips

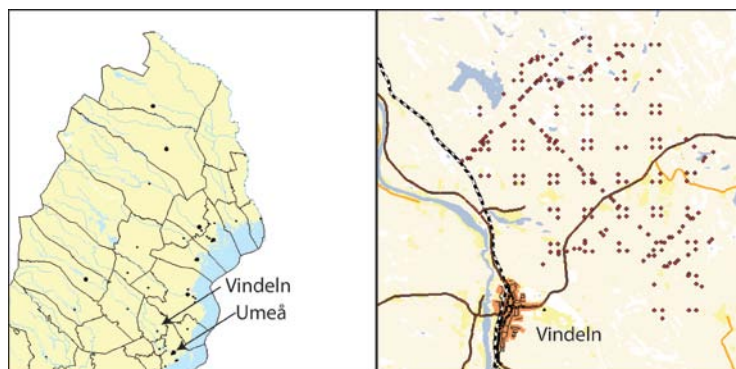
- Stencilen Regression i praktiken (geodata).
- Kapitel 6 i kurskompendiet

Studieområde

Data till den här uppgiften kommer från Krycklans avrinningsområde utanför Vindelns i Västerbotten, ca 80 km nordväst om Umeå. Figur 1 visar var området ligger. De röda punkterna representerar provytorna.

Data

Krycklan laserskannades i oktober 2010 inom ramen för den rikstäckande NH-skanningen. För att täcka in hela det område där vi har provtytor behövs 17 st 2.5 * 2.5-km LAS-filer. Det innebär stora mängder data, som skulle vara opraktiskt för er att ladda ner från SLU-servern. Ni får därför en LAS-fil som innehåller punkter i kvadrater om 50*50 m runt varje provyta. Filen fungerar som vilken LAS-fil som helst, bara det att områdena mellan dessa 50*50-metersbitar saknar punkter. Det gör ingen skillnad för tillvägagångssättet jämfört med om ni hade haft en fil med heltäckande punktdata.



Figur 1. Krycklans avrinningsområde ligger strax nordost om Vindeln i Västerbotten. De röda punkterna i högra delen av figuren visar provtytorna.

Fältdata har samlats in för provtytor med 10 meters radie. Vilka variabler som mätts framgår av tabell 1 (egentligen har mycket mer än så mätts, men ingår inte i den fil ni får till den här övningen).

Tabell 1. Variabler som mätts i fält för provtytorna

Variabel	Beskrivning	Enhet
X_Sweref	X-koordinat, Sweref99TM	m
Y_Sweref	Y-koordinat, Sweref99TM	M
VolumeTotal	Total virkesvolym	m ³ sk/ha
VolumePine	Volym tall	m ³ sk/ha
VolumeSpruce	Volym gran	m ³ sk/ha
VolumeBirch	Volym björk	m ³ sk/ha
VolumeAspen	Volym asp	m ³ sk/ha
VolumeCont	Volym contorta	m ³ sk/ha
BasalArea	Grundyta	m ² /ha
HGV	Grundtyvägd medelhöjd	m
DGV	Grundtyvägd medeldiameter	cm

Data ligger i mappen uppgift_8 (geodata). **Kopiera allt som finns i mappen till din dator!**
Mappen innehåller följande filer:

- KrycklanToPlots.las, LAS-fil med punktdata kring provytorna
- list_clipped_plots.txt, en lista med filnamn och koordinater att ge som input till FUSION-programmet ClipData, se nedan. Filnamnen avser de filer, en per provyta, som ClipData ska skapa. **Sökvägen måste ändras från *path_clipped_plots* till den mapp där du vill lägga filerna! Skapa en egen mapp för dessa filer, och lägg inget annat i mappen.** Sökvägen ändras enklast genom att öppna filen i Notepad (Anteckningar), gå till Edit | Replace och välja att ersätta *path_clipped_plots* med lämplig sökväg.
- Krycklan_fieldplots.txt, fältdata från provytorna.
- Krycklan_fieldplots.shp, fältdata i form av en shapefil.
- Regression_intro.R, ett R-skript som du kan utgå ifrån vid den inledande regressionsanalysen.

Varje provyta har ett unikt ID-nummer, PlotID, som återfinns i Krycklan_fieldplots.txt och i filnamnen i list_clipped_plots.txt.

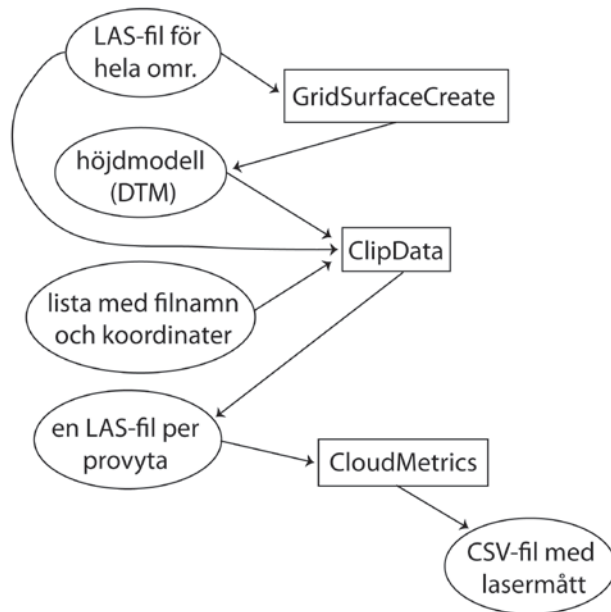
Vill du titta på en karta över området kan den hämtas hem med hjälp av GET-verktyget, se <https://docs.google.com/file/d/0B3yuX8WpOEGESm5RNnBKV0ZfZFk/edit?pli=1> för hjälp och information.

Ta fram lasermått för provytor

De FUSION-verktyg ni kommer att använda är

- GridSurfaceCreate
- ClipData
- CloudMetrics

Flödesschemat i figur 2 beskriver processen.



Figur 2. Flödesschema som visar arbetsgången i FUSION för att ta fram lasermått för provytorna.

Alla tre FUSION-verktyg anropas från en och samma batch-fil. Formen för batch-filen beskrivs nedan (hela anropet av ett verktyg ska skrivas på en och samma rad, så att batch-filen innehåller tre rader). Läs i FUSION-manualen för att förstå vad som görs. Byt ut `path` och `path_clipped_plots` mot lämpliga sökvägar.

```
GridSurfaceCreate /class:2 path\KrycklanToPlotsDTM.dtm 1 M M 0
0 0 0 path\KrycklanToPlots.las
```

```
ClipData /shape:1 /dtm:path\KrycklanToPlotsDTM.dtm /height
/zmin:-10 /zmax:40 path\KrycklanToPlots.las
path\list_clipped_plots.txt
```

```
CloudMetrics /above:2 /id /minht:2 path_clipped_plots\*las
path\KrycklanPlotMetrics.csv
```

Glöm inte att i `list_clipped_plots.txt` ändra sökväg enligt andra punkten på föregående sida. Det är sökvägen till den mapp du skapar som ska ersätta `path_clipped_plots` (i `list_clipped_plots.txt`) och därmed kommer filerna som `ClipData` skapar att hamna där du vill. I `CloudMetrics` kommer sedan ovanstående filer att användas förutsatt att du även där byter sökväg från `path_clipped_plots` till den mapp du skapat.

Verktyget `CloudMetrics` kan ta några minuter att köra eftersom det är mycket data som ska bearbetas. `*las` betyder att samtliga las-filer i aktuell mapp kommer att användas.

När du har kört batch-filen öppnar du CSV-filen i Excel. Fördela data på kolumner (fliken *Data*, knappen *Text to Columns*, bocka för *Delimited* och välj *Comma*) och sortera ID-nummer. Öppna ett nytt Excel-dokument där du först kopierar in innehållet i `Krycklan_fieldplots.txt`, dvs fältmätningarna, och direkt till höger om det kopierar du in innehållet i CSV-filen (dvs lasermåtten). Den ursprungliga `KrycklanPlotMetrics.csv` kan du stänga utan att spara ändringar.

Den nya filen kommer att innehålla två kolumner med PlotID/Identifier, en från Krycklan_fieldplots.txt och en från CSV-filen. Raderna bör vara i rätt ordning, men för säkerhets kan du kolla några rader i början, mitten och slutet av filen: har raden samma ID-nummer i båda ID-kolumnerna? Behåll PlotID och ta bort Identifier (högerklicka på bokstaven över kolumnen och välj *Delete*).

