



**Kandidatarbeten  
i skogsvetenskap**  
Fakulteten för skogsvetenskap

2012:21

## **Samband mellan laserdata och fältdata**

*Connections between laser data and field measured data*

Björn Karlsson

# Kandidatarbeten i Skogsvetenskap

Fakulteten för skogsvetenskap

Enhet/Unit	Institutionen för skogens ekologi och skötsel Department of Forest Ecology and Management
Författare/Author	Björn Karlsson
Titel, Sv	Samband mellan laserdata och fältdata
Titel, Eng	Connections between laser data and field measured data
Nyckelord/ Keywords	flygburen laser, skogsinventering, fjärranalys, skattningar av skogsegenskaper
Handledare/Supervisor	Kenneth Olofsson Institutionen för skoglig resurshushållning
Examinator/Examiner	Tommy Mörling Institutionen för skogens ekologi och skötsel/ Department of Forest Ecology and Management
Kurstitel/Course	Kandidatarbete i skogsvetenskap Bachelor Degree in Forest Science
Kurskod	EX0592
Program	Jägmästarprogrammet
Omfattning på arbetet/	15 hp
Nivå och fördjupning på arbetet	G2E
Utgivningsort	Umeå
Utgivningsår	2012

# Samband mellan laserdata och fältdata

*Connections between laser data and field measured data*

**Av: Björn Karlsson Jägmästarkurs 09/14**

**Handledare: Kenneth Olofsson**

## **FÖRORD**

Detta är ett kandidatarbete på jägmästarprogrammet vid skogshögskolan tillhörande Sveriges lantbruksuniversitet i Umeå. Arbetet kom till våren 2012 och omfattar 15 högskolepoäng.

Det är flera som hjälpt mig med detta arbete, men jag vill ge ett särskilt tack för värdefull hjälp till Kenneth Olofsson som har handlett uppsatsen, och Johan Holmgren som gav mig grunddatat och varit till hjälp under analysen.

## SAMMANFATTNING

För att underlätta för skoglig planering behöver man inom skogsbruket data som beskriver egenskaper hos skogen. Sådana data får man vanligtvis ifrån någon form av inventering. En relativt ny teknik som förmodligen kommer att bli viktig i framtiden är laserskanning, på grund av lägre kostnader. Men i och med att lasertekniken är relativt ny så bör man jämföra den med konventionell inventeringsteknik för att få bättre förståelse för de möjligheter lasertekniken ger och bättre förstå när och hur man kan utnyttja den. Detta arbete syftar till att skatta fältmätta egenskaper hos skogen utifrån laserdata genom regressionsanalys inom ett område i Västerbotten och sedan se vilka av de fältmätta variablerna som bäst kunde skattas, samt vilka laservariabler som gav bäst resultat.

Den här studien visar på att höjd skattas bäst, men att biomassa, diameter och volym också skattas ganska bra men lite sämre. Sämst blev skattningarna för stamantal och ålder. Vilket förmodligen beror på att man bara indirekt mäter egenskaper som är korrelerade med dessa data. Förklaringsgrad och RMSE % blev; höjd 88,6 %, 8 %; biomassa 79,1 %, 23,3 %; diameter 78,3 %, 13,7%; volym 73,8 %, 29 %; stamantal 66,4 %, 36 %; samt ålder 63,5 %, 27 %.

De laservariabler som verkar vara viktigast är percentiler tillsammans med olika kvoter som beskriver täthet av punkter. I framtiden finns stora möjligheter att laserskanning ska kunna användas för olika typer av inventeringar i skoglig planering.

**Nyckelord:** flygburen laser, skogsinventering, fjärranalys, skattningar av skogsegenskaper

## SUMMARY

To make it easier for forestry planning the forest owners needs data which describe characteristics of the forest. Such data you usually get from some type of inventory. A relatively new technology which probably is getting more important in the future, is laser scanning because of lower costs. But because the laser technology is relatively new you need to compare it with conventional inventory technologies to get better understanding of the possibilities the laser technology gives. This work intends to estimate field measured characteristics of the forest from laser data with help of regression analysis and then see which of the field measured properties that best could be estimated and which laser variables that gave the best results.

This study shows that heights has the best estimates, but biomass, diameter and volume also gives good estimates. How stem numbers and age does not give equally good estimates. Which probably depends on that these variables only measured indirect by characteristics which is correlated with these data. The coefficient of determination and RMSE % become; high 88,6 %, 8 %; biomass 79,1 %, 23,3 %; diameter 78,3 %, 13,7%; volume 73,8 %, 29 %; stem number 66,4 %, 36 %; and for age 63,5 %, 27 %.

The laser variables which seem to be most important were percentiles and different ratios which describe density of points. In the future there is a great potential for laser scanning which can be used for inventories in forest planning.

**Keywords:** airborne laser, forest inventory, remote sensing, estimates of forest characteristic

**Nota bene!**

**Summarium in lingua latina in appende I, est!**

# 1 INLEDNING

## 1.1 Bakgrund

För att underlätta för skoglig planering behöver man inom skogsbruket data som beskriver egenskaper hos skogen. Sådana data får man vanligtvis ifrån någon form av inventering. En relativt ny teknik som förmodligen kommer att bli viktig i framtiden är laserskanning, i och med att de totala kostnaderna är mindre jämfört med konventionella metoder (Eid m.fl. 2004). Studier har till exempel visat att bland annat höjd kan skattas mycket bra, möjligen även bättre än konventionella metoder (Næsset, 1997).

Lantmäteriet genomför idag en rikstäckande laserskanning för att på regeringens uppdrag kunna skapa en Ny Nationell Höjdmodell av marken. En del av mätpunkterna träffar marken, medan andra träffar exempelvis byggnader och vegetation (Lantmäteriet, 2011). Laserskanning fungerar så att man flyger över ett område och sänder ut pulser. När de sedan träffar något till exempel en byggnad, marken eller vegetation, reflekteras de sedan tillbaka och registreras av mätsensorn. Genom att man mäter tiden som pulsen har färdats kan man räkna ut avståndet (Dubaya & Drake, 2000) och därigenom få en tredimensionell koordinat. Laserdatat som bildats består av ett punktmoln med koordinater där varje punkt har en känt läge i markplanet X, Y och en höjd över marken. Dessa punkter kan delas upp i olika klasser (Lantmäteriet, 2011), exempelvis som tidigare nämnts mark eller vegetation.

