

Skogliga skattningar med fjärranalys och skördardata

Forest assessment based on remote sensing and harvester data



FOTO: ERIK VIKLUND/MELLANSKOG

Summary

There is a constantly growing need for geographical data about the Swedish forest. A solid body of data enables forest owners to plan and follow up forestry measures, from both environmental and production perspectives.

In 2009-2014, Lantmäteriet, the Swedish government agency for mapping and land registration, carried out a national airborne laser survey. The Swedish Forest Agency and the Swedish University of Agricultural Sciences combined this laser data with reference data from the Swedish National Forest Inventory and produced digital maps that describe the forest in detail. The map service, called Skogliga grunddata ('Basic Forest Data'), is used by the entire forest sector to improve forestry planning, and serves as a basis for decisions concerning many different and new applications. However, since the laser survey has not been repeated, Skogliga grunddata is becoming out of date, and other data sources must be identified that could keep the database updated.

An alternative to airborne laser scanning is stereo matching of aerial images, which, like laser scanning, produces 3-D data on the tree layer, but with somewhat lower quality.

Harvester data has proved to be a very good source of reference information about the forest and, in combination with data from the airborne laser scanning, is also useful in forestry operations for predicting the yield from planned harvests. In this case, harvester data replaces reference data in the form of surveyed sample plots. This method should be a very cost-effective way of producing updated, comprehensive, grid-based assessments of the forest condition, because expensive surveys can then be replaced by sample plots.

This project focused on methods for producing grid-based assessments of the forest condition using harvester data in combination with stereo-matched aerial photos. The project used Lantmäteriet's new standard product Ytmodell från flygbilder (3-D data from stereo-matching of aerial photos), together with harvester data from the forest companies, Sveaskog and Södra, in a study area east of Växjö.

Two different assessment methods were developed and evaluated. One was the existing and proven imputation method currently used for laser scanning data, and the other was a first application of a new method, Deep Learning, from the research field Artificial Intelligence/Machine Learning. The second method is expected to be particularly suitable for large-scale surveying using remote sensing and harvester data, because it is designed to utilise vast quantities of data for automatic, complex assessment.

The results were encouraging. The combination of harvester data and stereo-matched aerial photos produced results that were only of slightly lower quality than the laser scanning data. Deep Learning also produced reasonable results, but noticeably lower quality than the well-established imputation method. However, there is great potential to improve the Deep Learning model evaluated in this study, because the data material used in the project was probably too small for efficient assessment, and also because more development work is needed to find a Deep Learning model that works well for this particular application.

Förord

Denna rapport sammanfattar ett utvecklingsprojekt finansierat av Skogssällskapet, som genomförts under åren 2016–2018. Projektet visar hur skördardata kan användas som markreferens för fjärranalysskattningar av skogliga parametrar som exempelvis grundyta och virkesvolym. En väsentlig utgångspunkt är tillgången till skördardata och den skördardatabas för forskningsändamål som Skogforsk byggt upp. Liknande databaser byggs nu inom respektive skogsföretag, vilket gör det möjligt att sprida tillämpningar som använder skördardata. Detta projekt har använt skördardata från Södra och Sveaskog, vilka företagen välvilligt ställt till förfogande för projektet. Sveaskog har även bidragit med utvärderingsdata i form av objektiva provtytor som samlades in för en företags-taxering. Stort tack till alla inblandade!

Uppsala 2018-10-31

Erik Willén, Jon Söderberg, Johan Möller, Skogforsk
Jörgen Wallerman, SLU

Innehåll

Summary.....	2
Förord	3
Sammanfattning	5
Bakgrund	6
Syfte och mål.....	7
Material och metod.....	7
Studieområde.....	7
Fjärranalysdata	8
Skördardata.....	8
Utvärderingsdata	9
Metod	10
Introduktion.....	10
Imputering	10
Deep Learning.....	11
Utvärdering.....	11
Resultat	13
Bidrag från olika indata	13
Imputering	17
Deep Learning.....	21
Diskussion och slutsatser	24
Referenser.....	25

Sammanfattning

Behovet av geografiska data för den svenska skogen ökar ständigt. Med ett bra data-underlag blir det lättare för skogsägare att planera och följa upp skogsbruksåtgärder, ur både ett miljö- och produktionsperspektiv.

År 2009–2014 genomförde Lantmäteriet en nationell flygburen laserskanning. Dessa laserdata har Skogsstyrelsen och Sveriges lantbruksuniversitet sambearbetat med referensdata från Riksskogstaxeringen till digitala kartor som detaljerat beskriver skogstillståndet. Karttjänsten, som kallas Skogliga grunddata, utnyttjas av hela skogssektorn för att förbättra den skogliga planeringen och fungerar som beslutsunderlag för en mängd olika tillämpningar som inte tidigare funnits. Men eftersom laserskanningen inte upprepats är nu Skogliga grunddata på väg att bli inaktuellt och det söks efter andra data-källor för att uppdatera databasen.

En alternativ ersättning för flygburen laserskanning är stereomatchning av flygbilder, som i likhet med laserskanning levererar tredimensionella (3D) data över trädsiktet, men med något lägre kvalitet.

Skördardata har sedan tidigare visat sig vara en mycket bra källa för referensinformation om skogen och är, i kombination med data från flygburen laserskanning, även operativt användbar för att prediktera utfallet av planerade avverkningar. Skördardata fungerar i detta fall som en ersättning av referensdata i form av inventerade provytor. Denna metod bör vara mycket kostnadseffektiv för att producera uppdaterade, heltäckande rasterskattningar av skogstillståndet, eftersom dyra inventeringar då kan ersättas av provytor.

Det här projektet har fokuserat på metoder för att producera rasterskattningar av skogstillståndet med hjälp av skördardata i kombination med stereomatchade flygbilder. Projektet har använt Lantmäteriets nya standardprodukt ”Ytmodeller från flygbilder” (3D-data från stereomatchning av flygbilder), tillsammans med skördardata från Sveaskog och Södra i ett studieområde öster om Växjö. Två olika skattningsmetoder har utvecklats och utvärderats. Den ena är den befintliga och beprövade imputeringsmetoden som idag används för laserskanningsdata och den andra är en första tillämpning av en ny metod, Deep Learning, från forskningsfältet Artificiell Intelligens/Machine Learning. Den senare metoden förväntas vara mycket väl lämpad för storskalig kartering med fjärranalys och skördardata, eftersom den är designad för att utnyttja enorma datamängder för automatiska, komplexa skattningar.

Projektet visar goda resultat. Kombinationen av skördardata och stereomatchade flygbilder producerade bara något sämre resultat än laserskanningsdata. Deep Learning levererade också rimliga resultat, men klart sämre än den etablerade imputeringsmetoden. Dock finns en mycket stor förbättringspotential för Deep Learning-modellen som utvärderades här. Det beror dels på att projektets datamaterial sannolikt är för litet för en effektiv beräkning och dels på att en betydligt större utvecklingsinsats behövs för att hitta en Deep Learning-modell som fungerar bra för just denna tillämpning.

Bakgrund

Under 2015 avslutade Skogsstyrelsen arbetet med Skogliga Grunddata – en nationell kartering av skogliga data (parametrarna trädhöjd, grunddyta, medeldiameter och virkesvolym) baserat på Lantmäteriets laserdata och Riksskogstaxeringens inventeringsdata.

För att göra fjärranalysskattningar av skogliga parametrar krävs referensdata från fält. Genom att använda skördardata som fältreferens blir kostnaderna betydligt lägre än för fältinventeringar. Standardiserade skördardata samlas in kontinuerligt och kan sedan användas i etablerade metoder för värdeskapande skattning av skogstillståndet från fjärranalys. Med tillgång till heltäckande, aktuella och noggranna data över skogen kan planeringen av skogsbruket förbättras avsevärt. Samtidigt kan produktionskostnaderna för nya skogsbruksplaner sänkas, då en stor del av mätningarna i fält kan ersättas och systematiska fel från subjektiva inventeringsmetoder undviks. Detta bidrar till att mer tid kan användas i fält till att bestämma och korrigera de egenskaper som inte enkelt kan skattas med fjärranalys, såsom avdelningsgränser, trädslagsblandning och skötselalternativ.

Lantmäteriets laserskanning utfördes i huvudsak under perioden 2009–2014, vilket innebär att data för stora delar av landet börjar bli inaktuella, framförallt i södra Sverige. Den nya laserbaserade markmodellen har gett helt nya möjligheter att använda flygbilder för att producera likartad information om skogens tillstånd som laserskanning kan ge, det vill säga tre-dimensionella (3D) data med noggrann information om framförallt trädhöjd. Flygbilderna i kombination med markmodellen kan ge tillförlitliga trädhöjder. Med stereomatchning av Lantmäteriets flygbilder och markmodellen har forskning vid SLU visat skattningsnoggrannheter av skogliga data ungefär motsvarande Skogliga grunddata. Metoden är på stark frammarsch och Lantmäteriet har under 2015 gjort en pilotproduktion, som idag distribuerar trädhöjdsraster från stereomatchning av flygbilderna från 2015 och därefter.