FUSION beräknar väldigt många mått och vi kommer inte använda alla. Nästa steg är därför att rensa upp i filen för att den ska bli mer lättöverskådlig. Ta bort alla kolumner **UTOM** följande:

Total return count
Total return count above 2.00
Elev maximum
Elev mean
Elev stddev
Elev variance
Elev CV
Elev P01 till och med Elev P99
Percentage first returns above 2.00
Percentage all returns above 2.00
Percentage first returns above mean
Percentage all returns above mean
Total first returns

Flera av kolumnnamnen innehåller mellanslag, vilket kommer att ställa till problem när man sedan ska läsa in filen i R. Byt därför ut alla mellanslag mot understreck, `_`. Under *Home*-fliken i Excel, gå till *Find & Select* och välj *Replace*. Vid *Find what:* skriver du in ett mellanslag och vid *Replace with:* skriver du in understreck (`_`). Klicka *Replace all* och sedan *Close*. Kontrollera även att kommatecken är ersatt med punkter.

Spara filen som FieldPlotsMetrics.txt (File | Save As, vid Save as type väljer du *Text (Tab delimited)* (*.txt). Svara ja på alla frågor som dyker upp).

Inledande regressionsanalys

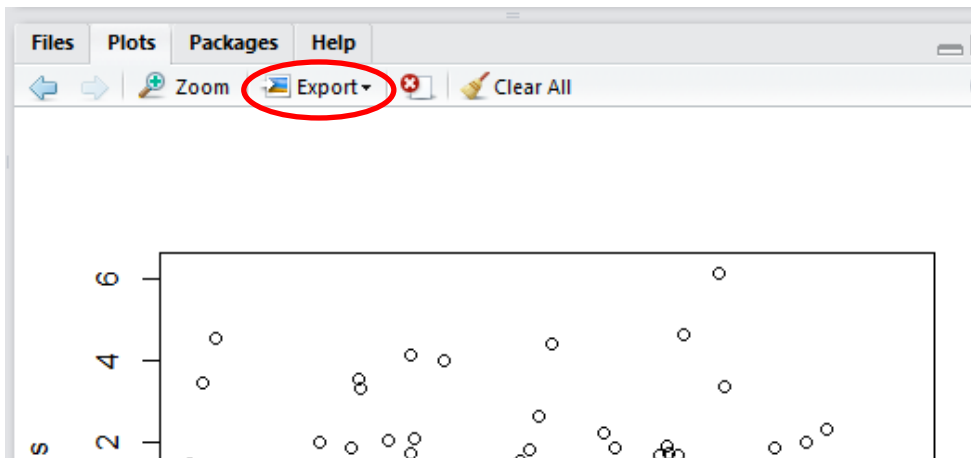
Starta RStudio och öppna Regression_intro.R. Titta igenom filen och läs noggrant vad som står i kommentarerna (kommentarer inleds med #). De lasermått som från början finns med i Regression_intro.R är inte nödvändigtvis de som kommer att vara med i de färdiga regressionsfunktionerna. Meningen är att du själv ska undersöka data och använda de givna exemplen som vägledning i R.

Skattning av grundtevägd medelhöjd

Grundtevägd medelhöjd finns i kolumnen HGV.

- Undersök korrelationen mellan HGV och ett antal lasermått som du tror skulle kunna användas för att skatta höjden (som i uppgift 5, men det är inte säkert att samma mått är "bäst").
- Rita upp ett spridningsdiagram med HGV på y-axeln och utvalt lasermått på x-axeln.
- Ta fram en regressionsfunktion som beskriver HGV som en funktion av detta lasermått. Rita in regressionslinjen i diagrammet. Spara diagrammet genom att klicka

på *Export* och *Save plot as image* (figur 3). **Denna bild ska du sedan klistra in i dokumentet som du lämnar in.**



Figur 3. Plotten sparas genom att gå till *Export | Save plot as image*.

- Plotta residualerna mot skattad HGV (`plot(reg)`).

Fråga 1: Ser du någon tydlig trend i spridningen hos residualerna (t ex större eller mindre spridning för skattad HGV, eller någon krökning)? Verkar en linjär modell fungera?

- Beräkna korrelationen mellan residualerna och andra lasermått som du tror kan vara till hjälp vid skattning av HGV.

Fråga 2: Hittar du något mått som är starkt korrelerat med residualerna? I så fall vilket, och hur stark är korrelationen?

Fråga 3: Plotta residualerna mot det mått som de är starkast korrelerade med. Ser du något samband mellan residualerna och lasermåttet? Baserat på korrelationskoefficienten och plotten, skulle du säga att det lasermåttet tillför något för skattningen av HGV, i tillägg till det mått som redan ingår i modellen?

- När du har bestämt dig för ett eller flera mått att använda för att skatta HGV: Ta fram den slutliga regressionsfunktionen och skriv på skärmen ut en sammanfattning av den (`summary(reg)`). Det ska se ut som i figur 4 men med andra siffror.

Fråga 4: Hur ser regressionsfunktionen ut? Skriv den på formen $y = a_0 + a_1 * x_1 + a_2 * x_2 + \dots$ med skattade värden för koefficienterna a .

Fråga 5: Vad är R^2 -värdet och vad betyder det?

```

> summary(reg2)

Call:
lm(formula = HGv ~ dfPlotData_clean$Elev_mean, data = dfPlotData_clean)

Residuals:
    Min       1Q   Median       3Q      Max
-6.7008 -1.4362  0.0722  1.0095  8.0292

Coefficients:
    (Intercept)          dfPlotData_clean$Elev_mean
    2.33142             1.15571

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.923 on 206 degrees of freedom
Multiple R-squared: 0.7368, Adjusted R-squared: 0.7355
F-statistic: 576.7 on 1 and 206 DF, p-value: < 2.2e-16

```

Figur 4. Sammanfattning av en regressionsfunktion i R. a) skattade värden på b) intercept och c) koefficienten framför lasermåttet som i detta exempel är Elev_mean. Justerad R^2 visas vid d).

Skattning av virkesvolym

Virkesvolym (sammanlagt för alla träslag) finns i kolumnen VolumeTotal.

- Undersök korrelationen mellan volymen och ett antal höjdmått som du tror kan användas för att skatta den. I slutänden är det troligt att det kommer med täthetsmått också, men börja med höjd.
- Börja med det laser(höjd)mått som verkar bäst för skattning av volymen. Rita upp ett spridningsdiagram med volymen på y-axeln och måttet på x-axeln.

Fråga 6: Vilket höjd mått har du valt och varför?

Fråga 7: Ser punkterna ut att följa en linje eller något slags kurva?

- Ta fram en regressionsfunktion med volymen som funktion av lasermåttet, och plotta residualerna mot skattad volym.

Fråga 8: Ser du någon trend i spridningen hos residualerna? Ökar eller minskar den med skattad volym, eller syns inget samband?

- Transformera volymen genom att logaritmera. Spara variabeln lnVolTot i din data frame dfPlotData (**tips på hur man gör** finns i R-koden, just under rubriken Regressionsanalys).
- Gör ett nytt spridningsdiagram med lnVolTot på y-axeln och samma lasermått som sist på x-axeln.

Fråga 9: Samma som fråga 7 – ser punkterna ut att följa en linje eller något slags kurva?

- Transformera den oberoende variabel (lasermått) du använt genom att logaritmera på samma sätt som ovan. Spara variabeln lnMått (byt ut mått mot måttets namn!) i din data frame dfPlotData.
- Gör ett spridningsdiagram med lnVolTot på y-axeln och 'lnMått' på x-axeln.
- Du kan om du vill göra någon annan transformation än logaritmering av höjdv variabeln och prova den istället.

Fråga 10: samma fråga igen – ser punkterna ut att följa en linje eller något slags kurva? För att göra en linjär regression, vad skulle du välja som oberoende variabel ('Mått' eller 'transformeratMått') och vad skulle du välja som beroende variabel (VolumeTotal eller lnVolTot)?

- Plotta residualerna mot skattade värden för modellen du valde i fråga 10. Spara residualplotten (residualer mot skattad VolumeTotal eller lnVolTot) för vald regressionsfunktion och klistra in i dokumentet du lämar in.

Fråga 11: Ser du någon trend i residualerna? Ser det bättre eller sämre ut än residualplotten mellan fråga 6 och 7, när varken volym eller lasermått var logaritmerade?