Just de här punkterna som träffar högre vegetation är intressanta för skogsbruket på grund av att de kan utnyttjas för att skatta beståndsegenskaper som höjd, volym, diameter med mera. I och med att Lantmäteriverket gör laserdatat tillgängligt för alla, kan fler komma att utnyttja tekniken och det gör att intresset förmodligen kommer att öka. Därmed kan lasermätningar i framtiden bli en allt vanligare teknik inom skogsbruket än idag, även om Lantmäteriverket inte skulle göra några fler omdrev. Intresset kan ha blivit så stort att åtminstone de större skogsbolagen kan kosta på att laserskanna sina skogar som en inventeringsmetod. Det finns olika metoder för att utnyttja laserdata för att skatta skogsegenskaper, i den ena baseras modellerna på enskilda träd och då behövs en tätare mängd laserpunkter (Hyypä m.fl. 2001). Den modell jag har använt mig av är baserad på provytor i stället för enskilda träd (Næsset, 2007). Den kallas areametoden och används vid låg täthet av laserpunkter likt studien av Næsset m.fl. (2004).

En tidigare studie som gjorts i norra Sverige på det här ämnesområdet är Martin Sjödens *Skattningar i gallringsskog med hjälp av flygburen laserskanning* (2010). I den här studien har han kopplat ihop laserdata med fältmätta värden via regressionsanalys och k-nearest neighbour-modeller för ett barrdominerat område i närheten av Älvsbyn i Norrbotten. Han skattade höjd och massaslutenheten och sedan beräknade han bland annat volymen. För att skatta den grundtyevägda höjden så använde han percentil 70 och

percentil 95 till förklarande variabler. Olika percentiler var även viktiga i funktionerna för övre höjd samt massaslutenhet. För massaslutenheten användes även skanningsvinkeln som förklarande variabel. Han fick hög förklaringsgrad på grundtyvägd höjd och volym, 93,7 %, 92,6 % och ett lägre RMSE för grundtyvägd höjd än för volym, 5,2 % och 15,9 %. Sjödin kom även fram till att laserskanning hade hög pålitlighet.

En annan studie på området är Martin Ferms *Prediktering av skogliga variabler med data från flygburen laser* (2010) som genomfördes i ett skogsområde i Östergötland. Även han har försökt att skatta egenskaper hos skogen med hjälp av både regressionsanalys och k-nearest neighbour-modeller. Volym och höjd kunde skattas bra med regressionsanalys medan den andra typen av modeller fungerade bra för att skatta grundtyvägd diameter samt trädantal. Även han fick högst förklaringsgrad på höjd, men även högt på volym, cirka 90 % och 85 %. Samt lägst på antal träd, cirka 68 %. Han kom bland annat fram till olika beståndsegenskaper är olika svåra att skatta. Exempelvis kom han fram till att antal träd/ha grundtyvägd diameter var svårare att skatta jämfört med volym och grundtyvägd medelhöjd.

Næsset, (2007) genomförde en studie i södra delen av Østlandet i Norge. Områdena var barrträdsdominerade. Han använde regressionsanalys för att skapa sina formler och kunna jämföra laserskattningar med fältmätta värden. Han kom fram till att man kan skatta skogliga egenskaper bra med hjälp av laserdata. RMSE blev för volym för de två studieområdena 10,6 % och 14 %, diameter 10,4 % och 12,9, stamantal 21,3 % och 19,0 %, samt övre höjd 3,8 % och 6,1 %. Næsset använde bland annat av percentil 70 för att kunna skatta övre höjd. Till volymen användes bland annat olika lasermått på krontakets täckning (krontätheter). För grundtyvägd diameter användes också både höjdpercentiler och täthetsmått som förklarande variabler och till stamantal användes ett mått på variation av första pulsens höjd.

2002 genomförde Næsset en annan studie i Våler som också ligger i Østlandet, Norge. Han fick för den här studien förklaringsgrader på 74-93 % för övre höjd, 39-78 % för medeldiameter, 50-68 % för stamantal samt 80-93 % för volym. Han fick olika resultat för olika typer av skog, ung och äldre skog, samt för äldre skog med god respektive dålig bonitet. I funktionerna var både diverse percentiler och krontakets täthet viktiga förklarande variabler hos laserdatat.

Lasertekniken kan ha sina fördelar som vi sett ovan, men i och med att den är relativt ny kan det vara viktigt att utvärdera hur resultaten skiljer sig jämfört med konventionell teknik. Genom att man förstår hur man på bästa sätt kan utnyttja de möjligheter lasertekniken ger och var den är tillämpbar så kan man på ett mycket bättre sätt förstå när och hur man ska använda den.



## 1.2 Syfte

Målet med denna studie är att se hur pass bra olika fältmätta egenskaper hos skogen, ålder, grundtevägd diameter, grundtevägd höjd, stamantal, volym och biomassa kan skattas med hjälp av data från flygburen laser genom att tillämpa regressionsanalys och areametoden (Næsset, 2007).

Specifikt vill jag svara på dessa frågeställningar:

- Hur pass bra går det att skatta fältmätta egenskaper med hjälp av laserdata?
- Vilka fältmätta egenskaper hos skogen går bäst att skatta utifrån laserdata?
- Vilka laservariabler är viktiga för att skatta beståndsegenskaper?
- Hur liknar eller skiljer sig, dessa resultat från andra liknande studier?

Dessa resultat kan ge en bättre förståelse för hur laserdata kan användas för att skatta skogliga egenskaper som behövs vid exempelvis skoglig planering.

## 2 MATERIAL OCH METOD

### 2.1 Studieområde

Datat som studien bygger på är insamlat i ett område kallat Krycklan beläget cirka 3 km nordöst om Vindeln i landskapet Västerbotten. Provytorna har en trädslagsblandning på cirka 70 % tall (*Pinus sylvestris* L.), 20 % gran (*Picea abies* (L.) Karst.) och 10 % löv, avrundat till jämna tiotal. Arealen av det totala laserskannade området är 70 km<sup>2</sup> (Blom, 2008).