Skördardata samlas in kontinuerligt för i stort sett alla avverkningar i Sverige. Dessa data lagras enligt StanForD-standarden (Arlinger m.fl. 2012), vilket möjliggör att de kan nyttjas i en rad olika tillämpningar. De senaste åren har gallringsuppföljning (Hannrup m.fl. 2015, Möller m.fl. 2011) med skördardata och olika verktyg för produktionsuppföljning implementerats och mer utveckling pågår, bland annat för användning som referensdata för skattningar av virkesutfall enligt hprYield-metoden (pers ref Bhuiyan 2015). Möjligheten finns även att beskriva den kvarvarande skogen efter gallring. I ett samarbete mellan SLU och Skogforsk har en kombination av skördardata och laserdata för att kartera skogliga parametrar visat mycket goda resultat (Holmgren m.fl. 2012; Söderberg 2015). Skogforsk har utvecklat nya verktyg för att enkelt sammanställa skördardatafiler. Dessa verktyg gör det praktiskt möjligt att hantera data från den egna verksamheten samt geografiskt koppla till fjärranalysdata och producera tematiska kartor.

Syfte och mål

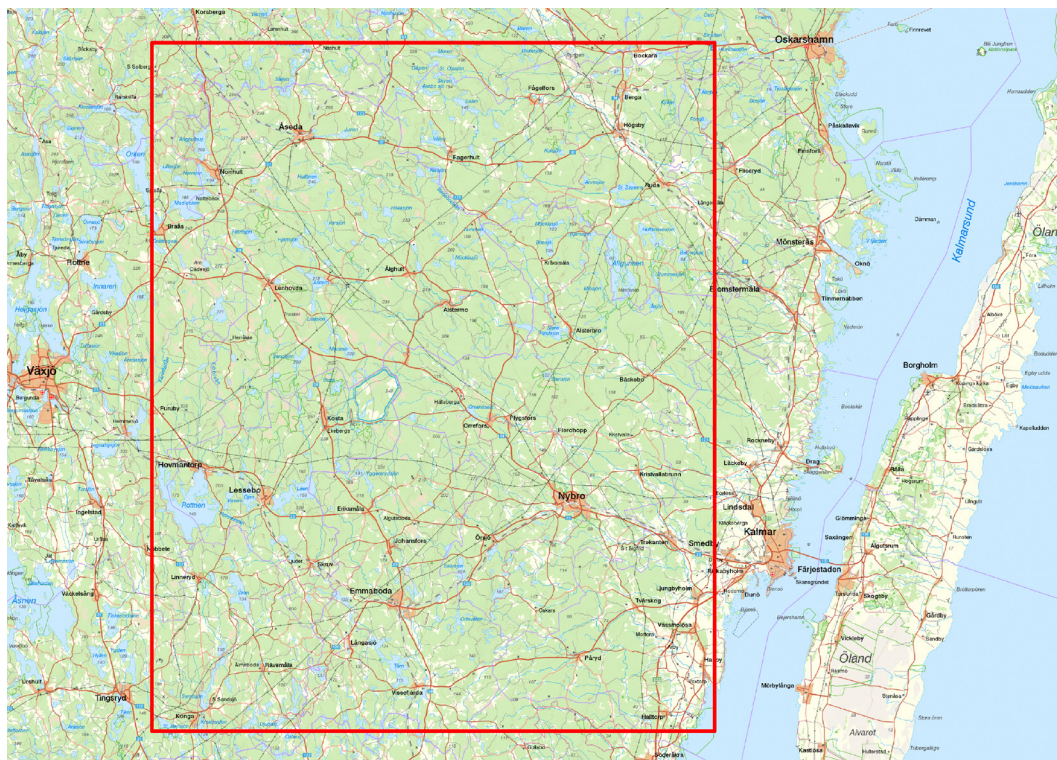
Syftet med projektet var att förbättra dataförsörjningen till skogsbruket genom att utnyttja möjligheterna att använda befintliga och noggranna datakällor till kostnads-effektiv fjärranalysbaserad kartering av skogliga data. För att fånga den skogliga variationen bättre än med beståndsmedelvärden blir karteringen yttäckande. Speciellt intresse riktas mot små och medelstora skogsägare som ännu inte kunnat motivera kostnaderna att använda modern fjärranalys i sin verksamhet. Projektets resultat förväntas generera nya möjligheter för värdeskapande planering av kundkommunikation, avverkning, skogsskötsel och naturvård.

Projektets mål var att ta fram en mycket kostnadseffektiv metod att producera aktuella rasterskattningar av skogliga data baserad på den löpande insamlingen av flygbilder och skördardata. Metoden utvärderades tillsammans med skogsföretag i Götaland mot nyligen gjorda kontrolltaxerade beståndsdata samt objektiva skördarmätningar på beståndsnivå. De skogliga parametrar som karterades var grundyta, grundytavägd medelhöjd, grundytavägd medeldiameter och stamvolym per hektar.

Material och metod

STUDIEOMRÅDE

Studieområdet var beläget öster om Växjö (Figur 1) med en area om drygt 6000 km². Denna yta bedömdes inkludera en tillräckligt stor mängd av skördardata för att fånga den skogliga variationen. Området valdes även då fjärranalysdata i form av flygbilder var registrerade samma år över hela området.



Figur 1. Studieområde (©Lantmäteriet).

FJÄRRANALYSDATA

Studieområdet omfattades av Lantmäteriets ordinarie omdrevsfotografering som genomfördes 2016-04-11 till 2016-04-12 med mätkameran Ultracam Eagle och levererade IR-flygbilder med en pixelstorlek av 0,24 m. Lantmäteriet genomför idag en löpande stereomatchning av alla nya flygbilder till en ny produkt, ”Ytmodell från flygbild” (Lantmäteriet, 2018), vilken användes för detta projekt. Produkten innehåller en höjd över havet för varje punkt i ett regelbundet grid med 0,5 m avstånd mellan närliggande punkter och är beräknade med semi-global stereomatchning av flygbilderna i programvaran SURE (Lantmäteriet, 2018; Rothmel m.fl., 2012).

Lantmäteriets ordinarie markhöjdsmodell användes för att beräkna ett flertal variabler som bör beskriva ståndortsegenskaper som exponering och markfuktighet. Dessa variabler är idag enkla att beräkna med nationell täckning och har sannolikt potential att förbättra skattningsnoggrannheten för storskalig kartering baserad på fjärranalys och så detaljerade data som skördarsystemen levererar. Två olika mått som beskriver markfuktigheten beräknades från Lantmäteriets laserskannade markmodell: depth-to-water (Leach m.fl., 2017), topographic wetness index (Sørensen m.fl., 2006) och en klassificering av markens lutningsförhållanden relaterat till dess omgivande landskap: Pennocks klassificering (Pennock m.fl., 1987). För den senare beräknades lutningsklasser för fyra olika spatiala upplösningar: 50, 30, 20 och 10 m.

Laserdata samt skogliga grunddata från Skogsstyrelsen nyttjades också, vilka baserades på skanningar över området från 2012.

SKÖRDARDATA

Södra och Sveaskog tillhandahöll skördardata från avverkningsuppdrag inom studieområdet, som utfördes under en period från hösten 2015 fram till våren 2017. Den beräknade arealen för de insamlade åtgärderna blev 6226 ha, varav 1857 ha kom från slutavverkningar (Tabell 1) och 4369 ha från gallringar (Tabell 2).

För att kunna använda skördardata från gallringsåtgärder behövs en prognos av kvarvarande bestånd efter utförd åtgärd. Skogforsk har utvecklat ett verktyg som kan göra detta inom ett projekt för gallringsuppföljning. I materialet ingår områden som gallrats före flygfotograferingen (11–12 april 2016), där endast prognosen av kvarvarande bestånd använts för att beskriva skogen. Även områden som gallrats efter datumen för flygfotona har använts. Där läggs prognosen för kvarvarande bestånd ihop med en skoglig beskrivning av uttaget.

Tabell 1. Skördardata från slutavverkning, 1716 beräkningsytor.

	Volym m³sk/ha	Grundyta m²/ha	HGV (m)	DGV (mm)	Antal stammar	Tall	Gran	Löv
medel	306,2	29,4	23,5	306,9	555	0,27	0,68	0,05
min	104,5	13	14,9	183	211	0	0	0
max	773,7	61,1	31,9	457	1605	1	1	0,98
SD	102,5	7,8	2,8	38,3	162	0,25	0,25	0,1

Tabell 2. Skördardata från gallring, totalt 4923 beräkningsytor.

	Volym m³sk/ha	Grundyta m²/ha	HGV (m)	DGV (mm)	Antal stammar	Tall	Gran	Löv
medel	142,3	18,7	16,3	198,5	869	0,46	0,41	0,13
min	44	12,1	9,4	116	266	0	0	0
max	327,1	29,4	25,5	370	1862	1	1	0,98
SD	42,5	2,7	2,8	39,7	281	0,34	0,29	0,17

UTVÄRDERINGSDATA

Som utvärderingsdata användes den företagstaxering som Sveaskog genomförde i studieområdet 2017. Den bestod av 840 provytor insamlade i 89 olika bestånd, 6–10 provytor per bestånd. Datainsamlingen gjordes enligt IPAK-modellen. Provytedata bearbetades i Heureka och sammanställdes beståndsvis för utvärdering mot skattningarna (Tabell 3).