- Här kan det vara läge att prova ännu ett lasermått (kom ihåg tipset om höjd och täthet!) enligt samma mönster. Fortsätt så långt du vill, och på träffen i Umeå tar du vid där du slutade.

Inlämning

Uppgiften lämnas in i form av ett Word-dokument eller en pdf-fil.

FUSION-delen

- Förklara vad de switchar och parametrar som här har använts i GridSurfaceCreate, ClipData och CloudMetrics fyller för funktion. Försök också att förklara varför värdena har satts som de gjort.

Regressionsdelen

- Klistra in spridningsdiagrammet du sparade när du tittade på HGV, som visar HGV på y-axeln, ett lasermått på x-axeln samt regressionslinjen.
- Klistra in spridningsdiagrammet du sparade vid fråga 10, som visar residualer mot skattade värden för den modell du valt.
- Besvara frågorna 1-11.

Lycka till!

Uppgift 9 - Tät skanning, enskilda träd

Vid tät laserskanning (minst 5-10 pulser/m²) får man så pass många träffar per träd att det går att detektera enskilda träd i laserdata och att skatta variabler för dessa. De inledande stegen av databehandlingen, fram till beräkning av punkternas höjd över marken, går till på samma sätt som vid gles skanning för areabaserad skattning. Därefter görs följande:

- Segmentering för att detektera enskilda trädkronor
- Beräkning av lasermått som beskriver de enskilda träden
- Hopkoppling av träd (segment) i laserdata med träd som inventerats och koordinatsatts i fält
- Regressionsanalys baserad på lasermått och fältdata för fältinventerade träd
- Skattning, dvs tillämpning av regressionsfunktionerna på alla träd som detekterats i laserdata.

I den här övningen får du titta på tätt skannat laserdata och resultatet av en segmentering. Därefter får du färdiga lasermått och fältdata och får undersöka sambanden mellan dessa. Du behöver inte göra en fullständig regressionsanalys.



Viktigt

Innan du börjar: kopiera de filer som anges under rubriken *Data och mjukvara* nedan och lägg dem på din dators hårddisk (oftast C:). Enklast är om du skapar mappen C:\Uppgift9 alt C:\Fusion\Uppgift9 och lägger allt som rör övningen där.

Om du sitter i en datasal kan saker på C: försvinna när du loggar ut. Därför måste du flytta hela mappen till din SLU-hemkatalog (Windows startknapp -> Documents) innan du loggar ut. Du kan då komma åt mappen från vilken SLU-dator som helst. Alternativt kan du flytta övningsmappen till en sticka / extern disk.

Läsning

Innan du sätter igång bör du ha läst inledningen till kapitel 7 i kompendiet.

Data och mjukvara

Du kommer att använda Fusion för att titta på punktmolnet och R (via RStudio) för att studera sambanden mellan lasermått och fältdata.

Laserdata kommer från en tät laserskanning som gjorts alldeles utanför SLU-byggnaden i Umeå. Fältdata kommer från två provytor där träden koordinatsatts och inventerats. Data finns i mappen uppgift_9 (geodata). Kopiera hela mappen till din dator.

Mappen innehåller följande filer:

- SingleTree.R, en R-fil med några färdiga kommandon att använda för att undersöka data
- SLU_forest.las, laserdata från tät skanning

- HeightImg_SLU_forest.bmp, en referensbild baserad på punkternas höjd över marken (till skillnad från tidigare övningar då referensbilden har baserats på intensitetsvärden)
- poly2287.shp, en shapefil med segmenteringsresultat från en del av det skannade området
- ytor_Buf.shp, en shapefil med konturerna för de provytor där fältdata samlats in
- tree_ID1.shp, tree_ID5.shp och tree_ID7.shp, segmenten som motsvarar tre av de fältmätta träden
- Ett antal foton (jpg) på en tall, en gran och ett dött träd (träd nr 1, 5 och 7 i shapefilerna ovan)
- TreeData.txt, en textfil med fältdata från träden på provytorna, samt höjd och vidd för motsvarande segment. Varje rad motsvarar ett träd.

Lasermått

- **H:** höjden hos det segment som kopplats till det fältmätta trädet (m)
- **W:** vidden hos det segment som kopplats till det fältmätta trädet (m)

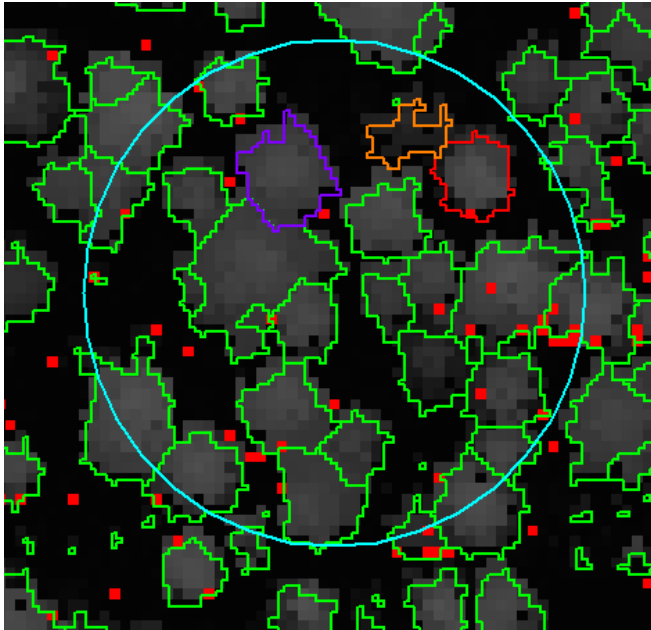
Fältdata

- **XP_GROUND, YP_GROUND:** koordinater för det fältmätta trädet
- **SPECIES:** trädslag. 1=tall, 2=gran, 3=björk, 4=asp
- **DBH:** brösthöjdsdiameter mätt i fält (cm)
- **TH:** trädhöjd mätt i fält (m)
- **V:** Stamvolym skattad från fältmätningar mha volymfunktioner (dm³)

Övningsinstruktioner

Starta Fusion och läs in referensbild, laserdata och de olika shapefilerna. Shapefilerna läser du in genom att klicka på knappen POI (under Image och Raw data i Fusionfönstret). POI står för *Point of Interest* men kan vara vilken sorts shapefil som helst: punkt, linje eller polygon. Det är bra om du läser in shapefilerna i den ordning de listas ovan. Då hamnar segmenteringen i botten (men visas överst i listan...) och skymmer inte de tre trädsegmenten 1, 5 och 7. Se till att rutorna framför Image och POI är ikryssade, men lämna rutan vid Raw data okryssad så slipper du se det stora kryss som visar var rådata finns.

För att det ska bli tydligare vad som är vad kan du ge de olika shapefilerna olika färg. Klicka på POI-knappen så kommer det upp en lista över shapefilerna. Dubbelklicka på en rad så att fönstret *Data layer properties* kommer upp. Där kan du ändra *Line color* för aktuell shapefil. Använd till exempel grön för segmenteringen (poly2287), turkos för provytorna (ytor_Buf) och tre olika färger för de tre trädsegmenten 1, 5 och 7:



Gör Fusion-fönstret stort, gärna hela skärmen. Zooma in genom att hålla in ctrl och rulla på scrollhjulet. Flytta dig runt i bilden med hjälp av piltangenterna. Om du tappat bort dig kan du alltid klicka *Zoom to extents* i menyn till vänster.

Zooma in på provytorna och titta på punktmolnet för vart och ett av träden 1, 2 och 3. Jämför med fotona. Kom ihåg att du kan ändra storlek och form på punkterna i Lidar Data Viewer genom att högerklicka och gå till *Marker* i menyn.

Fråga 1: Vilket segment (1, 5 eller 7) hör till granen, tallen respektive det döda trädet?

Titta nu på punktmolnet i olika delar av området. Titta på segmenten och på referensbilden. Hittar du något träd som delats upp i mer än ett segment? Något segment som verkar innehålla mer än ett träd? Du behöver inte besvara några frågor på den här delen av uppgiften, men ta dig ändå tid att titta runt en stund! I Lidar Data Viewer kan du högerklicka och välja *Image plate* så syns referensbilden under punktmolnet.

Starta nu RStudio och öppna R-skriptet *SingleTree.R*. Titta igenom innehållet för att få en bild av vad de olika kommandona gör. De funktioner som används är samma som i tidigare uppgifter.