### 2.2 Fältdata

Fältdatat som har använts har inventerats enligt Heureka's fältdatainsamlingsmetod (Heurekaenheten SLU, 2011) och Ivent (Wikström, 2011). Provytorna hade 10 meters radie och var systematiskt utlagda. På provytorna klavades alla träd samt vissa höjdmättes. Det fältdata som har använts i denna studie var insamlat sedan tidigare och var fördelat på 404 stycken provytor. De egenskaper hos skogen som studerats som responsvariabler samt vad de mättes i var:

- Grundtyvägd medeldiameter (cm)
- Grundtyvägd medelhöjd (dm)
- Medelstamantal (stammar/ha)
- Medelvolym per hektar (m<sup>3</sup>/ha)
- Medelålder (år)
- Biomassa (ton/ha)

### 2.3 Laserdata

Laserdatat som är använt i denna studie är beställt av Sveriges lantbruksuniversitet, SLU och Totalförsvarets forsknings institut, FOI. Det är insamlat med hjälp av helikopter 5-6 augusti 2008 av operatören Blom. Den utrustning som användes var TopEye system S/N 425 och efter har en visuell täckningskontroll genomförts. Vid GPS-mätningen låg SWEPOS referensstation i Vindeln alltid inom ett avstånd på mindre än 15 km. Flyghöjden var 500 m men tvärstråken flögs på 250 m. Den totala mängden laserpunkter var 940 000 000 stycken och punkttätheten blev då 5 punkter/m<sup>2</sup> och tvärstråk 15 punkter/m<sup>2</sup> (Blom, 2008).

De laservariabler som utvärderats som förklarande variabler var (i meter över mark, samt andel):

- Percentilerna, 50, 70, 80, 90, 95, 100 av vegetationspulserna
- Medelhöjd av vegetationspulserna
- Andel vegetationspulser av det totala antalet
- Standardavvikelse för vegetationspulserna

## 2.4 Analys av datat genom regression

Datat erhöles i en tabell med medelvärden för de fältmätta egenskaperna volym, grundtyevägd medelhöjd, grundtyevägd medeldiameter, stamantal, biomassa och ålder för varje provyta. För varje provyta erhöles också olika laservariabler i bearbetad form, diverse percentiler, standardavvikelse av alla vegetationspulser samt andel vegetationspulser, se avsnitt 2.3 Laserdata.

För att kunna skapa modeller, för att skatta egenskaperna hos skogen som man är intresserad av med hjälp av laservariablerna användes linjär regression och areametoden. De olika laservariablerna prövas för att se hur väl de beskriver fältvärdena och de som bäst beskriver fältvärdena används för att ingå i de slutliga funktionerna. Oftast användes multipel linjär regression i och med att oftast förbättrades skattningen om man tog med fler olika laservariabler, förutom vid skattningen av höjd då enkel linjär regression användes. P-värdet för samtliga förklarande variabler var mindre än 0,001, dock fanns en del uteliggare. Under analysen skattades fältmätta värden med laservariabler och olika modeller för skogsegenskaperna skapades. Den naturliga logaritmen, LN utnyttjades för att få ett bättre resultat samt att så nära som möjligt försöka följa de vanliga statistiska antagandena (Hadi & Chatterjee, 2006). Regressionsanalysen genomfördes i datorprogrammet Minitab 16 där sambanden kan hittas genom både grafiska-, exempelvis via spridningsdiagram, och statistiska metoder, exempelvis p-värde.

## 2.5 Validering av modellerna

Funktionerna skapades på ett separat träningsdataset och utvärderades sedan på ett separat utvärderingsdataset. Träningsdatasetet utgjordes av 324 provytor och utvärderingsdatasetet bestod av 79 stycken provytor. En provyta användes inte vid regressionen på grund av att höjdvärdet var orimligt, förmodligen beroende på fel vid inventeringen. De 79 utvärderingsprovytorna valdes ut genom att slumpa ett startnummer, sedan blev var femte yta utvald. Genom att 79 avdelningar utelämnats vid skapandet av funktionen kunde man sedan skatta beståndsegenskaperna på dessa provytor och se hur stor skillnad och hur bra anpassade formlerna blev.

På utvärderingsdatasetet uträknades noggrannhetsmått för att kunna jämföra de olika skattningarna. RMSE och standardavvikelse är några vanliga mått som ofta används vid jämförelser och utvärderingar av modeller. Funktionerna för dessa presenteras nedan (Næsset, 2007). De mått som använts för jämförelse och som är i procent, står relativt mot medelvärdet av de fältmätta värdena.

$y_i =$  fältmätt värde för provyta  $i$

$\hat{y}_i =$  skattat värde för provyta  $i$

$n =$  antal provytor

$D_i =$  differans mellan fältmättvärde och skattatvärde

Medeldifferensen:

$$\bar{D} = ME = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)}{n}$$

Standardavvikelsen räknades ut i Minitab 16, här är en allmän formel:

$$SD = s = \sqrt{\frac{\sum_{i=1}^n (D_i - \bar{D})^2}{n - 1}}$$

RMSE, Root mean square error, MSE räknades ut i Minitab 16

$$RMSE = \sqrt{MSE} = \sqrt{\frac{\sum_{i=1}^n D_i^2}{n}}$$

Förklaringsgraden som även den räknades ut av Minitab 16 kan räknas ut så här (Hadi & Chatterjee, 2006):

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2}$$

$\bar{y} =$  medelvärdet för de fältmätta värdena

Justerad förklaringsgrad (Hadi & Chatterjee, 2006):

$$R_a^2 = 1 - \frac{n-1}{n-p-1} (1 - R^2)$$

*p = antal parametrar*

## 3 RESULTAT

### 3.1 Noggrannhet och förklaringsgrad

**Tabell 1.** Mått på noggrannheten i skattningarna. RMSE = root mean square error, RMSE % = RMSE i procent av det fältmätta medelvärdet, s = standardavvikelse, s % = standardavvikelsen i procent av det fältmätta medelvärdet, ME = medeldifferensen,  $R^2$  = förklaringsgrad,  $R^2_a$  = justerad förklaringsgrad

*Table 1. Measurements of accuracy in the estimates. RMSE=root mean square error, RMSE%=RMSE in percentage of the field measured mean, s= standard deviation, s%=standard deviation in percentage of the field measured mean, ME= Mean difference,  $R^2$  = coefficient of determination,  $R^2_a$ =adjusted R-squared*