Tabell 3. Egenskaper för 89 utvärderingsbestånd inventerade av Sveaskog.

	Volym m³sk/ha	Grundyta m²/ha	HGV (m)	DGV (mm)
medel	167,3	19,89	14,71	18,59
min	0	0	0	0
max	823,9	76,94	25,17	33,57
SD	123,6	11,6	6,6	8,81

Metod

INTRODUKTION

Den främsta utmaningen inom projektet gäller användningen av skördardata som fält-referens istället för fältmätta provytor. Metodutvecklingen fokuserade på hur dessa nya indata kan nyttjas och ge svar på frågor som; Hur mycket skördardata som behövs, storlek på beräkningsytor, aktualitet på skördardata, geografiskt avstånd mellan referensdata och skattning samt nyttan med att använda kompletterande indata (stöd till skördardata) baserat på markmodellen.

Skördardata som samlas in från tiden direkt efter flygbilden är registrerade för slutavverkningar. För gallringar, där den kvarstående skogen beskrivs, tilläts även skördardata från den tidigare vegetationssäsongen och fram till registreringstidpunkten för flygbilderna. Sedan delas utförda avverkningar in i små (0,1–1 hektar) beräkningsytor baserat på den skogliga beskrivningen i skördardata.

Den nya markmodellen används tillsammans med flygbilder för att producera motsvarande information om skogens tillstånd som med laserskanning, det vill säga 3D-data med noggrann mätning av framförallt trädhöjd. 3D-data beräknas genom stereomatchning av nationellt tillgängliga flygbilder, exponerade innan avverkning, med fotogrammetri-programmet SURE.

Karteringen av skogliga data från flygbildernas 3D-information genomfördes med area-metoden (Næsset m.fl. 2004), som är en etablerad standard (även kallad imputering). Även en annan metod, Deep Learning, testades. Bägge metoderna redovisas separat nedan. Metoderna innebär i korthet att sambandet mellan flygbildernas 3D-data och skördarmätningarna av skogliga parametrar modelleras matematiskt med hjälp av data från beräkningsytorna. Dessa modeller används därefter för att prediktera de skogliga målparametrarna i alla områden med flygbildstäckning. Resultatet blir yttäckande skattningar av de skogliga parametrarna.

De framtagna skattningarna analyserades och utvärderades mot olika oberoende material. Med hjälp av skördaravverkade objekt kan noggrannheten i skattningarna utvärderas, dels generellt och dels avseende skillnader mellan olika skogstyper. Den objektiva företagstaxering som Sveaskog genomförde var ett lämpligt utvärderingsmaterial över området.

IMPUTERING

Vi har valt att använda termen ”imputering” för tekniken att söka ett antal närbelägna avverkade beräkningsytor med liknande fjärranalysmetriker och överföra en skoglig beskrivning från dessa. Metoden för imputering som användes kallas kMSN och är en utveckling av Moeur & Stages (1995) ”Most Similar Neighbour”-metod, där ett antal (k) av de mest liknande referensytorna (grannarna) väljs ut för varje yta som ska skattas. Medelvärdena av skattnings-variablerna från dessa ”grannar” överförs till prognosytan. Imputering kräver dock en omfattande referensdatabas. Detta eftersom metoden inte kan extrapolera skattningar, utan endast väljer ur det tillgängliga referensmaterialet och beräknar ett viktat medelvärde av skogliga parametrar från detta urval.

För att skatta skogliga parametrar för segmenten användes kodbiblioteket "yaImpute" (Crookston & Finley 2007) för programpaketet "R" (Anon. 2016) som är fritt tillgängligt. "yaImpute" utför imputering på multivariata data, vilket innebär att hela den skogliga beskrivningen kan skattas med en och samma funktion.

DEEP LEARNING

Det finns idag en mångfald av komplexa problem som har lösts med hjälp av självlärande algoritmer utvecklade inom forskningsfältet Machine Learning (ML) och artificiell intelligens (Goodfellow m.fl., 2016). Inom ML har speciellt nya metoder baserade på neurala nätverk fungerat mycket väl för att med hjälp av stora mängder data modellera mycket svårstuderade beroenden i mätningar. Neurala nätverk är inget nytt forskningsfält. Det har snarare fått en renässans tack vare dagens kraftfulla processkapacitet som gjort det möjligt att använda djupa nätverk. Detta är grunden i deep learning, som är den snabbast växande ML-metoden. Google har släppt källkod för metoden, Tensor Flow (Abadi m.fl., 2015), som är tillgänglig för alla. En grund i ML-konceptet är att utnyttja stora mängder data för att automatiskt "träna" en algoritm att bli en effektiv prediktor. Algoritmen blir alltså bättre och bättre allt eftersom den utsätts för, bearbetar och analyserar ny information. På så sätt blir den inte låst till förhållandevis enkla matematiska modeller som inte säkert beskriver multivariata samband som är svåra att upptäcka manuellt.

Geografisk kartering av stora skogsområden baserat på enkelt tillgängliga fjärranalysdata och skördardata från avverkningar som registreras rutinmässigt bör vara en tillämpning där ML-metoder är effektiva alternativ till dagens inventerings- och modellbaserade karteringsmetoder. I detta projekt användes ett feedforward-neuralt nätverk implementerat i Tensor Flow (via KerasLab) för en första utvärdering av metoden. Flera olika nätverk utvärderades, där vart och ett simultant skattade DGV, HGV, grundyta och volym från krontakshöjdsdata från de stereomatchade flygbilderna (för detta användes medelvärde och standardavvikelse för den 20:e, 60:e och 95:e höjdpercentilen inom beräkningsytan, samt även medelvärde och standardavvikelse för typvärdet av höjden).

UTVÄRDERING

De framtagna modellerna utvärderades dels med korsvalidering och dels med hjälp av ett oberoende datamaterial (kontrolltaxerade bestånd från Sveaskog). Genom att beräkna standardavvikelse (ekvation 1 & 2) erhöles mått på den genomsnittliga variationen i skattningarnas avvikelse från uppmätta värden. För att jämföra precisionen för olika varianter på skattningar användes absoluta och relativa medelfel (ekvation 3 & 4). Systematiska avvikelser för skattningar identifierades och kvantifierades i absoluta och relativa mått med hjälp av ekvation 5 & 6.

$$SD = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}} \quad [1]$$

$$SD\% = 100 * \frac{SD}{\bar{y}} \quad [2]$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (x_i - y_i)^2}{n}} \quad [3]$$

$$RMSE\% = 100 * \frac{RMSE}{\bar{y}} \quad [4]$$

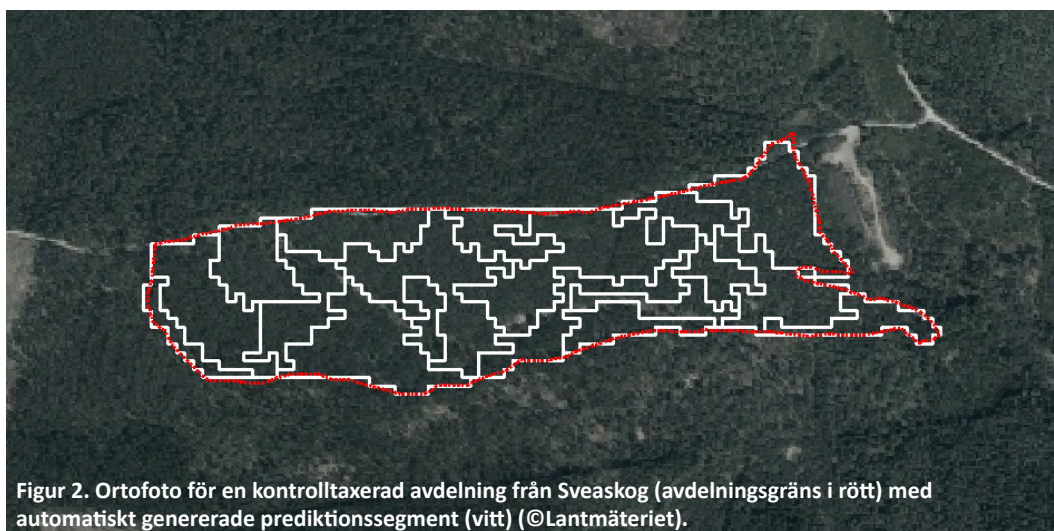
$$Bias = \frac{\sum_{i=1}^n (x_i - y_i)}{n} \quad [5]$$

$$Bias\% = 100 * \frac{Bias}{\bar{y}} \quad [6]$$

I ekvationerna representerar n antalet prognosytor, x_i det skattade värdet för prognosyta i , y_i det observerade värdet för prognosyta i , \bar{x} medelvärdet av de skattade värdena och \bar{y} medelvärdet av de observerade värdena.

Korsvalidering genomförs med samma datamaterial som använts för att anpassa skattningsmodellerna och ger en uppskattning av modellernas noggrannhet utan ytterligare datamaterial. Den utförs genom att en (eller flera) observationer undantas från datamaterialet. Modellerna anpassas sedan till det resterande datamaterialet och de anpassade modellerna används för att prediktera målparametrarna på de undantagna observationerna. Detta upprepas tills alla observationer har tilldelats predikterade värden. Därefter kan modellernas skattningsnoggrannheter slutligen beräknas genom att jämföra verkliga med predikterade värden för alla observationer.