Skapa två nya variabler utifrån lasermåtten W och H , och lägg dessa i din data frame *dfTrees*: den ena variabeln beräknar du som $W*H$ och kallar WH . Den andra beräknar du som W^2*H (dvs W^2*H) och kallar $W2H$. Undersök nu sambanden mellan de olika lasermåtten (W , H , WH , $W2H$) och fältmätt trädhöjd, brösthöjdsdiameter och volym. Använd spridningsdiagram och korrelationskoefficient. Leta efter linjära samband.

Fråga 2: Vilket lasermått har starkast linjärt samband med fältmätt trädhöjd, och vad är korrelationskoefficienten?

Fråga 3: Vilket lasermått har starkast linjärt samband med fältmätt brösthöjdsdiameter, och vad är korrelationskoefficienten?

Fråga 4: Vilket lasermått har starkast linjärt samband med fältmätt volym, och vad är korrelationskoefficienten?

Inlämning

Lämna in svar på frågorna 1-4.

Lycka till!

Uppgift 10– jämför punktmoln från olika sensorer

Såväl laserskanning som digital fotogrammetri kan användas för att skapa tredimensionella punktmoln. Dessa har flera likheter men också många olikheter. Beroende på vilka instrument som används och hur databearbetningen görs finns det dessutom skillnader mellan olika fotogrammetriska punktmoln, och mellan olika laserpunktmoln. Dessa skillnader gör att olika tekniker passar för olika tillämpningar. I den här övningen får du titta närmare på två punktmoln från laserskanning och två från matchning av digitala stereoflygbilder (digital fotogrammetri). **Läs igenom hela övningsinstruktionen innan du sätter igång!**



Viktigt!

Innan du börjar: kopiera de filer som anges under rubriken Data nedan och lägg dem på din dators hårddisk (oftast C:). Enklast är om du skapar mappen **C:\Uppgift_10** alt **C:\Fusion\Uppgift_10** och lägger allt som rör övningen där.

Om du sitter i en datasal på SLU kan saker på C: försvinna när du loggar ut. Därför måste du flytta hela mappen till din SLU-hemkatalog (Windows startknapp -> Documents) innan du loggar ut. Du kan då komma åt mappen från vilken SLU-dator som helst. Alternativt kan du flytta övningsmappen till en sticka /extern disk.

Läsning

Innan du börjar med övningen bör du ha läst följande delar av kompendiet:

- Inledningen till kapitel 7
- Inledningen till kapitel 8

Under övningens gång kan du ha nytta av att titta i resten av kapitel 7 och 8. Avsnitt 8.5 beskriver hur man framställer punktmoln från stereobilder. Avsnitt 8.4 är ganska tekniskt och läses bäst i lugn och ro hemma.

Data

I den här övningen används data från ett ca 1 km² stort område i Krycklan i Västerbotten, en del av det område som används i skattningsuppgifterna. Data till övningen finns i mappen uppgift_10 (geodata). Kopiera innehållet i mappen till en egen mapp som du skapar på datorns hårddisk (oftast C:). Mappen innehåller följande filer:

- Punktmoln_A.las, Punktmoln_B.las, Punktmoln_C.las: tre punktmoln, alla framställda med olika tekniker. Projektionen är RT90.
- Punktmoln_D.las: ett punktmoln framställt med en fjärde teknik. Projektionen är Sweref 99 TM.
- RefImgHeight_ABC.bmp: den referensbild som ska användas till punktmolnen A, B och C. Baserad på punkternas höjd över marken. 1 m pixelstorlek. Projektionen är RT90.
- RefImgHeight_D.bmp: den referensbild som ska användas till punktmoln D. Baserad på punkternas höjd över marken. 2 m pixelstorlek. Projektionen är Sweref 99 TM.
- RefImgHeight_ABC.bmpw och RefImgHeight_D.bmpw är så kallade *world*-filer som behövs för georefereringen av bmp-filerna.

- TopographicMap.jpg: en kartbild i jpg-format över området, för att underlätta orienteringen. Bilden är inte georefererad så använd ett vanligt bildvisningsprogram för att titta på den, t ex Windows Photo Viewer.

Punktmolnen har framställts på olika sätt. En av uppgifterna nedan är att bestämma vilket punktmoln (A-D) som framställts med vilken teknik (1-4):

1. Laserskanning från flygplan på ca 2000 m höjd. Flyghastighet ca 280 km/h.
2. Laserskanning från helikopter (= låg flyghastighet) på ca 250 och 500 m höjd. Området innehåller överlappande stråk. Punkterna har dessutom färgats med hjälp av digitala flygbilder (grönt, rött, NIR).
3. Pixelvis matchning av digitala flygbilder (*semi-global matching*).
4. Objektsbaserad matchning av digitala flygbilder (*feature based matching*).

Samma flygbilder har använts för både pixelvis och objektsbaserad matchning. De är tagna med Lantmäteriets Zeiss/Intergraph DMC kamera, med 60 % överlapp i flygriktningen och 30 % överlapp mellan stråken. Vid **objektsbaserad matchning** identifierar matchningsprogramvaran lämpliga objekt i bilderna och matchar dessa. **Pixelvis matchning** är en nyare metod som innebär att var n :te pixel matchas, där n kan vara till exempel 3 för var 3:e pixel. Du kan läsa mer om matchning i **kompendiet, avsnitt 8.5**.

Övningsinstruktioner

När du ska undersöka punktmolnen kan du lägga upp arbetet på två olika sätt. Det spelar ingen större roll vilket – läs och välj själv!

- Ett sätt är att göra ett Fusion-projekt för varje punktmoln. Öppna först Punktmoln_A (läs först in referensbilden) och spara Fusion-projektet som Punktmoln_A.dvz (Gå till File | Save As i Fusions huvudfönster). Gör sedan samma sak för övriga 3 punktmoln. Du kan nu enkelt ”zappa” mellan projekten genom att gå till File | Open och välja önskat projekt.
- Alternativt kan du göra ett projekt för punktmoln A, B och C (som använder samma referensbild) och ett för punktmoln D (som använder en annan referensbild pga annan projektion). Du måste då vara noga med att **bara ha ett punktmoln i taget inläst som Raw Data**, eftersom du annars kommer se punkterna från alla punktmoln samtidigt i Lidar Data Viewer.

När du ska titta på en del av punktmolnet i Lidar Data Viewer är det en bra idé att börja med ett mycket litet område, innan du vet hur mycket datorn klarar av utan att det tar för lång tid. **För att kunna titta på större områden utan att det tar för lång tid:** klicka på *Sample options* i Fusions huvudfönster. Under *Decimation* skriver du in t ex 10 för att bara visa var 10:e punkt. Hur stora områden du kan titta på utan att glesa ur punktmolnet beror på datorn och punkttätheten.

För att få hjälp med orienteringen i Lidar Data Viewer kan du högerklicka och bocka för *Image plate*. Då visas referensbilden under punktmolnet. Använd musens scrollhjul för att höja och sänka bilden.

Punkternas storlek i Lidar Data Viewer kan ändras genom att högerklicka och välja *Marker*. I fönstret som kommer upp kan man ändra på *Marker size* och *Marker type*.

Titta på följande egenskaper hos punktmolnen:

- Punkttäthet
- Punkternas fördelning i höjd
- Eventuellt mönster som bildas av punkterna när du tittar rakt uppifrån

Du kan också ha nytta av att **jämföra LAS-filernas header**: Tools | Miscellaneous utilities | Examine LAS file headers...

Fråga 1: Vilken teknik 1-4 har använts för att framställa punktmoln A? Varför tror du det?

Fråga 2: Vilken teknik 1-4 har använts för att framställa punktmoln B? Varför tror du det?

Fråga 3: Vilken teknik 1-4 har använts för att framställa punktmoln C? Varför tror du det?

Fråga 4: Vilken teknik 1-4 har använts för att framställa punktmoln D? Varför tror du det?

Fråga 5: Hur bra lyckas de olika fotogrammetriska metoderna ”tränga igenom” trädskronorna?

Fråga 6: Hur påverkar detta a) höga percentiler, b) låga percentiler och c) täthetsmått som beräknas ur fotogrammetriska punktmoln istället för ur laserpunktmoln?

Fråga 7: I området finns en tjärn. Titta på det fotogrammetriska punktmoln som gjorts med objektbaserad matchning, och titta på hela tjärnen i Fusions Lidar Data Viewer. Hur ser punktmolnet ut över tjärnen? Varför?

Fråga 8: Som frågan ovan, men titta istället på det punktmoln som gjorts med pixelvis matchning.