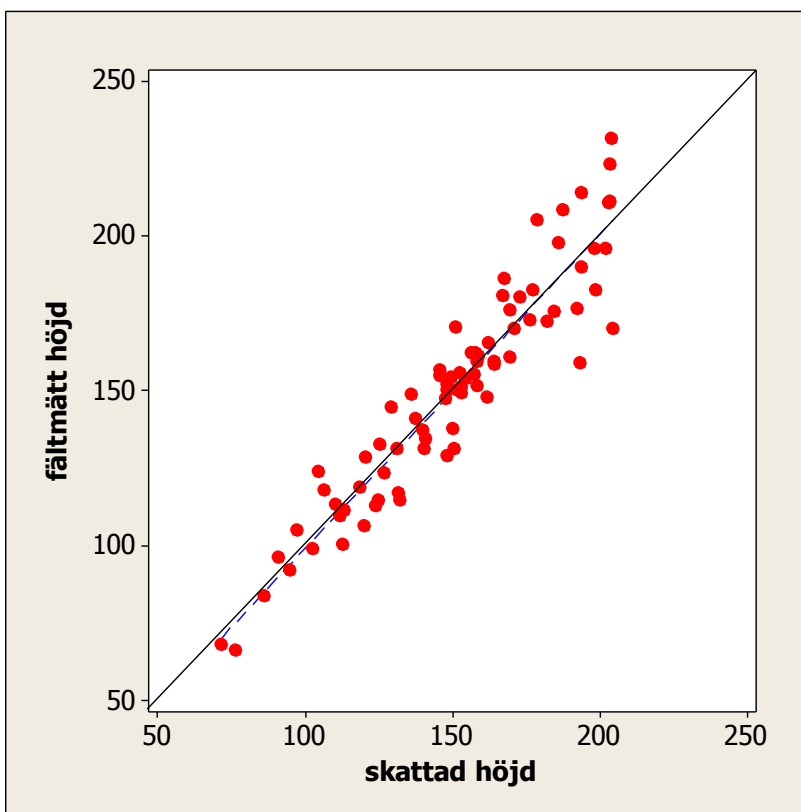
Skattad variabel	RMSE	RMSE %	s	s %	ME	$R^2$	$R^2_a$
Ålder (år)	24,88	27	24,89	27	0,23	63,5	63,1
Grundtevägd diameter (cm)	2,86	13,7	2,87	13,7	0,18	78,3	78
Grundtevägd höjd (dm)	12	8	12,02	8	-0,45	88,6	88,5
Stammar/ha	458,5	36	458,5	36	87,6	66,4	65,9
Volym (m <sup>3</sup> /ha)	51,6	29	51,6	29	1,43	73,8	73,5
Biomassa (ton/ha)	20,98	23,3	20,98	23,3	1,66	79,1	78,9

Nedan beskrivs noggrannheten i skattningarna för ålder, grundtevägd diameter, grundtevägd höjd, stamantal, volym och biomassa efter validering (Tabell 1), för definitioner se avdelning: 2.5 Validering av modellerna. Förklaringsgraden ligger för samtliga variabler över 63 %. Den justerade förklaringsgraden ligger något lägre än förklaringsgraden, men ligger så pass nära varandra ändå, så hädan efter kommer jag bara hänvisa till ”förklaringsgrad” för bägge två. Den skattning som bäst förklaras är den av grundtevägd höjd på cirka 89 %. Biomassa och grundtevägd diameter kommer sedan med cirka 78-79 %. Volymen skattas också ungefär lika fast något sämre, cirka 74 %. De skattningar som får sämst förklaringsgrad är stamantal respektive ålder med cirka 66 % och 63 %.

Resultatens RMSE och standardavvikelse var nästan identisk och därför gäller nedanstående tolkning dem båda. Högst RMSE och standardavvikelse har stamantalet både vanligt cirka 459 st/ha och relativt sett 36 %. Den lägsta har grundtevägd höjd och grundtevägd diameter, beroende på om man ser på det relativt eller absolut. Grundtevägd höjd hade ett RMSE och standardavvikelse på 12 dm och 8 %. Den grundtevägda medeldiametern hade ett RMSE och standardavvikelse på cirka 3 cm samt 14 %. Även biomassa 21 ton/ha, 23 %; volym 52 m<sup>3</sup>/ha, 29 % och ålder 25 år, 27 % hade ganska höga RMSE och standardavvikelser.

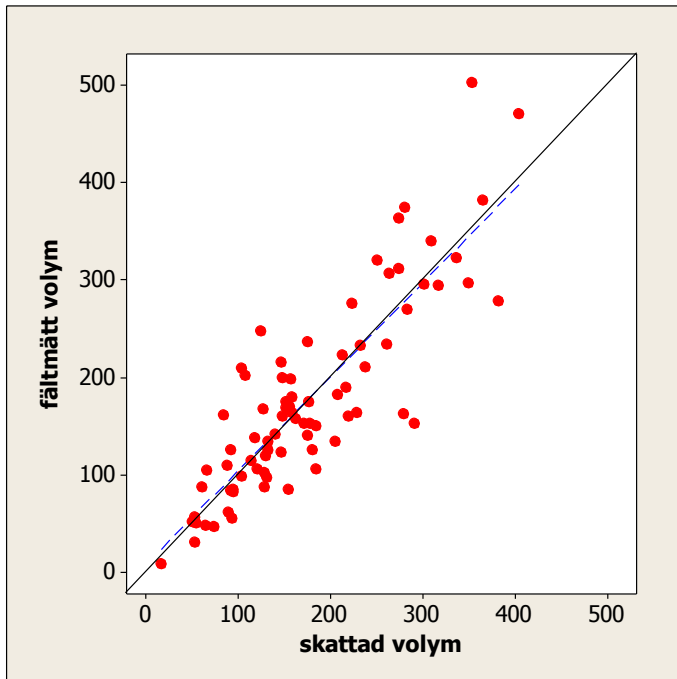
Medeldifferensen var inte så stor för de skattade egenskaperna hos skogen förutom för stamantalet där den var mycket stor. Stamantalets differens var cirka 88 st/ha. Den absolut minsta medeldifferensen, närmast 0, hade den grundtevägda medeldiametern på 0,18 cm och ålder på 0,23 år.

Nedan presenteras de respektive fältmätta egenskaperna mot de med hjälp av ovanstående modeller skattade egenskaperna hos skogen (Figur 1-6). Detta för att kunna visuellt bedöma hur bra skattningarna blev och huruvida det finns systematiska fel. För alla dessa skogliga egenskaper finns ett mer eller mindre tydligt linjärt samband. Mest önskvärt hade varit om alla röda punkter hade varit på den svarta 1:1 linjen i och med att då hade skattningarna varit exakt desamma som de fältmätta. Nu är det inte helt så. De röda punkterna har en viss spridning runt 1:1 linjen och om man gör en linjär regression mellan dessa värden, den blå streckade linjen, kan man se hur väl de stämmer överens med 1:1 linjen. Ju mer lika den blå regressionslinjen är 1:1 linjen desto bättre enligt ovanstående resonemang. Trots tidigare nämnd transformering med hjälp av den naturliga logaritmen, LN så finns viss heteroskedastisitet kvar. Det vill säga att residualerna ökar med högre värden och punkterna kan ha ett lite trattformat mönster. De flesta graferna visar inga större systematiska fel, men ålder och särskilt stamantal uppvisar systematiska fel i och med att den blå streckade regressionslinjen och 1:1-linjen skiljer sig från varandra och de röda punkterna uppvisar dålig passform. Biomassa, volym och höjd stämmer dock ganska väl överens med 1:1 linjen.