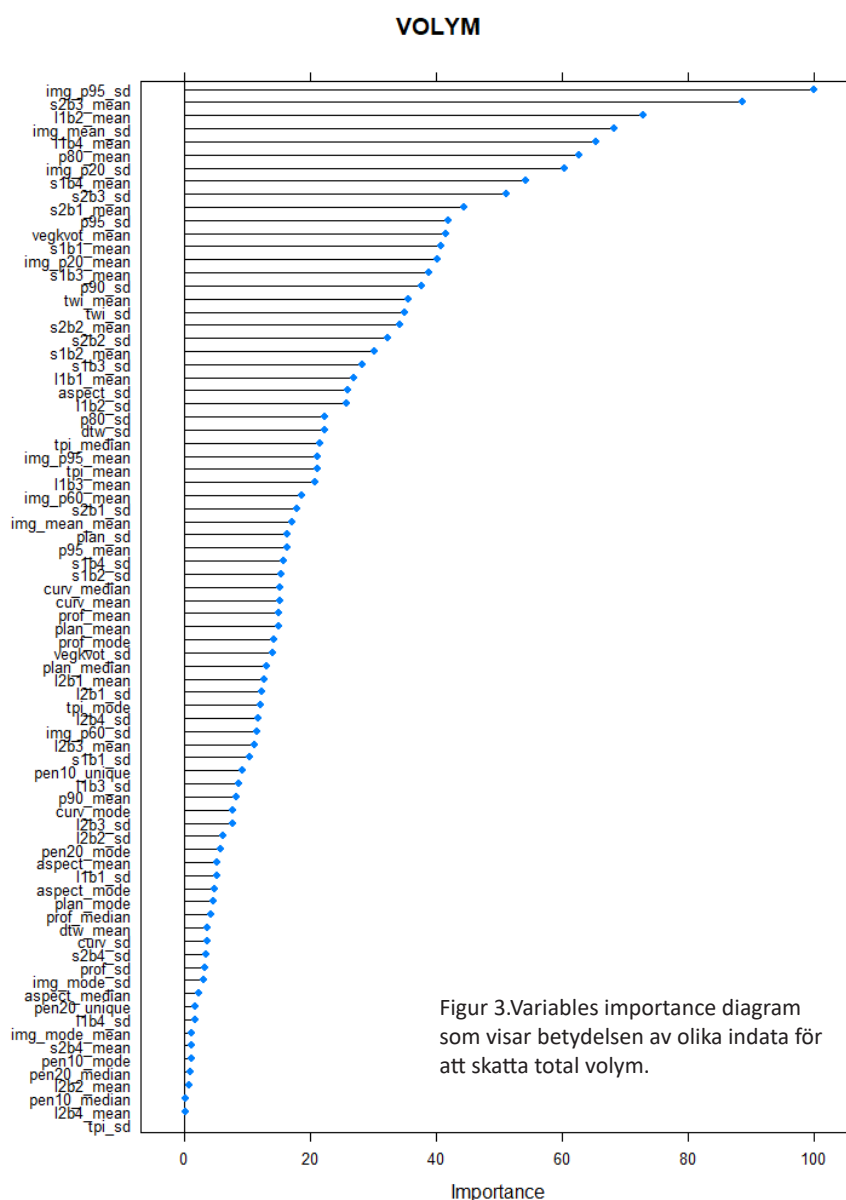
Utvärdering med ett oberoende datamaterial genomfördes genom att göra skattningar för segmenterade delar i 89 taxerade bestånd (Figur 2) inom försöksområdet och jämföra en arealviktad sammanställning av skattade värden för segmenten i varje bestånd med beståndsmedelvärden från taxering.



Resultat

BIDRAG FRÅN OLIKA INDATA

För att bedöma vilka indata som bidrar mest till skattningarna skapades regressionsmodeller för volym, grundyta, medelhöjd och medeldiameter. "Variable importance", som beskriver prediktorvariablernas statistiska signifikans för modellen, beräknades för varje regressionsmodell. Resultaten presenteras som i figur 3 med en sjunkande betydelse för skattningen. Det är ofta en eller ett par variabler som bidrar mest. I figur 3 bidrar ett höjdmått (p95) från fotogrammetriskattningen mest, följt av satellitdata. Många av de insamlade variablerna samvarierar, exempelvis de olika trädhöjdsmått, och bidrar då inte lika mycket till förklaringsgraden i modellerna. En sammanfattning över de tio mest värdefulla indata för olika skogstyper redovisas i tabell 4–7. Här skiljer sig slutavverkningsskog från gallringsskog genom vilka skördardata som användes som referens. Generellt sett var det förvånande hur mycket det äldre laserdata fortfarande kan förklara.



Figur 3. Variables importance diagram som visar betydelsen av olika indata för att skatta total volym.

För volymsskattningar, tabell 4, bidrar olika höjdmått från fotogrammetri samt laserdata mest. För gallringsskog är inte fotogrammetrimätningarna lika betydelsefulla. Detta kan förklaras med att de inte lika bra fångar skogens täthet. För slutavverkning noteras att topografimåtten *profile* och *curvature* ger ett betydande bidrag.

Tabell 4 Betydelsen av indata för volymsskattning.

Rangordning	Totalt	Gallringsskog	Slutavverkning
1	Fotogrammetri höjd (p95) standardavvikelse	Laserdata vegetationskvot medelvärde	Fotogrammetri höjd (p95) standardavvikelse
2	Fotogrammetri medelhöjd, standardavvikelse	Satellitdata s2(b1) medelvärde	Satellitdata s (b4) medelvärde
3	Fotogrammetri Höjd (p20), standardavvikelse	Satellitdata s2(b3) medelvärde	Profile, standardavvikelse
4	Laserdata vegetationskvot standardavvikelse	Satellitdata s1(b1) medelvärde	Laserdata vegetationskvot medelvärde
5	Fotogrammetri Höjd (p20), medelvärde	Laserdata percentil (80)	Satellitdata tidpunkt l1 (b4) medelvärde
6	Laserdata höjd (p80) medelvärde	Satellitdata l1b2 medelvärde	Fotogrammetri Höjd (p20), standardavvikelse
7	Laserdata höjd (p95) medelvärde	Fotogrammetri höjd (p95) standardavvikelse	S1b2 satellitdata medelvärde
8	Fotogrammetri höjd medelvärderna från olika höjder	Satellitdata s1(b3) medelvärde	Curvature standardavvikelse
9	TWI, medelvärde	Satellitdata l1(b4) medelvärde	Satellitdata band 2 medelvärde
10	Laserdata höjd (p95) standardavvikelse	Fotogrammetri höjd (p95) medelvärde	Fotogrammetri medelvärde standardavvikelse

Vid skattningen av grunddyta, tabell 5, är det många olika typer av indata som bidrar och på lite olika sätt. Höjd, satellitdata samt vegetationskvot är viktiga. Satellitdata ger ett betydande bidrag som sannolikt kan hjälpa till att förklara vissa trädslagsskillnader.

Tabell 5. Betydelse av olika indata för skattning av grunddyta.

Rangordning	Totalt	Gallringsskog	Slutavverkning
1	Fotogrammetri höjd (p95) standardavvikelse	Satellitdata s2b1 medelvärde	Laserdata vegetationskvot medelvärde
2	Fotogrammetri medelhöjd, standardavvikelse	Fotogrammetri höjd (p95) standardavvikelse	Satellitdata s 1(b4) medelvärde
3	Satellitdata l1b2 medelvärde	Satellitdata s1(b2) medelvärde	Fotogrammetri höjd (p95) standardavvikelse
4	Satellitdata l1b4 medelvärde	Fotogrammetri medelhöjd, standardavvikelse	Profile standardavvikelse
5	Satellitdata s1b4 medelvärde	Fotogrammetri Höjd (20) standardavvikelse	Satellitdata tidpunkt l1 (b4) medelvärde
6	Laserdata vegetationskvot standardavvikelse	Laserdata vegetationskvot medelvärde	Curvature standardavvikelse
7	Laserdata höjd (p95) medelvärde	Satellitdata 11(b1) medelvärde	Satellitdata S1b2 medelvärde
8	Fotogrammetri Höjd (20) medelvärde	Satellitdata s1(b1) medelvärde	Aspect (från höjddata) standardavvikelse
9	Laserdata (p95) standardavvikelse	Satellitdata s2(b3) medelvärde	Fotogrammetri Höjd (20) standardavvikelse
10	Satellitdata s2b3 standardavvikelse	Laserdata höjd (p80) medelvärde	TWI standardavvikelse

För den grundytevägda medelhöjden, tabell 6, ser den mer noggranna lasermätningen fortfarande ut att kunna bidra, även om höjderna i sig inte stämmer. Den nya höjden underskattas. Även här är det möjligt att satellitdata hjälper med förklaringen mellan olika trädslag. Trädhöjder mäts bra även från fotogrammetri, men sannolikt används de mer noggranna lasermätningarna i modellerna och då viktas betydelsen av nyare data från fotogrammetriskattningar ner.