Fråga 9: Vilken typ av skanner har använts vid skanningen från helikopter på låg höjd? Varför tror du det? Tips: titta på punktmolnet uppifrån i ett område utan skog. Kom ihåg att du kan gå till *Sample options* och ändra siffran vid *Decimation* för att titta på ett utglesat punktmoln (snabbare, och gör att du kan titta på ett större område).

Fråga 10: Vilket punktmoln skulle du välja för att göra en skattning på enskilda träd-nivå?

Fråga 11: Skulle ett sådant punktmoln fungera för areabaserade skattningar?

Fråga 12: Oavsett ditt svar på fråga 11: vad talar emot att använda sådana punktmoln för areabaserade skattningar?

Inlämning

Svara på frågorna 1-12.

Lycka till!

Uppgift 11 – Areabaserad skattning och utvärdering (bygger på uppgift 8)

I uppgift 8 har du tagit fram regressionsfunktioner för skattning av grundytavägd medelhöjd HGV (m) och volym V (m³sk/ha).

I uppgift 11 ska du ta fram en regressionsfunktion för grundytavägd medeldiameter DGV (cm), på ett liknande sätt som du gjorde för HGV och Volym. För att underlätta har vi skrivit ihop en ny R kod som du kan använda.

De regressionsfunktioner som du har tagit fram för HGV, V, och DGV ska du sedan använda för att göra skattningar på beståndsnivå. Arbetsgången kan sammanfattas i följande steg:

- Applicera funktionerna på raster över hela området för att skatta HGV, V och DGV. Skippa grundyta eftersom denna variabel saknas i utvärderingsdata.
- Aggregera skattningsresultaten till beståndsnivå genom att beräkna medelvärden inom bestånden
- Utvärdera resultatet mot fältinventerade bestånd
- Diskutera resultaten i mindre grupper (se detaljer under Inlämning nedan).

Läsning

Kapitel 6 i kurskompendiet.

Data

Data till övningen finns i mappen uppgift_11 (geodata).

Du kommer att använda data från samma laserskanning som i uppgift 8, men i den här uppgiften får du färdiga raster. Rastercellerna har 18 m sidlängd, för att få ungefär samma area som provytorna som användes i uppgift 8. Rastren har skapats med hjälp av Fusion. Filformatet är ASCII. Vilka raster du behöver använda beror på vilka regressionsfunktioner du har kommit fram till i uppgift 8.

I datamappen finns också katalogen field_stands som innehåller en shapefil med utvärderingsbestånd. Filen innehåller bland annat fältmätta beståndsmedelvärden för volym, medelhöjd och medeldiameter. Kopiera hela field_stands-katalogen till din dator.

Dessutom innehåller mappen ett utsnitt av terrängkartan och ortofoto från Lantmäteriet.

Mjukvara

I den här uppgiften kommer du att använda ArcMap. Fusion behövs inte eftersom du får färdiga raster. I ArcMap måste du gå till Customize | Extensions och se till att Spatial Analys är ikryssat.

För utvärderingen används R (Rstudio). Gör ditt eget R-skript, men använd R-skripten från tidigare uppgifter, samt tips nedan, för att få till syntaxen.

Instruktioner

Regressionsfunktion för DGV

Öppna filen Regression_DGV.R i R-Studio och följ kommentarerna med instruktioner. Kom fram till en regressionsfunktion för DGV som du är nöjd med, och skriva ner regressionsfunktionen med koefficienterna. (Kom ihåg att om det behövs, kan du titta tillbaka på uppgift 8:s instruktioner – men R koden är ganska instruktiv).

Ta fram dina resultat från uppgift 8, dvs de regressionsfunktionen du gjorde till Total Volym och HGV.

Några ord om logaritmer

Kanske har du logaritmerat den beroende variabeln vid regressionsanalysen för att skapa en bra funktion till TotVolym, HGV eller DGV. Eftersom det inte är de logaritmerade värdena man vill skatta egentligen, t ex $\ln(\text{volym})$, utan bara vanlig volym rakt upp och ner, måste man transformera tillbaks till vanliga värden. **Kom ihåg detta när du skattar en variabel som du har logaritmerat i regressionsfunktionen!**

Ett (påhittad) exempel på en logaritmerad regressionsfunktion skulle vara:

$$\ln(\text{DGV}) = 0,33 + 2,88 * \ln(\text{P80}) - 0,55 * \ln(\text{Elev_mean})$$

Återtransformering skulle då vara:

$$\text{DGV} = e^{0,33} * \text{P80}^{2,88} * \text{Elev_mean}^{-0,55}$$

Men notera följande skillnad när den oberoende variabeln inte är logaritmerad:

$$\ln(\text{DGV}) = 0,33 + 2,88 * \ln(\text{P80}) - 0,55 * (\text{Elev_mean})$$

Återtransformering skulle då vara:

$$\text{DGV} = e^{0,33 - 0,55 * \text{Elev_mean}} * \text{P80}^{2,88}$$

För att förstå detta, jämför ekvation 1 och 3 samt 2 och 4 i texten ”regression i praktiken” som ingick som litteratur till uppgift 8.

För att mjukstarta är det bra om du börjar med att skatta en variabel som har en enkel regressionsfunktion utan logaritmer.

När man använder logaritmer i regressionsfunktionen och sedan går tillbaks till vanliga värden uppstår ett fel som kallas **logaritmiskt bias**. Det brukar vara ganska litet, och för att inte krångla till saker för mycket bortser vi från det i den här kursen.

Förberedelse

Lägg in terrängkartan och ortofotot i ArcMap så att du får en överblick över området.

Bestäm nu en skoglig variabel som du vill börja med att skatta, förslagsvis höjden eftersom den brukar ha en enkel regressionsfunktion. Kolla vilka lasermått som ingår i funktionen och läs in motsvarande raster i ArcMap.

Ifall rastren ser jämngrå eller helsvarta ut måste du kontraststräcka dem. Dubbelklicka på rastrets namn för att öppna *Properties*-fönstret. Gå till fliken *Symbology*. Under *Stretch*, vid

Type öppnar du rullgardinsmenyn och väljer *Custom*. Får du frågan om du vill skapa histogram svarar du ja. I regel räcker det att sedan klicka OK och stänga *Properties*-fönstret. Annars måste du fixa kontraststräckningen manuellt, alternativt välja t ex *Standard Deviations* istället för *Custom*. Fråga om du inte vet hur man gör.

Tänd och släck rastren och jämför med ortofotot och terrängkartan för att kontrollera att allt ser normalt ut.

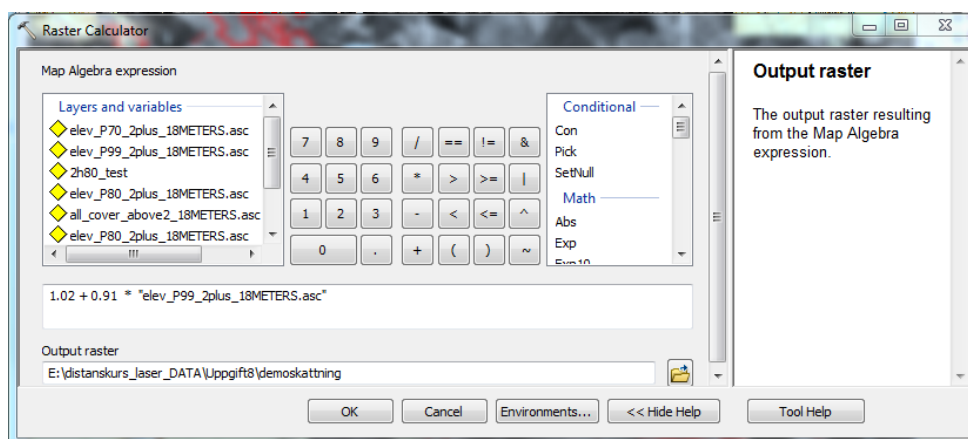
Skattning

Nu kommer själva skattningen. Du har tillgång till

1. En regressionsfunktion som beskriver en skoglig variabel som funktion av ett eller flera lasermått
2. Ett raster för vart och ett av dessa lasermått.

Idén är då att för varje rastercell stoppa in cellvärdet/cellvärdena (från rastren med lasermått) i regressionsfunktionen och få ut ett skattat värde på den skogliga variabeln i just den cellen. Detta görs för alla celler och resultatet blir ett raster med skattade värden för aktuell variabel.