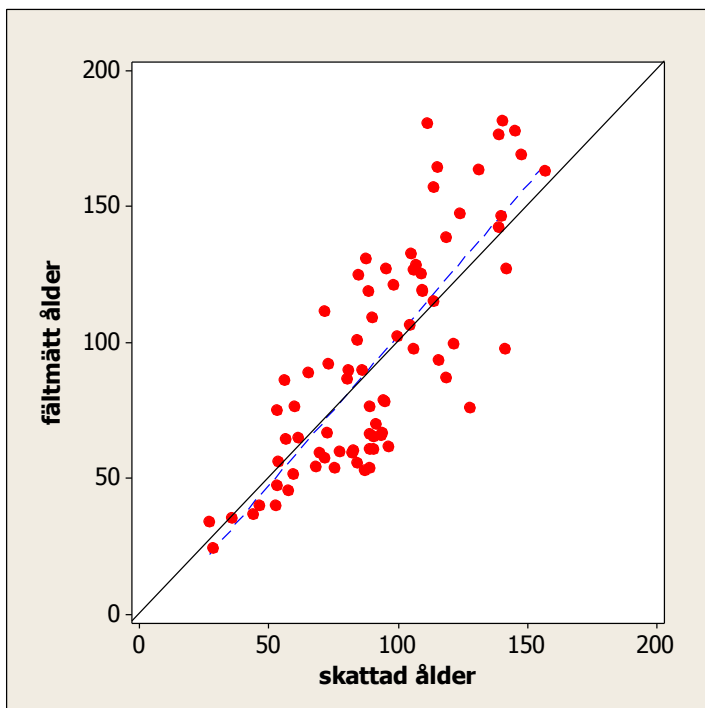
**Figur 1.** Spridnings diagram över fältmätt höjd mot skattad höjd, i dm. Svart linje: 1:1. Blå streckad linje: regressionslinje

*Figure 1. Scatterplot of field measured height versus estimated height, in dm. Black line: 1:1. Blue broken line: regression line*



**Figur 2.** Spridningsdiagram över fältmätt volym mot skattad volym, i  $\text{m}^3/\text{ha}$ . Svart linje: 1:1. Blå streckad linje: regressionslinje

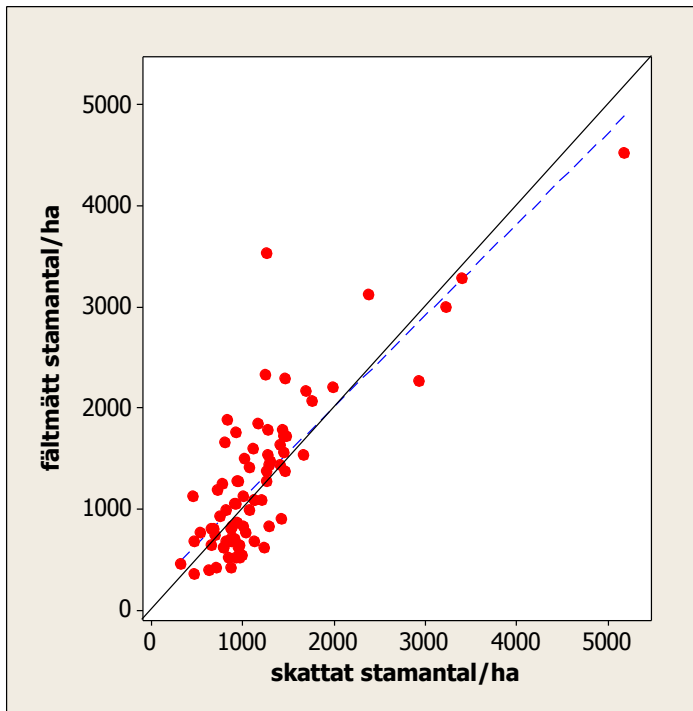
*Figure 2.* Scatterplot of field measured volume versus estimated volume, in  $\text{m}^3/\text{ha}$ . Black line: 1:1. Blue broken line: regression line



**Figur 3.** Spridningsdiagram över fältmätt ålder mot skattad ålder, i år. Svart linje: 1:1. Blå streckad linje: regressionslinje

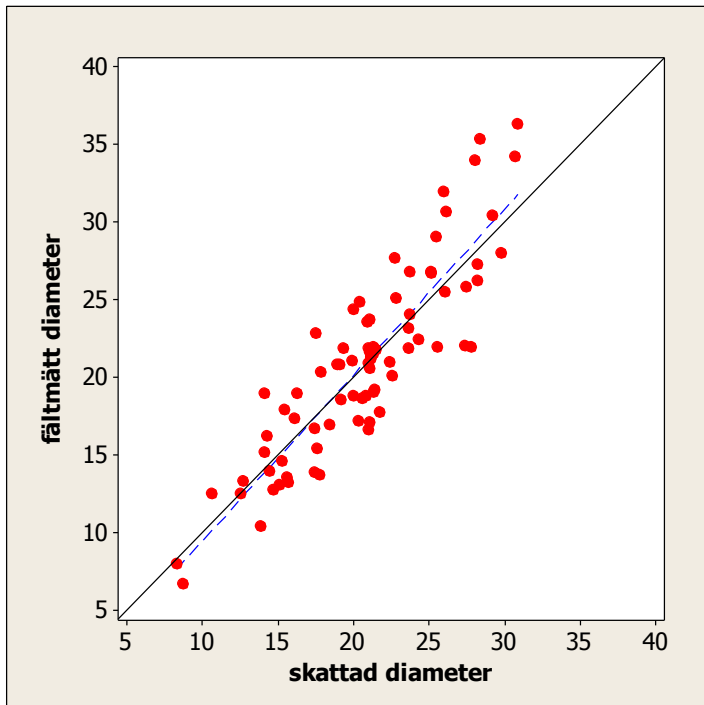
*Figure 3.* Scatterplot of field measured age versus estimated age, in years. Black line: 1:1. Blue broken line: regression line