Tabell 6. Betydelse av olika indata för skattning av hgv.

Rangordning	Totalt	Gallringsskog	Slutavverkning
1	Laserdata vegetationskvot medelvärde	Laserdata vegetationskvot medelvärde	Satellitdata s2b3 medelvärde
2	Satellitdata s2b3 medelvärde	Laserdata (p80) medelvärde	Laserdata vegetationskvot medelvärde
3	Laserdata (p80) medelvärde	Satellitdata s1(b1) medelvärde	Satellitdata l1b2 medelvärde
4	Fotogrammetri (p95) standardavvikelse	Satellitdata s2(b1) medelvärde	Laserdata vegetationskvot medelvärde
5	Satellitdata s2b2 medelvärde	Satellitdata s2(b3) medelvärde	Satellitdata tidpunkt s2 (b3) standardavvikelse
6	Satellitdata s2b1 medelvärde	Laserdata (p95) medelvärde	Laserdata (p95) medelvärde
7	Satellitdata s1b1 medelvärde	Laserdata (p90) medelvärde	Satellitdata s1b2 medelvärde
8	Fotogrammetri medelvärde standardavvikelse	Fotogrammetri (p95) standardavvikelse	Laserdata (p80) standardavvikelse
9	Satellitdata s2b3 medelvärde	Satellitdata l1(b4) medelvärde	Satellitdata s2b1 medelvärde
10	Satellitdata s2b2 medelvärde	Satellitdata s1(b3) standardavvikelse	Satellitdata s2b2 standardavvikelse

Sambanden mellan höjd och grundtevägd medeldiameter (dgv) är tydliga, vilket även märks på så sätt att det är liknande indata som bidrar mest till skattningen, tabell 7.

Tabell 7. Betydelse av olika indata för skattning av dgv.

Rangordning	Totalt	Gallringsskog	Slutavverkning
1	Laserdata vegetationskvot medelvärde	Laserdata vegetationskvot medelvärde	Satellitdata s2b3 medelvärde
2	Satellitdata s2b3 medelvärde	Laserdata (p80) medelvärde	Satellitdata s2b2 medelvärde
3	Fotogrammetri Höjd (p95), standardavvikelse	Satellitdata s1b1 medelvärde	Satellitdata s2b3 standardavvikelse
4	Fotogrammetri medelvärde standardavvikelse	Fotogrammetri (p95) medelvärde	Laserdata (p95) medelvärde
5	Satellitdata s2b2 standardavvikelse	Fotogrammetri medelvärde standardavvikelse	Laserdata vegetationskvot medelvärde
6	Satellitdata s2b2 medelvärde	Satellitdata s2b2 standardavvikelse	Laserdata (p90) medelvärde
7	Fotogrammetri (p20) standardavvikelse	Satellitdata l1b1 medelvärde	Satellitdata S2b2 standardavvikelse
8	Satellitdata s2b3 standardavvikelse	Fotogrammetri (p20) standardavvikelse	Satellitdata s1b2 standardavvikelse
9	Satellitdata s1b1 medelvärde	Satellitdata s1(b4) standardavvikelse	Satellitdata s2b4 medelvärde
10	Satellitdata s1b3 medelvärde	Pennock 20	Fotogrammetri (p95) medelvärde

IMPUTERING

Under arbetet med att ta fram en modell för imputering av skogliga parametrar utfördes korsvalidering av olika modeller för jämförelse. Skattningar gjordes för över 6000 enskilda beräkningsytor och resultaten utvärderades mot den skogliga beskrivningen av beräkningsytan med ekvation 1–6. Utvärderingen av en modell som inkluderade både laser- och fotogrammetrimetriker samt satellit- och geodata presenteras i tabell 8.

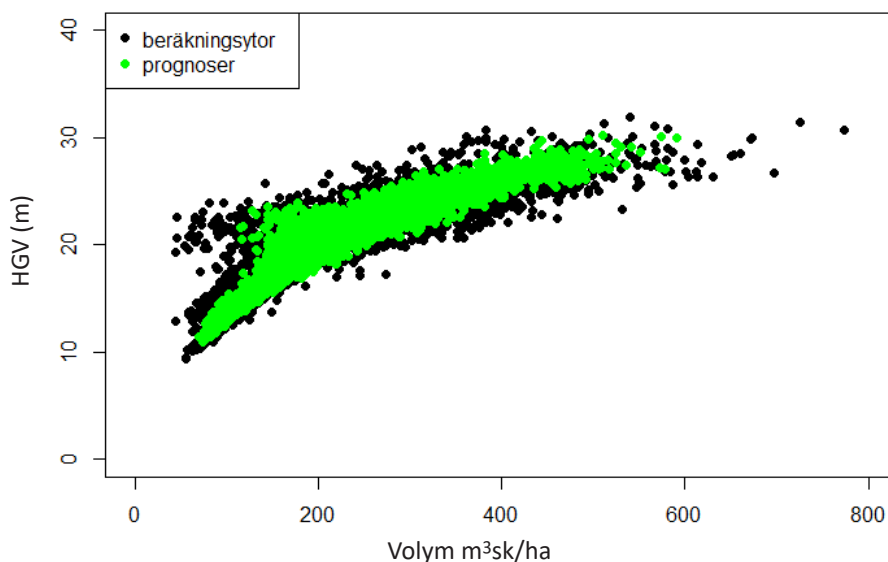
Sambanden mellan volym och medelhöjd (Figur 4) samt volym och grunddyta (Figur 5) fångas bra med imputering, men fröträdsställningar är svårskattade (uppe till vänster i Figur 4). Genom att använda medelvärdesbildning fångas dock inte hela variationen i skogen.

För att skatta utvärderingsbestånden gjordes separata prognoser för delar (segment) av utvärderingsbestånden. Förhållandet mellan volym och grundytavägd medelhöjd (Hgv) kan ses i Figur 6, medan Figur 7 illustrerar sambanden mellan volym och grunddyta.

Tabell 8. Resultat för korsvalidering av skattningar för enskilda beräkningsytor.

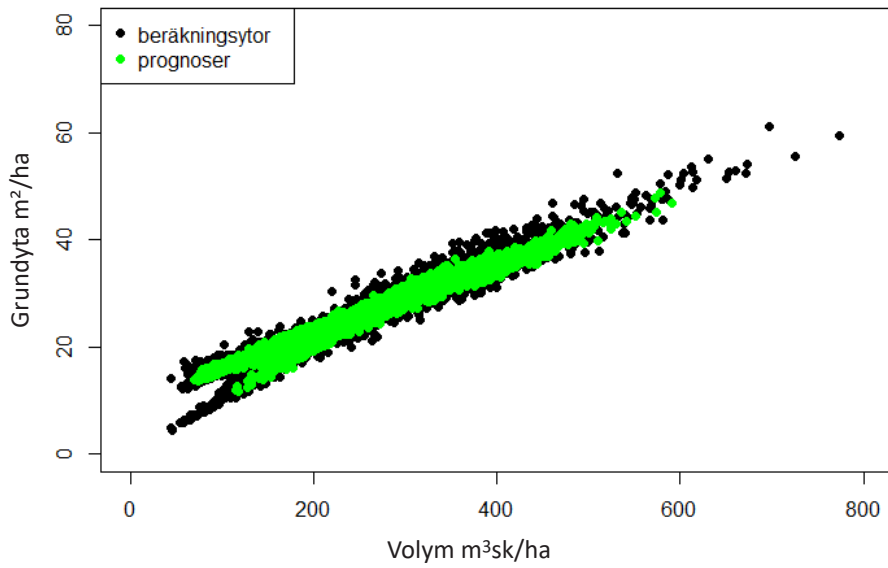
	Enhet	RMSE	RMSE%	Bias%
Volym	m ³ sk/ha	35,5	19,2	-0,3
Hgv	m	1,2	6,6	-0,1
Dgv	cm	2,5	11,2	-0,1
Grunddyta	m ² /ha	3,1	14,6	-0,1

Samband mellan Volym & HGV



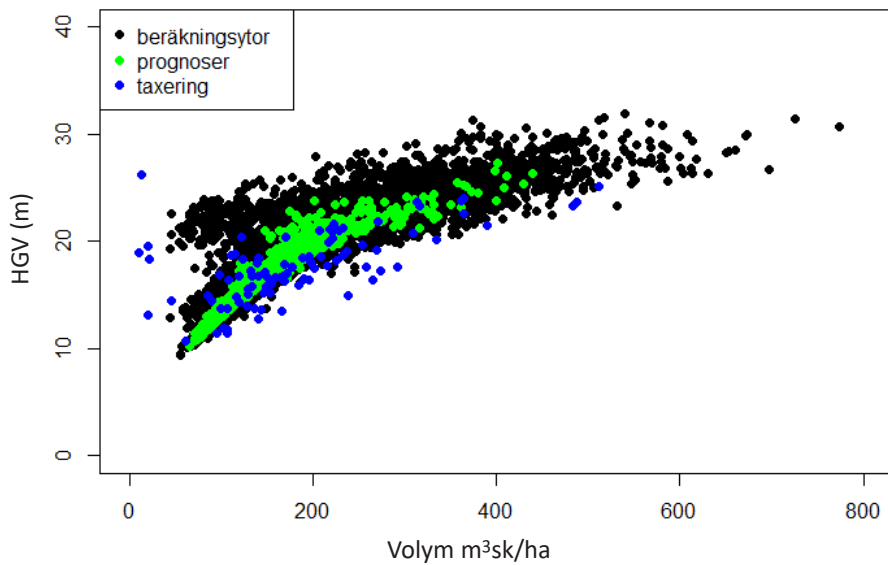
Figur 4. Samband mellan volym och grundytavägd medelhöjd (Hgv) för beräkningsytor (svart) och prognoser för enskilda beräkningsytor baserade på medelvärdesbildning av 5 beräkningsytor (grön).