Verktyget för detta heter *Raster Calculator* och finns i ArcToolbox | Spatial Analyst Tools | Map Algebra. Starta Raster Calculator och förstora fönstret så att det ser ut som i figur 1.




Figur 1. Raster Calculator.

Den övre vänstra rutan innehåller de raster som finns inlästa i ditt ArcMap-projekt. I mitten finns ”miniräknarknappar”. I rutan nedanför skriver du in regressionsfunktionen, och i den nedersta rutan bläddrar du till den mapp där du vill spara rastret med skattningsresultatet. Döp det till skattad_hojd eller liknande.

För att skriva in regressionsfunktionen: Ställ dig i den tomma rutan och skriv in funktionen. När du kommer till den oberoende variabeln (det lasermått som ingår i funktionen) letar du upp motsvarande raster i den övre vänstra rutan och dubbelklickar. Rastrets namn dyker då upp i funktionsrutan omgivet av citationstecken. Observera att när du ska skriva in exponenter i funktionen bör du använda två gångertecken (**) istället för det vanliga tecknet för upphöjt (^), detta är vanligt förekommande i programmeringsspråk. Skriv in tecknen med tangentbordet istället för knappen i *Raster Calculator* för att undvika mellanslag. Ett tips är även att arbeta med parenteser när du skriver in funktionen. Det är även viktigt att välja rätt

decimalavskiljare (punkt eller komma). ArcMap vill troligtvis att det ska vara punkt men fungerar inte detta använd det som datorn är inställd på. Om du klickar på knappen för decimalavskiljare, till höger om nollan i sifferdelen i *Raster Calculator*, så borde det bli rätt.

Välj var du vill spara ditt ”skattade” raster och vad det ska heta. Klicka OK och vänta. Rastret ska automatiskt öppnas i ArcMap när det är klart.

Använd nu *Identify*-knappen  för att undersöka om de skattade värdena verkar rimliga. I *Identify*-fönstret som kommer upp när du klickar i bilden väljer du rätt raster under *Identify from*, och tittar vad som står vid *Pixel value*. Titta på några olika delar av området, med olika typer av skog, och använd ortofotot för att få en uppfattning om ifall värdena verkar rimliga.

Beräkning av beståndsmedelvärden

Nästa steg är att beräkna medelvärden på beståndsnivå. Du kommer att göra detta för utvärderingsbestånden, de som finns i *field_stands.shp*. Det går till så att ArcMap kollar vilka rasterceller som ligger inom gränserna för ett bestånd, beräknar medelvärdet av dessa och tilldelar hela beståndet detta medelvärde.

Verktyget som används heter *Zonal Statistics as Table* och finns i ArcToolbox | Spatial Analyst Tools | Zonal.

Under *Input raster or feature zone data* anges vilket lager som ska användas för att definiera de områden där medelvärdet ska beräknas. Här väljer du shapefilen med beståndsgränser (*field_stands*).

Vid *Zone field* låter du FID vara kvar som det står.

Vid *Input value raster* anges för vilket raster du vill beräkna medelvärdet. Välj ditt raster med skattade värden, t ex *skattad_hojd*.

Vid *Output table* bläddrar du till den mapp där du vill spara utdata, och döper tabellen till exempelvis *skattad_hojd.dbf*.

Låt *Ignore NoData in calculations* vara ikryssat. Annars räcker det att ett bestånd har en enda cell utan data för att hela beståndet ska få värdet NoData.

Under *Statistics type* väljer du MEAN, dvs medelvärde.

Klicka OK. Tabellen med genomsnittlig skattad höjd per bestånd kommer nu att dyka upp i *Table of Contents* i ArcMap (troligen längst ner i listan).

Lägg ihop skattade data med fältdata

För att kunna utvärdera resultatet av skattningen är det praktiskt att ha skattade data och fältdata i en och samma tabell. Detta kan ordnas genom att göra en så kallad *Join* mellan *field_stands.shp* och *skattad_hojd.dbf*. För att rätt rader, dvs bestånd, ska kopplas ihop använder vi oss av fältet FID, ett slags id-nummer som ursprungligen fanns i *field_stands.shp* och som fördes över till tabellen *skattad_hojd.dbf* i och med att vi valde FID som *Zone field* när vi beräknade medelvärden ovan.

Högerklicka på tabellen skattad_hojd i *Table of Contents*. Välj Joins and Relates | Join. Under punkt 1, *Choose the field in this layer that the join will be based on* väljer du FID. Under punkt 2, *Choose the table to join...* väljer du field_stands. Under 3, *Choose the field in the table to base the join on*, väljer du FID. Klicka OK. Klicka Yes i rutan som kommer upp.

Om du nu öppnar skattad_hojd ser du att den nu även innehåller kolumnerna (fälten) från field_stands.shp. De finns egentligen inte i själva tabellen, utan visas bara där på grund av kopplingen (*Join*).

För att skapa en fil som innehåller alla kolumner ”på riktigt” högerklickar du på tabellen skattad_hojd och väljer Data | Export. Vid Export väljer du *All records* och vid *Output table* väljer du sökväg och namn till den fil du vill spara. Exporten ska genererar en dbf-fil. Om du inte har Excel kan du redan här skapa en txt-fil (Tab delimited) istället för dbf-fil.

Tips vid krångel: Om du har klickat på mapp-knappen, bläddrat dig fram till rätt katalog och kanske bytt filtyp så kan det hända att du bara får *Error* när du sedan klickar OK. Prova i så fall att bara ändra sökväg och namn direkt i *Output table*-fältet genom att manuellt skriva in sökväg mm.

Nu kan du i ArcMap högerklicka på tabellen skattad_hojd, gå till Joins and Relates | Remove Join(s) och välja *Remove all Joins* så kommer kopplingen till field_stands att försvinna.

Utvärdering

En enkel utvärdering av resultatet kan göras i R. Skapa ditt eget R-skript i RStudio (använd encoding UTF8 om du får frågan när du sparar). Filändelse för R-skript är .R.

Innan du börjar med utvärderingen i R måste du spara om dina dbf-filer med skattade beståndsmedelvärden till txt-filer. Öppna din dbf-fil med skattade beståndsmedelvärden i Excel. Gå till "Spara som" och spara filen som textfil (Tab delimited), stäng filen och svara ja till alla frågor. Öppna den sparade textfilen med Notepad. Gå ner till slutet av filen och ta bort den sista tomma raden med tangenten "backsteg". Gör detta med alla dina dbf-filer med skattade beståndsmedelvärden innan du börjar med utvärderingen i R.

Två vanliga mått på skattningens kvalitet är RMSE och bias som beskrivs nedan. För att utvärdera höjdsfattningen läser du in din textfil innehållande både fältdata (i kolumnen MHeight om det är höjd du tittar på) och skattad variabel (i kolumnen MEAN). Läs in data till en data frame som du kallar dfHojd. Använd exempelkod från tidigare uppgifter för att få syntaxen rätt!

RMSE står för *Root Mean Square Error* och är detsamma som roten ur medelvärdet av de kvadrerade felen. Man vill alltså att RMSE ska vara lågt. I R skrivs kvadratroten som sqrt, medelvärde som mean och ”upphöjt till två” som ^2. Använd uttrycket

```
rmse_hojd <- sqrt(mean((dfHojd$MEAN-dfHojd$MHeight)^2))
```

för att beräkna RMSE för höjdsfattningen (namnet rmse_hojd kan du givetvis byta ut till vad du vill). För att skriva ut RMSE-värdet på skärmen skriver du bara rmse_hojd och

trycker enter. Enheten på RMSE är samma som på den skattade variabeln, i det här fallet meter. Det är också intressant att känna till relativt RMSE, uttryckt i %. Detta beräknar du som RMSE dividerat med medelvärdet för fältmätt höjd (och multiplicerar med 100 för att få procent), dvs

```
rel_rmse_hojd <- rmse_hojd/mean(dfHojd$MHeight) * 100
```

Bias betyder ungefär *skevhets*. Positiv bias innebär systematisk överskattning, och negativ bias innebär systematisk underskattning. Man vill därför ha ett bias nära noll. Bias beräknas som genomsnittlig skillnad mellan skattade och fältmätta värden. I R kan du använda uttrycket

```
bias_hojd <- mean(dfHojd$MEAN-dfHojd$MHeight)
```

För att få bias i procent av skattat värde använder du

```
rel_bias_hojd <- bias_hojd/mean(dfHojd$MHeight) * 100
```

Även **spridningsdiagram** är illustrativa vid utvärderingen. Börja med att plotta skattad beståndsmedelhöjd mot fältmätt höjd,

```
plot(dfHojd$MHeight, dfHojd$MEAN)
```

Vid en perfekt skattning utan några fel skulle alla punkterna i diagrammet hamna på en linje som går genom origo och har lutningen 1, det vill säga $y = 0 + 1 * x$. Om detta inte känns självklart så fundera en stund på saken! Du kan enkelt rita in en sådan linje i spridningsdiagrammet:

```
abline(0,1)
```

Nollan står för intercept=0 och ettan för lutning=1.