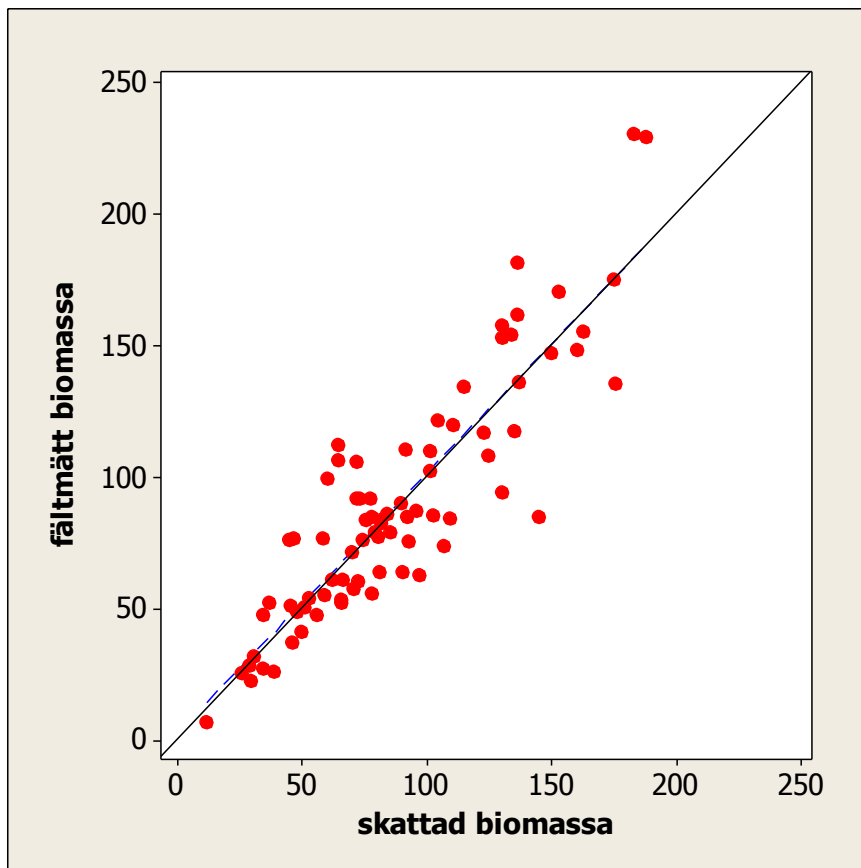
**Figur 4.** Spridningsdiagram över fältmätt stamantal mot skattat stamantal, i stammar/ ha. Svart linje: 1:1. Blå streckad linje: regressionslinje

*Figure 4. Scatterplot of field measured stem numbers versus estimated stem number, in stems/ha. Black line: 1:1. Blue broken line: regression line*



**Figur 5.** Spridningsdiagram över fältmätt diameter mot skattad diameter, i cm. Svart linje: 1:1. Blå streckad linje: regressionslinje

*Figure 5. Scatterplot of field measured diameter versus estimated diameter, in cm. Black line: 1:1. Blue broken line: regression line*



**Figur 6.** Spridningsdiagram över fältmätt biomassa mot skattad biomassa, i ton/ha. Svart linje: 1:1. Blå streckad linje: regressionslinje

*Figure 6.* Scatterplot of field measured biomass versus estimated biomass, in ton/ha. Black line: 1:1. Blue broken line: regression line

### 3.2 De skattade modellerna och deras laservariabler

För grundtytevåg höjd skapades detta samband där percentil 90 utnyttjades:

$$LN(\text{fältmätt höjd}) = 2,81 + 0,808 LN(\text{laserpercentil } 90)$$

För volym skapades detta samband, där både percentil 50 och andelen vegetationspulser var viktiga.

$$LN(\text{fältmätt volym}) = 2,43 + 1,32 LN(\text{laserpercentil } 50) + 1,07 LN(\text{andel vegetationspulser})$$

Åldern fick ett annat samband, där percentil 95, standardavvikelsen hos laserpulserna och andelen vegetationspulser var viktiga.

$$LN(\text{ålder}) = 1,25 + 0,987 LN(\text{laserpercentil } 95) + 0,314 LN(\text{laser standardavvikelse}) - 0,235 LN(\text{andel vegetationspulser})$$

Skattningen av stamantal fick denna modell. Här var andelen vegetationspulser samt percentil 95 viktiga.

$$LN(stamantal) = 10,5 + 1,71 LN(andel\ vegetationspulser) - 0,952 LN(laserpercentil\ 95)$$

För den grundtevägda diametern var percentil 80, standardavvikelsen hos laserpulserna samt andelen vegetationspulser viktiga.

$$LN(fältmätt\ diameter) = 0,781 + 0,714 LN(laserpercentil\ 80) + 0,207 LN(laser\ standardavvikelse) - 0,208 LN(andel\ vegetationspulser)$$

Biomassan fick detta samband. Här var percentil 80, andelen vegetationspulser och standardavvikelsen hos laserpulserna viktiga.

$$LN(biomassa) = 1,15 + 1,59 LN(laserpercentil\ 80) + 1,07 LN(andel\ vegetationspulser) - 0,308 LN(laser\ standardavvikelse)$$

## 4 DISKUSSION

### 4.1 Skattningarnas noggrannhet och förklaringsgrad

Liktydigt med tidigare studier om laserskanning (Næsset 1997a; Næsset 1997b) har denna studie visat att det går att skatta fältmätta egenskaper hos skogen utifrån laserdata. Dock har olika fältmätta egenskaper kunnat skattas olika bra. Denna studie fick i fallande ordning från bäst till sämst skattning, grundtyevägd medelhöjd, biomassa, grundtyevägd medeldiameter, volym, stamantal och sist ålder, utifrån förklaringsgrad.

Mina resultat visar på att höjden verkar kunna vara den egenskap som skattas bäst. Den har lägst procentuell standardavvikelse, 8 % och RMSE, 8 %. Den har också högst förklaringsgrad, 88,5 % av alla egenskaperna. Högst förklaringsgrad för höjden verkar också Ferm (2010) och Sjödin (2010) fått. De fick bägge en förklaringsgrad på runt 90 % för höjd vilket är mycket likt vad den här studien har lyckats uppnå. Även Næsset (2002) uppnådde liknade förklaringsgrader för höjd, dock något lägre för äldre skog. I en annan studie av Næsset (2007) fick han RMSE på 3,8 och 6,1 %. Dock för övre höjd, men det bör kunna användas för en ungefärlig jämförelse mot grundtyevägd höjd. Hans RMSE var något lägre än vad som uppnåtts i denna studie. Att höjden kan skattas så pass bra som den tror jag kan bero på att det i princip är höjden för olika punkter i krontaket man mäter med hjälp av lasern (Dubaya & Drake, 2000). Då borde det rent logiskt kunna ge en relativt bra skattning av höjden. Dock har tidigare studier påvisat att höjden kan underskattas något (Næsset 1997b), vilket även kan antydans i denna studie i och med att regressionslinjen för höjd ligger något lägre än 1:1 linjen (figur 1). Detta skulle kunna förklaras av att trädens toppar träffas med mindre sannolikhet än övriga kronan på grund av ytan (Nelson m.fl. 1988).