Samband mellan Volym & Grundyta



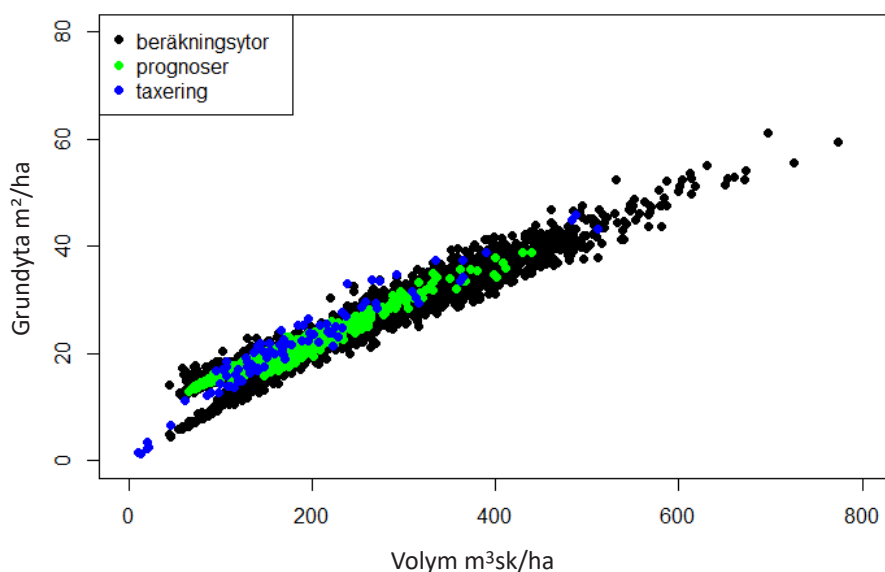
Figur 5. Samband mellan volym och grundyta för beräkningsytor (svart), skattningar för beräkningsytor (grön).

Samband mellan Volym & HGV



Figur 6. Samband mellan volym och Hgv för beräkningsytor (svart), imputerad prognos för delar (segment) av utvärderingsbestånd (grön) samt taxeringsdata från utvärderingsbestånd (blå).

Samband mellan Volym & Grundyta



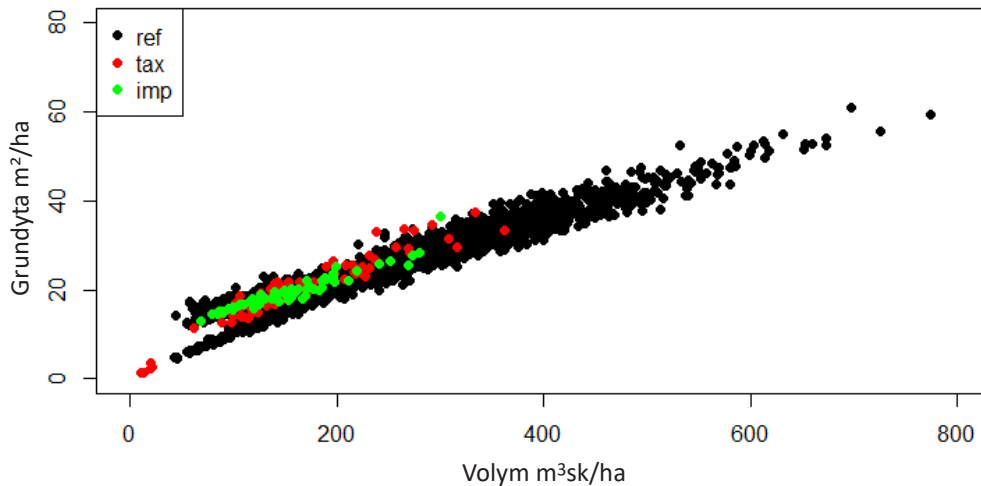
Figur 7. Samband mellan volym och grundyta för beräkningsytor (svart), skattningar för delsegment av utvärderingsbestånd (grön) samt taxeringsdata från utvärderingsbestånd (blå).

Utvärdering mot oberoende referensdata visar att skattningar av grundyta och volym varierar betydligt mer än testerna på beräkningsytor ovan, medan hgv och dgv skattas med liknande precision som tidigare (tabell 9, figur 8 och 9).

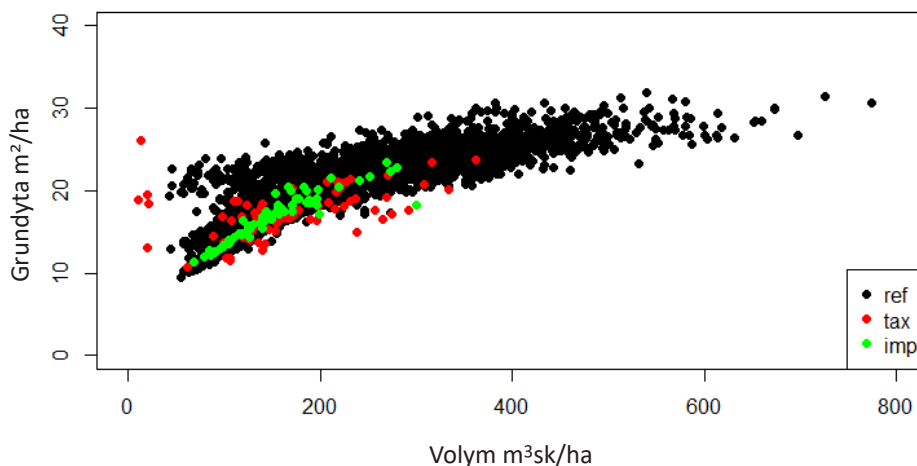
Fröträdsställningar är svåra att skatta, då de gamla laserdata är missvisande. De beskriver tillståndet innan avverkning och lämnande av fröträden. Även de nya flygfotometrierna har svårt att beskriva så glesa ytor.

Tabell 9. Avvikelse för skattningar på utvärderingsbestånd, respektive segment/delar av utvärderingsbestånd i förhållande till taxerade beståndsuppgifter.

Variabel	Enhet	Segment			Bestånd		
		RMSE	RMSE (%)	Bias (%)	RMSE	RMSE (%)	Bias (%)
Volym	m ³ sk/ha	74,27	44,17	-8,90	69,61	41,40	-8,46
Grundyta	m ² /ha	8,13	39,55	-5,31	7,60	36,95	-5,93
Hgv	m	1,30	7,49	-2,83	1,40	8,09	-1,94
Dgv	cm	3,92	17,05	-7,79	3,97	17,27	-7,60

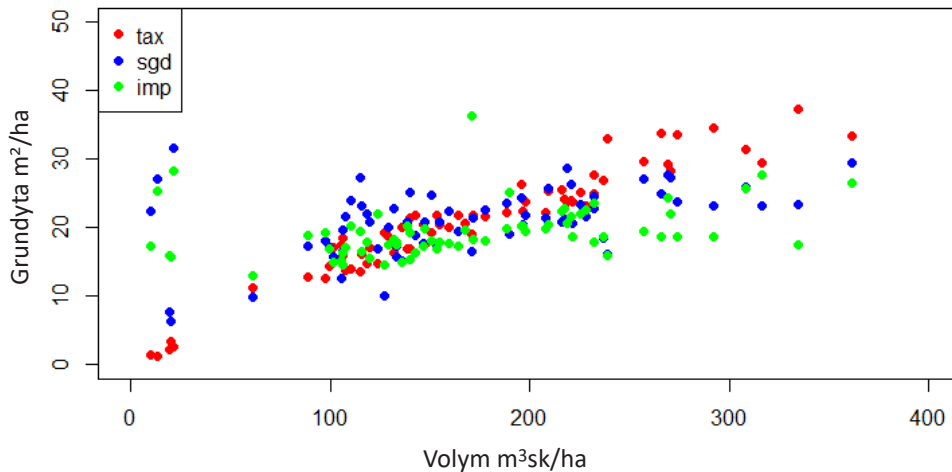


Figur 8. Samband mellan volym och grunddyta för beräkningsytor baserade på skördardata (ref), taxeringsdata för utvärderingsbestånd (tax) samt beståndsskattningar aggregerade från segmentsprognoser (imp).