Notera dina värden på RMSE, relativ RMSE, bias och relativ bias. Gör nu om proceduren för volym och medeldiameter.

Inlämning

Redovisningen av den här uppgiften görs i mindre grupper.

Börja med ett inlägg (**steg 1**) där du redovisar

- dina regressionsfunktioner från uppgift 8 och 11, inklusive justerade R^2 -värden
- hur du har gått till väga när du tagit fram funktionerna: vilka oberoende variabler har du med och varför? När och varför bestämde du dig för att funktionen var klar? Du kan hålla det ganska kortfattat med försök att vara konkret och beskriva hur du har tänkt!
- RMSE och bias (även relativa) från din utvärdering.

När du har läst vad övriga gruppmedlemmar skrivit gör du **minst** ett reflekterande inlägg i forumet som **Steg 2**. Du ska diskutera följande:

- likheter och olikheter i era val av variabler och funktioner, och dina tankar om detta.
- vilka variabler som var lätta respektive svåra att ta fram en regressionsfunktion för. Har alla upplevt samma variabler som lätta/svåra eller är det olika inom gruppen?
- har de "lätta" variablerna ovan också fått bäst skattningsresultat? Titta både på R^2 , RMSE och bias. Har de "svåra" fått sämre resultat?
- varför tror du att de variabler som fick bra resultat fick det, och varför tror du att de som fick dåligt resultat fick det?
- vad har du lärt dig av att ta del av övriga gruppmedlemmars resultat oh funderingar?
- finns det några variabler vars funktioner är väldigt lika varandra? (Alltså inte att Kalle och Pelle har fått snarlika funktioner för höjd, utan att Kalle har fått nästan samma funktion för skattning av variabel A och variabel B.) Varför?

Diskussionen (Steg 2) utgör den sista delen av uppgiften.

Lycka till!

Uppgift 12 – Areabaserad skattning med Random Forest

I uppgift 8 och 11 har ni använt regressionsanalys för att skatta skogliga variabler för rasterceller med 18 m sidlängd, aggregerat skattningen till beståndsnivå och utvärderat resultatet. Man kan få goda resultat med relativt enkla metoder, men det är ett hantverk som kräver både statistikkunskap och ämneskunskap. Som ni har märkt är det inte alltid så enkelt att bestämma sig för vilken regressionsfunktion man ska använda, och olika personer kommer fram till olika funktioner. Tänk om det räckte att trycka på en knapp istället, och låta datorn göra jobbet! Faktum är att det finns sådana metoder också.

Ett exempel är *best subset regression* där man använder regression men låter datorn välja modell enligt vissa kriterier. En nackdel med denna metod är att datorn inte har några som helst ämneskunskaper och att användaren inte riktigt har koll på vad som händer. Man bör därför inte lita blint på datorns resultat även om det kan vara en god hjälp.

Det finns också metoder som inte har med regression att göra. En sådan är Random Forest, som beskrivs i avsnitt 6.4 i kompendiet. Till fördelarna med metoden hör att man kan använda ett stort antal oberoende variabler även när dessa är starkt korrelerade med varandra, att man slipper bygga modeller manuellt och att resultaten brukar bli ungefär lika bra som för regression. Till nackdelarna hör att det är något av en "svart låda" – det är inte så lätt att se hur det egentligen går till.

Väldigt kortfattat går metoden ut på att datorn själv konstruerar ett antal regressionsträd med hjälp av data från provytorna. Regressionsträd liknar klassningsträd som illustreras i figuren i avsnitt 6.4, men slutnoderna representerar inte klasser. Istället är träden konstruerade så att varje slutnod innehåller en liten, relativt homogen (med avseende på den variabel som ska skattas) grupp av provytor. När man sedan ska skatta exempelvis höjd för varje pixel i ett raster, så körs varje rastercell genom alla träd, och skattat värde beräknas som medelvärde av resultaten från träden.

Syftet med den här uppgiften är att ni ska få

- prova på att använda Random Forest för skattning
- jämföra resultaten med vad ni fått fram med hjälp av regression

Övningen består av följande delar

- använda fältdata och laserdata från provytor (Krycklan-området) för att träna Random Forest-algoritmen för skattning
- applicera resultatet på raster för skattning av hela området
- aggregera resultatet till beståndsnivå (beräkna beståndsmedelvärden)
- utvärdera resultatet och jämföra med uppgift 8 och 11
- undersöka vilka lasermått som vägt tyngst i Random Forest-skattningen, och jämföra det med vilka mått ni använt i regressionsanalysen i uppgift 8 och 11.

Läs igenom alla instruktioner innan du börjar, även inlämningsdelen.

Läsning

Avsnitt 6.4 i kurskompendiet.

Den som vill sätta sig in i metoden mer i detalj kan titta på <http://www.stat.berkeley.edu/~breiman/RandomForests>.

Data

Kopiera data från mappen uppgift_12 (geodata) till din dator:

- KrycklanPlots.txt, med fältdata och laserdata från provytorna
- KrycklanRaster72.gri och KrycklanRaster72.grd: en fil bestående av 72 band (raster) ”staplade” på hög, ett band per lasermått (samma mått som i provytefilen ovan).
- RF_skattning.r: R-skript för Random Forest-skattning

Ta också hem följande data från mappen uppgift_11 (geodata):

- Fältdata från utvärderingsbestånd, i mappen field_stands

Mjukvara

För Random Forest-skattningen används R (RStudio). Du får ett färdigt skript att använda som mall.

Beräkningen av beståndsmedelvärden görs i ArcMap. Du måste gå till Customize | Extensions och se till att Spatial Analyst är ikryssat.

För utvärderingen används R (Rstudio). Utvärderingen görs på samma sätt som i uppgift 11.

Instruktioner

För skattningen används Random Forest-metoden. Använd R-skriptet RF_skattning.r som mall. Du måste själv se till att sökvägarna stämmer, och anpassa skriptet efter vilken variabel du ska skatta. Läs noggrant igenom skriptet, inklusive kommentarer, för att förstå vad som görs.

Förberedelser

Läs igenom rad 1-9 i skriptet och installera de paket som står på rad 14-20. Observera att installationen görs via *Tools* i övre menyraden i RStudio, och bara behöver göras en gång per dator. Prata med din IT-support ifall du får problem med detta i datasalen.

Kör rad 14-20 för att ”ladda in” paketen. Detta måste göras varje gång du startar R-studio.

Se till att alla sökvägar stämmer med var du har sparat/vill spara data. Det gäller raderna 25, 26 och 32.

Skattning

Du kommer att skatta volym (VolumeTotal), grundyta (BasalArea), grundytvägd medelhöjd (HGV) och grundytvägd medeldiameter (DGV). 72 stycken lasermått kommer att användas vid skattningen av varje variabel. Som tidigare nämnts är en fördel med Random Forest att man kan slänga in väldigt många variabler av olika typer, utan att bry sig om eventuell korrelation mellan dem.

Börja med skattning av grundtyevägd medelhöjd, HGV.

1. Byt filnamn i sökvägen på `outputRaster` (rad 32) så att det är tydligt vilken variabel det gäller (nu står det HGV där). Gör detta inför varje skattad variabel. `outputRaster` är alltså namnet på den fil där du vill spara rastret med skattningsresultat.
2. Rad 65: Nu står redan HGV som beroende variabel (`...randomForest(HGV ~ ...)`), men när du ska skatta en annan variabel måste du ändra här.
3. Rad 111: Nu står det redan `myData$HGV` som `y`, men när du ska skatta en annan variabel måste du ändra här.
4. Första gången, alltså inför skattningen av HGV: **kör raderna 25, 26, 36 och 43** för att läsa in provyte- och rasterdata. Detta behöver bara göras en gång. **Kör också rad 38** för att se vilka lasermått som kommer att användas i skattningen (jämför med instruktioner till tidigare uppgifter och FUSION-manualen för att ta reda på vilka mått det rör sig om).
5. Inför varje ”varv”, alltså varje variabel som ska skattas: kör rad 32 för att definiera vad utdatafilen ska heta (efter att ha ändrat filnamnet till något lämpligt).