På runt 76 % förklaringsgrad var biomassa och volym. Biomassan, 78,9 % fick något högre förklaringsgrad än volymen 73,5 %. Volymen fick Sjödin (2010) och Ferm (2010) till cirka 90 % förklaringsgrad vilket är bättre än vad den här studien lyckades uppnå. RMSE var också lägre. Vilket var något högt i den här studien. Dock fick Næsset (2002) förklaringsgrad på 80 % på äldre skog vilket är mindre än Sjödens (2010) och RMSE på cirka 20 % vilket är lite mer i linje med vad jag fick fram. Næsset (2007) fick ett ännu bättre RMSE på 11 %.

Den grundtyevägda medeldiametern fick också ganska bra resultat av skattningarna. Förklaringsgraden blev 78 %. 39-78% fick Næsset (2002) som förklaringsgrad på diameter. Där min skattning ligger inom den övre delen av intervallet, vilket var den förklaringsgrad Næsset (2002) fick för yngre skogar.

De helt klart sämsta skattningarna erhöles i denna studie för ålder och stamantal. Förklaringsgraden blev cirka 63 % för ålder och 66 % för stamantal. Den procentuella RMSE och standardavvikelsen var även mycket höga 27 % för ålder och 36 % för stamantal. Vilket ligger i samma klass som Hjelmér (2009) RMSE för validering genom ett nytt data material vilket var på mellan 30-32 %. Ferm fick en förklaringsgrad på cirka 69 % och RMSE på 31 % vilket också ligger i linje med den här studien. Tydligt kan alltså stamantal och ålder skattas relativt dåligt med hjälp av laserskanning.

Alltså uppvisar denna studie liknande mönster som andra liknande studier om laserskanning. Höjder och volymer kan skattas bättre och stamantal och även ålder skattas dåligt. Vilket kan bero på som ovan nämndes att det är höjden som mäts tillsammans med vissa täthetsmått, exempelvis andel vegetationspulser. Därför borde då höjd skattas bra. Volym och biomassa som båda beror på höjd och täthet kan skattas ganska bra, medan stamantal som borde vara mer oberoende av höjd, skattas sämre. Även hos åldern är det förståeligt att det skattas dåligt. Det är ju bara ett indirekt samband mellan ålder och höjd, äldre träd är ju ofta högre än små plantor. I figur 3 ser vi också att spridningen runt linjen ökar med ökad ålder, vilket kan bero på att skogar med olika ståndortsindex får olika maximal höjd (Hägglund & Lundmark 1982).

## 4.2 Laservariabler

Den grundtyvägda medelhöjden skattades med hjälp av en enkel linjärregression, där percentil 90 utnyttjades. Sjödin (2010) använde däremot två percentiler 70 och 95. I denna studie undvek dock jag att använda fler än en percentil i modellerna på grund av att de förklarande variablerna var korrelerade, se även 4.3 Osäkerhet. Men i alla fall ligger min percentil 90 ganska nära hans 95.

En annan typ av variabler användes av Ferm (2010), exempelvis skilde han på olika variabler som berodde på första, mitten, sista och singelreturer och använde olika kombinationer av dem till percentiler och kvoter. De mest signifikanta variablerna för höjd han fick fram var percentilerna 50, 90 och kvoter mellan första/sista och singelreturer genom totalt antal. Han har alltså också använt sig av percentiler liksom jag gjort men även av andra variabler.

Næsset (2007) skattade övre höjd i två strata och i dessa olika strata användes olika laservariabler. Alltså kan olika variabler vara viktiga på olika platser, men som Næsset också skriver är att variablerna bör vara lika om man vill använda det över större områden. För stratum I användes max och ett täthetsmått, andel av första laserpulsen av totalt antal, för stratum II percentil 70 och max. Han har alltså både använt sig av både höga percentiler och ett täthetsmått. Man behöver kanske även täthetsmått i vissa fall för att skatta en del skogar bra. Beroende på hur skogen ser ut kan det kanske ibland finnas ett samband mellan täthet och höjd.

För de andra skattningarna av egenskaper användes även standardavvikelsen och andel vegetationspulser för alla utom för volym där bara percentil 50 och andel vegetationspulser användes. Att det behövs någon form av täthetsmått som andel vegetationspulser för volym, biomassa, diameter och stamantal är ganska logiskt i och med att till exempel volym beskriver en mängd och den mängden beror både på vertikal höjd och hur den är spridd över ytan. Exempelvis använder även Næsset (2007) både höjdmått och täthetsmått vid diameter- och volym-skattning och (2002) använde han kvantil 80 som motsvarande percentil 80 täthetsmått över 80 kvantilen. Han använde även den naturliga logaritmen som transformering för att få ett bättre resultat vilket han har hittat i tidigare studier. Det pekar även min studie på i och med att skattningarna blev bättre efter transformering.

Sjödin (2010) nämner att höjdvariabler var viktiga och även bland annat skanningsvinkel men nämner också att man inte använder den så ofta i skogliga sammanhang på grund av att vinklarna oftast inte är så stora. Även Næsset (2007) tog bort de punkter som hade högre vinkel än 15 grader. Det man kan säga verkar vara viktigt vid olika skattningar är olika percentiler och täthetsmått men att det även finns andra variabler och transformationer som vi sett ovan, man kan utnyttja. Förmodligen är olika variabler bra i olika sammanhang som om det är hög skanningsvinkel kan den variabeln bli viktigare. Men i den här studien har exempelvis skanningsvinkel inte kunnat testats på grund av utgångsdatat.

### 4.3 Osäkerhet

Det finns vissa osäkerheter i datat. Någon provyta fick tas bort ur datamaterialet på grund av orimliga höjdvärden, vilket kan indikera att någon provyta kan ha blandats ihop i fältinventeringen. Det fanns också en del uteliggare som har påverkat skattningarna, men i och med att värdena inte var alltför orimliga togs de ej bort. Tar man bort data utan en bra anledning kan det påverka resultatet och skattningen skulle bli felaktig (Hadi & Chatterjee, 2006).

En sak man får vara uppmärksam på då man jämför dessa resultat med andra är att, de här resultaten är kanske inte helt jämförbara i andra områden med annan skogsstruktur. Mina modeller är förmodligen inte helt giltiga i ett helt annat område med annan skogsstruktur därför att de inte är byggda för just de förhållandena. Olika skogsstruktur kan ge olika modeller (Næsset 2002).