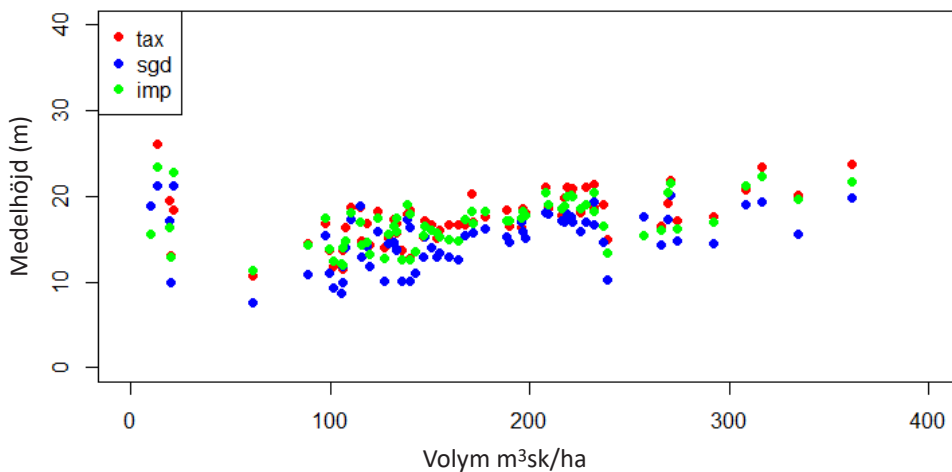


Figur 9. Samband mellan volym och grunddyta för beräkningsytor (svart), sammanslagna skattningar för delsegment av utvärderingsbestånd (grön) samt taxeringsdata från utvärderingsbestånd (röd).

I figur 10 ses skillnaderna i utvärderingsbeståndens grunddyta mellan taxering, skogliga grunddata och skattningen i förhållande till beståndens taxerade volym. Skillnaderna mellan taxering och skattning är markanta för fröträdsställningar och för bestånd med volymer över 250 m³sk/ha. För medelhöjd blir skattningarna mer i linje med taxeringsuppgifter, medan skogliga grunddata, som inte uppdateras, underskattar medelhöjden (figur 11).



Figur 10. Spridning i grunddyta per hektar för utvärderingsbestånd i förhållande till taxerad volym för tre källor; taxeringsdata från utvärderingsbestånd (tax), skogliga grunddata (sgd) samt viktad grunddyta (imp) från imputering på delsegment.

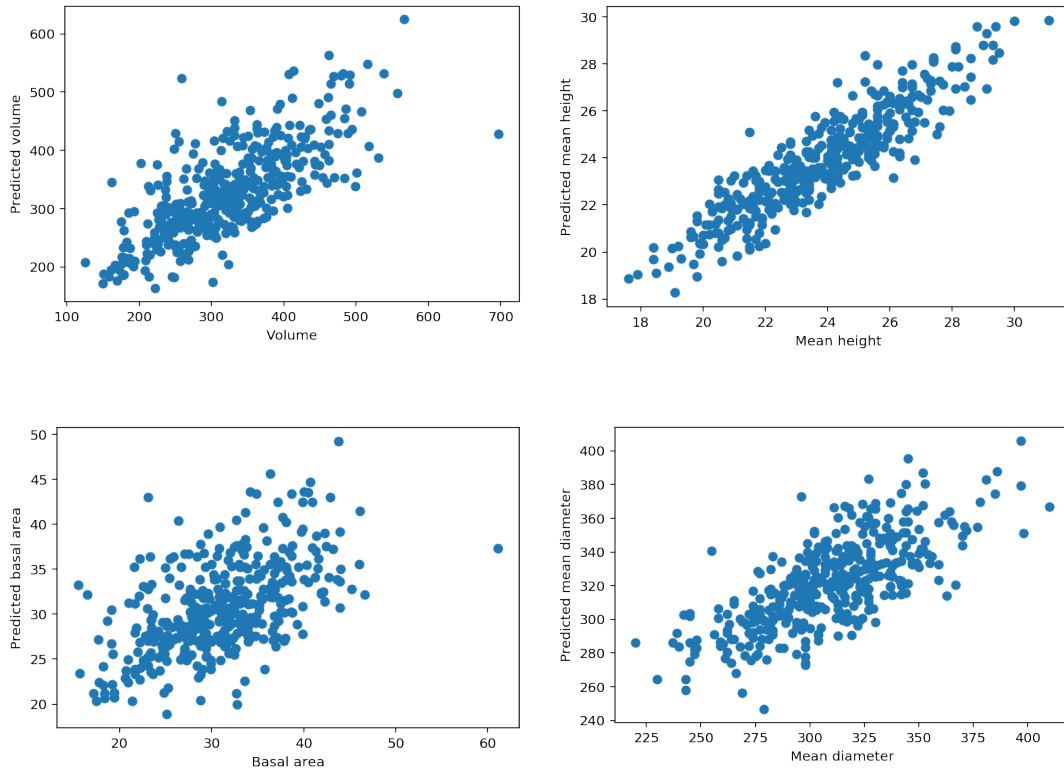


Figur 11. Variation i grundtyevägd medelhöjd för utvärderingsbestånd i förhållande till taxerad volym mellan tre källor; taxeringsdata från utvärderingsbestånd (tax), skogliga grunddata (sgd) samt imputerad medelhöjd (imp) från skattningar på delsegment.

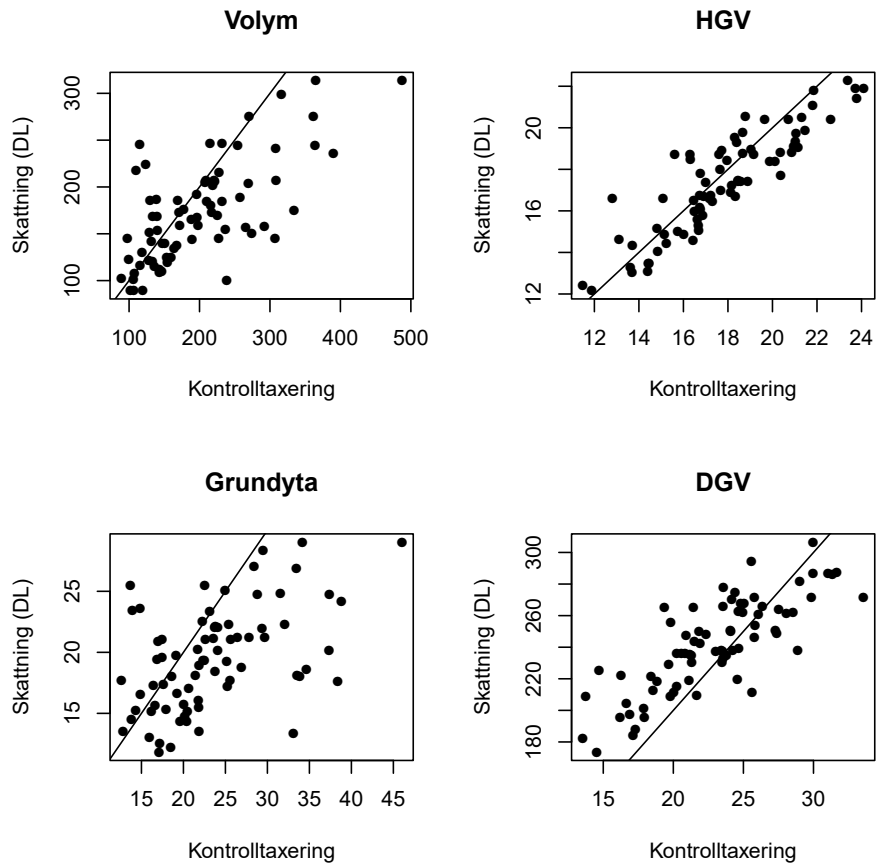
DEEP LEARNING

Ett flertal olika neurala nätverk utvecklades i Tensor Flow. Dessa utvärderades parallellt genom att iterativt träna nätverket på 60 procent av det tillgängliga datamaterialet från skördaravverkningarna inom studieområdet och utvärdera med resterande 40 procent (detta kan möjligen liknas vid en korsvalidering, men utförs som en grundläggande del av träningen av nätverket). Det bästa nätverket användes till sist för att prediktera de kontrolltaxerade avdelningarna från Sveaskog och från dessa prediktioner beräkna noggrannhetsmått (ekvationerna 3–6). Nätverk med djup mellan 2 och 8 provades och med 64 eller 128 noder. Det nätverk som fungerade bäst (Figur 12 och 13) består av tre lager och 64 noder. Större och mer komplexa nätverk tenderade att bli överanpassade och leverera sämre skattningar.

Datamaterialet i denna studie var sannolikt för begränsat för att metoden ska kunna presterar väl, men betydligt mer utvecklingsarbete behövs för att konstruera ett djupt neuralt nätverk som effektivt kan anpassas för denna tillämpning.



Figur 12. Resultat från korsvalidering med Deep Learning (DL).



Figur 13. Resultat från skattningar av utvärderingsbestånden med Deep Learning (DL).

Diskussion och slutsatser

En förutsättning för att kunna använda skördardata som markreferens är kvalitetssäkrad kalibrering av skördarmätningen. Skördarförarna klavar in två träd/stockar per dag och oberoende revisorer kontrollerar skördarmätningen mot inklavade stockar. Koordinater för avverkade träd är också nödvändiga. Idag är trädpositionerna osäkra då endast skördarens position registreras, men ny teknik, främst inom kranstyrning, ger högre precision för enskilda träd (Hauglin m.fl. 2017). Lämnade träd och naturvårdshänsyn är ett problem, men från version 3.1 av Stanford 2010 finns stöd för att dokumentera kvarlämnade träd och naturhänsynsåtgärder. Anpassning av användargränssnitt pågår hos maskintillverkare och ett fåtal testmaskiner är i drift idag.