Nu kommer själva skattningen. Börja som sagt med HGV.

6. Läs igenom rad 47-63 om du inte redan gjort det. **Kör Random Forest-funktionen på rad 65-78** (du kan till exempel kopiera alltihop, klistra in i konsolen och trycka enter). Det kan ta en liten stund.

Nästa steg är att välja inställningar på `n`tree och `m`try. Det är i stort sett den möjlighet man har att påverka resultatet.

7. Läs igenom rad 80-94. **Kör rad 95** för att plotta skattningsfel mot antal träd. Kör om `randomForest` (rad 65-78) med så många träd som datorn orkar (prova dig fram), upp till ca 500 eller mer om möjligt. Inställningen för `n`tree ändras på rad 80. Plotta felet mot antal träd igen och undersök grafen. Bestäm dig för ett antal träd som är tillräckligt stort för att felet ska stabiliseras på en låg nivå. Ha gärna lite marginal också. Notera vilket `n`tree du valt och spara anteckningarna.
8. Läs igenom rad 97-109. **Kör raderna 110-111** en gång för att definiera `x` (beroende variabel, t ex HGV) och `y` (oberoende variabler, dvs lasermått). **Kör rad 112** för att plotta skattningsfel för olika värden på `m`try. Välj det `m`try som ger lägst fel, notera värdet och spara anteckningarna.
9. Ändra `n`tree och `m`try på rad 78 till de värden du valt ovan. Kör om `randomForest` (rad 65-78) med de inställningarna.

Sedan är det dags att undersöka modellen och resultatet.

10. Läs rad 118-127. **Kör rad 129** för att skriva ut en sammanfattning av ditt `randomForest`-objekt, och anteckna värdena på ”Mean of squared residuals” och ”% Var explained”.

Jämför med "Mean square error" och "R2 adjusted" eller R2 från din regression ifall du har sparat de resultaten.

Undersök vilka oberoende variabler (lasermått) som spelar störst roll för skattningen.

11. Läs igenom rad 131-154. **Kör rad 139, 140, 149 och 154** och undersök vilka lasermått som ger högst "IncMSE", dvs de mått som verkar viktigast för skattningsresultatet. Jämför de mått som hamnar överst i listan med de mått du valt i din regressionsmodell i uppgift 11. Ser det ut som du skulle förvänta dig, eller får du några överraskningar? Varför tror du att det ser ut så?

Använd din Random Forest-modell för skattning av hela Krycklan-området.

12. Läs igenom rad 156-163. **Kör rad 160** för att applicera modellen på hela området och få fram ett raster med skattade värden. **Kör rad 163** för att skriva resultatet till en TIF-fil. **Viktigt:** dubbelkolla på rad 32 att outputRaster är satt till rätt filnamn, så att du inte skriver över något tidigare resultat.

Gör nu om punkt 1-12 (punkt 4 behövs egentligen bara första gången, eller om du har startat om RStudio sedan sist) för en i taget av **volym och DGV**. För **grundyta** behöver du bara göra punkt 1-11 eftersom det inte finns utvärderingsdata för den.

Utvärdering

Använd dina raster med skattningsresultat och beräkna beståndsmedelvärden på samma sätt som i uppgift 11. Gör också en likadan utvärdering (rmse, relativ rmse, bias och relativ bias) som i uppgift 11.

Inlämning

För varje variabel (HGV, DGV, volym och grundyta), besvara fråga 1 och 2:

Fråga 1: Vilka värden valde du på ntree och mtry?

Fråga 2: Vilka variabler visade sig viktigast för skattningsresultatet (nämna de 5-10 första)? Hittar du bland dessa igen de variabler du använde i din egen regressionsmodell? Dök det upp några överraskningar? Varför tror du att resultatet ser ut som det gör?

För HGV, DGV och volym, besvara dessutom fråga 3 och 4:

Fråga 3: Jämför resultatet efter utvärdering på beståndsnivå (rmse, relativ rmse, bias och relativ bias) med dina resultat från uppgift 14. Blev det bättre/sämre? Hur mycket?

Fråga 4: Titta på rastren med skattningsresultat från regression och Random Forest. Använd *identify*-knappen i ArcMap. Titta på några pixlar med väldigt låga skattningsvärden och några med väldigt höga. Verkar någon av metoderna "dra mot mitten" mer än den andra?

Samla dina svar i ett word-dokument.

Lycka till!

Appendix: Översikt av filer och mappar till övningsdata

Servern geodata/Laserövning innehåller följande filer:

[\(\storage-ume.slu.se\common\\$\student\geodata\Laserövningar\)](http://storage-ume.slu.se/common$/student/geodata/Laserövningar)

Installera programvaror

- Installera Fusion.pdf
- Installera OSGeo4W & OrfeoToolbox.pdf
- Tekniska tips.pdf

Uppgift 2 Reserv_10C033_66350_5300_25.las
Reserv_10C033_66350_5300_25.rar

Uppgift 4 09P001_67275_5925_25_height_segm.tif
09P001_67275_5950_25_height_segm.tif
kombi_height_ortho.tif
orthophoto_12_epsg3006.tif

Uppgift 5 plotData.txt
readPlotData.R
Scatterplot.R

Uppgift 8 Krycklan_fieldplots.shp
 Krycklan_fieldplots.dbf
 Krycklan_fieldplots.prj
 Krycklan_fieldplots.sbn
 Krycklan_fieldplots.sbx
 Krycklan_fieldplots.shp
 Krycklan_fieldplots.shx
Krycklan_fieldplots.txt
KrycklanToPlots.las
list_clipped_plots.txt
Regression_intro.R

Uppgift 9 Photos 10st jpg-filer
poly2287
 poly2287.dbf
 poly2287.sbn
 poly2287.sbx
 poly2287.shp
 poly2287.shp.idx
 poly2287.shp.xml
 poly2287.shx
tree_ID1.shp
 tree_ID1.dbf
 tree_ID1.sbn
 tree_ID1.sbx
 tree_ID1.shp
 tree_ID1.shp.idx

tree_ID1.shp.xml
 tree_ID1.shx
 tree_ID5
 tree_ID5.dbf
 tree_ID5.sbn
 tree_ID5.sbx
 tree_ID5.shp
 tree_ID5.shp.idx
 tree_ID5.shp.xml
 tree_ID5.shx
 tree_ID7
 tree_ID7.dbf
 tree_ID7.sbn
 tree_ID7.sbx
 tree_ID7.shp
 tree_ID7.shp.idx
 tree_ID7.shp.xml
 tree_ID7.shx
 ytor_Buf
 ytor_Buf.dbf
 ytor_Buf.sbn
 ytor_Buf.sbx
 ytor_Buf.shp
 ytor_Buf.shp.idx
 ytor_Buf.shp.xml
 ytor_Buf.shx
 HeightImg_SLU_forest.bmp
 HeightImg_SLU_forest.bmpw
 SingleTree.R
 SLU_Forest.las
 TreeData.txt

Uppgift 10 Punktmoln_A.las
 Punktmoln_A.ldi
 Punktmoln_A.idx
 Punktmoln_B.las
 Punktmoln_B.ldi
 Punktmoln_B.idx
 Punktmoln_C.las
 Punktmoln_C.ldi
 Punktmoln_C.idx
 Punktmoln_D.las
 Punktmoln_D.ldi
 Punktmoln_D.idx
 RefImgHeight_ABC.bmp
 RefImgHeight_ABC.bmpw
 RefImgHeight_D.bmp
 RefImgHeight_D.bmpw
 TopographicMap.jpg

Uppgift 11 field_stands

field_stands.dbf
field_stands.prj
field_stands.sbn
field_stands.sbx
field_stands.shp
field_stands.shp.idx
field_stands.shp.xml
field_stands.shx

Raster_Laser

82st asc-filer
orthophoto_1209_epsg3006.tif
orthophoto_1209_epsg3006.tfw
orthophoto_1209_epsg3006.tif.ovr
Regression_DGV.R
topographicmap_12_epsg3006.tif

Uppgift 12 KrycklanPlots.txt

KrycklanRaster72.grd
KrycklanRaster72.gri
RF_skattning.r

Introduktion till Fusion.pdf

NIR-ljus och Intensitet.pdf

Regression i praktiken.pdf