Andra studier pekar på att andra sätt att ta fram modeller kan ge lite olika resultat. Vilket kan medföra att andra metoder hade kunnat ge en bättre eller sämre skattning än vad som blev fallet i denna studie (Hjelmér 2009). I datamaterialet finns också vissa problem med korrelerade förklarande variabler. Så för att undvika de största korrelationerna beslutades att bara ta med en percentil i varje modell. Percentilerna var mycket korrelerade inbördes. Trots det kan viss korrelation förekomma mellan percentil, standardavvikelse och andel vegetationspulser, men de värsta korrelationerna är i alla fall borta. Det kan hända att även jag skulle ha fått högre förklaringsgrad med fler percentiler, men i och med att jag inte kunde lösa det på annat sätt så föredrog jag att minska kollinearitetsproblemet.

Som tidigare nämnts finns viss heteroskedastisitet, strutformighet i diagram, kvar trots transformering vilket medför att skattningarna inte blir lika säkra som de skulle ha varit annars (Hadi & Chatterjee, 2006). Efter återtransformeringen ökade detta problem också.

Det finns alltså kvar vissa problem i modellerna som i viss mån kan påverka skattningarna. Men detta visar också att det är svårt att få till riktigt bra modeller, och det fodrar goda kunskaper i regressionsanalys för att få till bra resultat som uppfyller de statistiska antagandena (Hadi & Chatterjee, 2006).

#### **4.4 Slutsats**

Det verkar som det går ganska bra att skatta egenskaper hos skogen med hjälp av laserdata, men olika egenskaper kan skattas olika bra. Andra studier på området har kommit fram till liknande resultat.

Den här studien visar på att höjd skattas bäst, men att biomassa, diameter och volym också skattas ganska bra men lite sämre. Sämst blev skattningarna för stamantal och ålder. Vilket förmodligen beror på att man bara indirekt mäter egenskaper som är korrelerade med dessa data.

De laservariabler som verkar vara viktigast skiljer sig åt mellan olika egenskaper, studier och områden, men percentiler har en dominerande roll tillsammans med olika kvoter som beskriver täthet av punkter. I framtiden borde det kunna finnas stora möjligheter att laserskanning ska kunna användas för olika typer av inventeringar inom skoglig planering.

## 5 KÄLLFÖRTECKNING

Blom (2008) Projektrapport 1110 Krycklan, (Rapport/Blom,1110 Krycklan). Göteborg:Blom Swe AB/ Top Eye AB

Dubayah R., & Drake J. (2000). Lidar Remote Sensing for Forestry. *Journal of Forestry*, 98:6, 44-46

Eid T., Gobakken T. & Næsset E. (2004). Comparing Stand Inventories for Large Areas Based on Photo-interpretation and Laser Scanning by Means of Cost-plus-loss Analyses. *Scandinavian Journal of Forest Research*, 19:6, 512-523

Ferm M. (2010), *Prediktering av skogliga variabler med data från flygburen laser - En jämförelse mellan multipla regressionsmodeller och k nearest neighbour-modeller*. Umeå, Umeå Universitet, examensarbete

Hadi A. & Chatterjee S. (2006). *Regression analysis by example*. Hoboken: Wiley

Heurekaenheten, SLU (2011) *Field inventory instruction/sv*.(online) Tillgänglig: [http://heureka.resgeom.slu.se/wiki/index.php?title=Field\\_inventory\\_instruction/sv](http://heureka.resgeom.slu.se/wiki/index.php?title=Field_inventory_instruction/sv) (2012-03-13)

Hjelmér M. (2009), *Skattning av antal träd baserat på data från flygburen laserskanning*. Umeå, Umeå universitet, examensarbete

Hyypä J., Kelle O., Lehtikainen M. & Inkinen M. (2001). A segmentation-based method to retrieve stem volume estimates from 3-D tree height models produced by laser scanners. *IEEE Transactions on Geoscience and Remote Sensing*. 39:5 969-975

Hägglund B. & Lundmark J-E. (1982). *Bonitering -Definitioner och anvisningar* Jönköping: Skogsstyrelsen

Lantmäteriet (2011), *Produkt beskrivning: Laserdata Dokumentversion: 1.3* 2011-12-01, Lantmäteriverket

Nelson R., Krabill W. & Tonelli J. (1988) Estimating Forest Biomass and Volume Using Airborne Laser Data. *Remote Sensing of Environment* 24:2 247-267

Næsset E. (1997a) Estimating Timber Volume of Forest Stands Using Airborne Laser Scanner Data. *Remote Sensing of Environment*, 61:2 246–253

Næsset E. (1997b) Determination of mean tree height of forest stands using airborne laser scanner data. *Journal of Photogrammetry & Remote Sensing*, 52:2, 49-56

Næsset E. (2002) Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sensing of Environment* 80:1 88-99



Næsset E. (2007) Airborne laser scanning as a method in operational forest inventory: Status of accuracy assessments accomplished in Scandinavia. *Scandinavian Journal of Forest Research*, 22:5 433-442

Næsset E., Gobakken T., Holmgren J., Hyypä H., Hyypä J., Maltamo M., Nilsson M, Olsson H., Persson Å. & Söderman U. (2004): Laser scanning of forest resources: the nordic experience, *Scandinavian Journal of Forest Research*, 19:6, 482-499

Sjödin M. (2010). *Skattningar i gallringsskog med hjälp av flygburen laserskanning - beräkningar med massaslutenhet*. Umeå, Sveriges lantbruksuniversitet, examensarbete

Wikström P., Edenius L., Elfving B., Eriksson L-O., Lämås T., Sonesson J., Öhman K., Wallerman J., Waller C. & Klintebäck F. (2011) The Heureka Forestry Decision Support System: An Overview. *Mathematical and Computational Forestry & Natural-Resource Sciences*, 3:2, 87-94

## **BILAGA I**

### **SUMMARIUM**

Hoc scriptum tractat quomodo bona, silvae propriees computantur possunt, fortis laser intuens, et regression analysis. In Botnia occidentalis in Suecia, opusculum perficetur. Coefficiens determinatio et RMSE % erat; altitudo 88,6 %, 8 %; biomoles 79,1 %, 23,3 %; diametrus 78,3 %, 13,7%; amplitudo 73,8 %, 29 %; numerus stirpiorum 66,4 %, 36 %; et aetas 63,5 %, 27 %. Laser variables, qui gravissimus erant, percentiles et densites quoti erant.