De utvärderingsdata som användes från Sveaskog håller hög kvalitet och fanns i stor kvantitet över studieområdet. Det gör att resultaten bör ha blivit ordentligt utvärderade, även om metoderna alltid kan utvecklas. Ett speciellt problem med den areabaserade skattningsmetoden som används här är att den baseras på beräkningsytor som inte finns tillgängliga för områden som är mål för prediktering. Den segmentering som utnyttjas för att skapa prediktiva enheter följer inte nödvändigtvis de avdelningsgränser som användes för Sveaskogs kontrolltaxering. Konsekvensen av detta är att projektet redovisar resultat som sannolikt ser ut att vara sämre än de är i verkligheten, eftersom en okänd del av noggrannheten gått förlorad på grund av diskrepansen mellan avdelningsgränserna och skattningssegmenten.

Skördardata samlas rutinmässigt in för i stort sett alla avverkningar i hela landet. Det produceras alltså löpande enorma mängder noggranna data över skogens tillstånd för helt andra ändamål än skoglig kartering. Detta, i kombination med stereomatchade flygbilder eller nya laserdata, som också produceras löpande för hela landet, är en enorm resurs för att beräkna rasterskattningar av skogliga parametrar med nationell täckning. Kostnaderna för data är små och möjligheterna för kontinuerliga uppdateringar av rasterskattningarna är stora. För detta ändamål är DL-metoder mycket lämpade, eftersom de kan utnyttja stora datamängder för att ersätta statistisk modellering med artificiell intelligens. De kan också automatiskt uppdateras med nya data och producera nya rasterskattningar. I detta projekt utprovades en första DL-algoritm som visade klart att metoden fungerar, men betydande ytterligare utvecklingsarbete kommer att behövas för att ta fram en operativ metod.

Resultaten från korsvalideringen var lovande. Användning av beräkningsytor som markreferens gav skattningar i klass med tidigare fotogrammetriska skattningar baserade på provytor från riksskogstaxeringen. Skattningarna gav heller inte några nämnvärda systematiska avvikelser. Skattningarna för segment av utvärderingsbestånden gav liknande samband, men lyckades inte fånga variationen i utvärderingsbestånden fullt ut. Framförallt var några bestånd med mycket låg grundyta problematiska, då de låg utanför referensmaterialets spännvidd. Detta indikerar att det kan finnas behov av kompletterande indata i vissa skogstyper, alternativt att bearbetningen av skördardata behöver justeras för att bättre fånga variationen.

Då skattningarna delvis baserades på lasermetriker från Lantmäteriets nationella laser-skanning inkluderades även skogliga grunddata i jämförelsen (Figur 10 och 11). Skattningarna av grundyta blev förvånande nog ofta sämre än skogliga grunddata i förhållande till taxeringen. När det gäller bestånd med större volymer har ännu inte någon tydlig felkälla kunnat identifieras. För fröträdsställningar har flera felkällor identifierats: 1)

areaberäkningen för beräkningsytor var dålig för glesa avverkningar, en ny version utvecklades av Skogforsk. 2) Flygfotometriker har ofta svårt att beskriva glesa skogar på ett bra sätt. 3) Stora förändringar har skett på ytorna som inte speglas i motsvarande lasermetriker. För medelhöjd blev skattningarna mer i linje med taxeringsresultaten, medan Skogliga grunddata tydligt saknar de senaste årens höjdtillväxt.

Produktionen av data från skördare och olika fjärranalysmetoder är mycket stor idag och kommer sannolikt att öka än mer i framtiden. För att dra full nytta av detta kommer nya analysmetoder att behövas. Deep Learning och andra ML-/AI-metoder har visat sig mycket kraftfulla i andra tillämpningar med liknande mängd data och komplexa frågeställningar. Sannolikt är detta vägen framåt även för de tillämpningar som det här projektet inriktats mot. Även betydligt mer intrikata egenskaper, som virkesegenskaper och skadefrekvens, bör vara möjliga att kartera med hjälp av Deep Learning. Stora mängder träningsdata behövs för att lyckas få de neurala nätverken att fånga dessa komplexa samband.

Referenser

- Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., ... Zheng, X. 2015. TensorFlow: Large-scale machine learning on heterogeneous systems. Retrieved from <https://www.tensorflow.org/> (Software available from tensorflow.org)
- Anon. 2016. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Arlinger, J., Nordström, M. & Möller, J.J. 2012. StanForD 2010. Modern kommunikation med skogsmaskiner. – StanForD 2010. – Modern communication with forest machines. Arbetsrapport 784. Skogforsk. 16 s.
- Bhuiyan, N. 2015. Pers medd. hprYield verktyg för stratifiering och beräkning av skogliga parametrar med hjälp av skördardata.
- Bhuiyan, N., Möller J.J. & Arlinger, J. 2018. hprYield – beräkningsmodul för generering av geografiskt uppdelade nyckeltal baserat på skördardata. Arbetsrapport XX, Skogforsk. XX s. Manuskript under publicering 2018.
- Crookston, N.L. & Finley, A.O. 2007. yaImpute: An R Package for k-NN Imputation. Journal of Statistical Software. 23(10): 1–16.
- Goodfellow, I., Bengio, Y., & Courville, A. 2016. Deep learning. MIT Press.
- Hannrup, B. Bhuiyan, N. & Möller, J.J. 2015. Rikstäckande utvärdering av ett system för automatiserad gallringsuppföljning. Arbetsrapport 857, Skogforsk.
- Hauglin, M., Hofstad Hansen, E., Naesset, E., Busterud, B.E., Omholt Gjevestad, J.G. & Gobakken, T. 2017. Accurate single-tree positions from a harvester: a test of two global satellite-based positioning systems. Scandinavian Journal of Forest Research, 32:8, p. 774–781.
- Holmgren, J., Barth, A., Larsson, H., & Olsson, H. 2012. Prediction of stem attributes by combining airborne laser scanning and measurements from harvesters. Silva Fennica, 46(2), 227 - 239.

- Lantmäteriet. 2018. Produktbeskrivning: Ytmodell från flygbilder och Ytmodell från flygbilder IRF. Dokumentversion 1.0 (2016-10-25). Lantmäteriet, Gävle. (http://www.lantmateriet.se/globalassets/kartor-och-geografisk-information/hojddata/produktbeskrivningar/ytmodell_fran_flygbild2.pdf).
- Leach J., Lidberg W., Kuglerová L., Kuglerova L., Peralta Tapia A., Ågren A. & Laudon H. 2017. Evaluating topography-based predictions of shallow lateral groundwater discharge zones for a boreal lake-stream system. *Water Resources Research*, 53 (7), pp.5420-5437.
- Moer, M. & Stage, A.R. 1995. Most Similar Neighbor: An Improved Sampling Inference Procedure for Natural Resource Planning. *Forest Science*, 41(2), pp 337–359.
- Möller, J. J, Arlinger, J., Barth, A., Bhuiyan, N., & Hannrup, B. 2011. Ett system för beräkning och återföring av skördarbaserad information till skogliga register och planeringssystem. Arbetsrapport 756, Skogforsk.
- Möller, J.J., Bhuiyan, N. & Hannrup, B. 2015. Utveckling och test av beslutsstöd vid automatiserad gallringsuppföljning. Uppsala: Skogforsk. (Arbetsrapport 862-2015).
- Möller, J.J., Söderberg, J., Arlinger, J., Eriksson, I., Bhuyian, N. & Willén, E. 2017. Utbytesprognoser baserat på skogliga data och skördarinformation – Modell och systembeskrivning för ”Skapande av stambanker och imputerad utbytesprognos baserat på skogs- och skördardata på objektsnivå”. Uppsala: Skogforsk. (Arbetsrapport 961-2017).
- Næsset, E., Gobakken, T., Holmgren, J., Hyyppä, H., Hyyppä, J., Maltamo, M., . . . Soderman, U. 2004. Laser scanning of forest resources: The Nordic experience. *Scandinavian Journal of Forest Research*, 19(6), 482-499.
- Pennock, D.J., Zebarth, B.J., & De Jong, E. 1987. Landform classification and soil distribution in hummocky terrain, Saskatchewan, Canada. *Geoderma* 40(3): 297 – 315.
- Rothermel, M., Wenzel, K., Fritsch, D., & Haala, N. 2012. SURE - photogrammetric surface reconstruction from imagery. In *Lc3d workshop* (p. 1-9). Berlin, Germany.
- Söderberg, J. 2015. A method for using harvester data in airborne laser prediction of forest variables in mature coniferous stands, Arbetsrapport 443. Sveriges Lantbruksuniversitet, Inst. för skoglig resurshushållning, 901 83 Umeå.
- Sørensen, R., Zinko, U., Seibert, J. 2006. On the calculation of the topographic wetness index: evaluation of different methods based on field observations. *Hydrology and Earth System Sciences*. 10: 101–112.