

# Utvärdering av utbytesprognoser baserade på skördardata hos Sveaskog

*Evaluation of pre-harvest forecasts based on harvester data at  
Sveaskog*

Linnéa Hemmingsson



Examensarbete • 30 hp

Jägmästarprogrammet

Rapport från Institutionen för skogens biomaterial och teknologi, 2020:2

Umeå 2020



# Utvärdering av utbytesprognoser baserade på skördardata hos Sveaskog

*Evaluation of pre-harvest forecasts based on harvester data at Sveaskog*

Linnéa Hemmingsson

**Handledare:** Dan Bergström, Sveriges lantbruksuniversitet, Institutionen för skogens biomaterial och teknologi

**Examinator:** Dimitris Athanassiadis, Sveriges lantbruksuniversitet, Institutionen för skogens biomaterial och teknologi

**Omfattning:** 30 hp

**Nivå och fördjupning:** Avancerad nivå, A2E

**Kurstitel:** Masterarbete i skogsvetenskap

**Kursansvarig inst.:** Institutionen för skogens biomaterial och teknologi

**Kurskod:** EX0956

**Program/utbildning:** Jägmästarprogrammet

**Utgivningsort:** Umeå

**Utgivningsår:** 2020

**Omslagsbild:** Linnéa Hemmingsson

**Serietitel:** Rapport från Institutionen för skogens biomaterial och teknologi

**Delnummer i serien:** 2020:2

**Elektronisk publicering:** <https://stud.epsilon.slu.se>

**Nyckelord:** utbytesberäkning imputering utfall prognos planering

**Sveriges lantbruksuniversitet**

Fakulteten för skogsvetenskap

Institutionen för skogens biomaterial och teknologi



## Sammanfattning

För att på ett effektivt sätt styra ett kundanpassat virkesflöde krävs god information om tillgången av virke. Vid en skoglig åtgärd så som gallring eller föryngringsavverkning görs vanligtvis en skattning av detta i fält, vilken sedan även kan kompletteras med en utbytesberäkning. En ny metod att göra utbytesberäkningar baserade på historiskt skördardata har en kort tid använts inom Sveaskogs verksamhet.

Syftet med detta arbete var att utvärdera denna metod att göra utbytesberäkningar genom att: 1) kartlägga hur metoden har implementerats i Sveaskogs verksamhet, 2) kvantifiera noggrannheten och precisionen i ingångsdata för utbytesberäkningen och 3) kvantifiera noggrannheten och precisionen för utbytesberäkningen.

Arbetet har genomförts genom en workshop med insatta personer för att kunna göra en kartläggning av arbetsprocessen där utbytesprognoser tas fram. Sedan har ett datamaterial innehållande alla åtgärder inom ett givet tidsspänn studerats för att kunna kvantifiera noggrannheten och precisionen. För den skogliga skattningen har även en regressionsanalys genomförts för att undersöka vilka variabler som bidrar till att avvikelser mellan planerat och avverkad volym uppstår.

Resultaten av studien visar att det för de skogliga skattningarna finns vissa systematiska fel, vilka är olika stora för de olika skogliga parametrarna. Framförallt stamantal och grundyta utmärker sig genom att visa både något sämre precision och noggrannhet än övriga parametrar. Avvikelser i volym tycks framförallt påverkas av vilket Resultatområde åtgärden tillhör samt åtgärdens volym per ha. Dock visar den framtagna modellen att det finns fler faktorer som kan påverka avvikelsen än vad som tagits med i studien. Imputeringen visar bra noggrannhet men något sämre precision.

*Nyckelord:* utbytesberäkning, imputering, utfall, prognos, planering

## Abstract

In order to handle a customized wood flow in an efficient way high quality information about wood resources are important. Before harvests an estimation of the stand is usually conducted which then can be complemented with a pre-harvest forecast. A new method to conduct these forecasts based on historical harvester data has been implemented at Sveaskog.

The aim of this study was to evaluate these pre-harvest forecasts by: 1) map the work-process of the forecasts, 2) quantify the precision and accuracy of the forestry measurements that are used as input for the forecast and 3) quantify the precision and accuracy for the forecasts.

The work has been conducted by arranging a workshop with people working with the forecasts in order to do the mapping. Then a set of data containing all the harvests carried out during a given time period was analyzed to be able to quantify the precision and accuracy. For the forest estimations a regression analysis was conducted to see which variables mostly affected the differences between planned and harvested volume.

The results from this study show that there are systematic errors for several of the estimated forest variables. Number of stems and basal area had the lowest precision and accuracy. Deviation between planned and harvest volume seems to mostly be affected by the volume per hectare and region. However, the module only explains some part of the deviation. The volumes predicted with historical harvester data show a good accuracy and a lower precision.

*Keywords:* yield, imputation, result, prediction, planning

# Förord

Detta examensarbete omfattande 30 hp har utförts vid institutionen för skogens biomaterial och teknologi vid Sveriges lantbruksuniversitet.

Jag vill rikta ett stort tack till Christer Ranvald på Sveaskog, Johan Möller och Jon Söderberg på Skogforsk för alla era bidrag som gjorde detta arbete möjligt. Jag vill även rikta ett stort tack till min handledare Dan Bergström för stort engagemang och många goda råd.

Umeå, mars 2020

Linnéa Hemmingsson

# Innehållsförteckning

1	Inledning.....	1
1.1	Om Sveaskog .....	1
1.2	Planering hos ett skogsföretag .....	2
1.2.1	Skoglig planering .....	3
1.3	Utbyteprognoser .....	3
1.3.1	Tidigare metoder för utbytesprognoser.....	3
1.3.2	Utbytesprognoser baserade på skördardata .....	4
1.3.3	Tidigare studier .....	4
1.4	Syfte.....	6
2	Material och metoder.....	7
2.1	Material .....	7
2.1.1	Kartläggning av prognosmetoden hos Sveaskog .....	7
2.1.2	Datamaterial .....	8
2.1.3	Datasammanställning - skogliga variabler .....	9
2.1.4	Datasammanställning - direktutbyte .....	11
2.2	Metoder.....	12
2.2.1	Formler .....	13
2.2.2	Regressionsanalys .....	13
2.2.3	Precision och noggrannhet.....	15
3	Resultat .....	16
3.1	Kartläggning av arbetsprocessen för utbytesprognoser hos Sveaskog .....	16
3.2	Utvärdering av skogligt data .....	20
3.3	Utvärdering av direktutbyte .....	22
4	Diskussion .....	25
4.1	Kartläggning av arbetsprocessen för utbytesprognoser hos Sveaskog .....	25
4.2	Utvärdering av skogligt data .....	26
4.3	Utvärdering av direktutbyte .....	27
4.4	Styrkor och svagheter i studien.....	28
4.5	Implementering och framtida studier.....	29
4.6	Slutsatser .....	30
	Referenslista .....	31



# 1 Inledning

## 1.1 Om Sveaskog

Sveaskog äger 14 % av den svenska skogsmarksarealen och är därmed Sveriges största skogsägare (Sveaskog 2018). Som ett skogsägande företag är deras främsta mål med verksamheten att skapa primärprodukterna timmer, massaved och biobränsle, från den egna skogen (Nilsson *et al.* 2017). För att Sveaskog ska kunna tillfredsställa sina kunders virkesbehov köper de även in virke från externa leverantörer. Dels från privata skogsägare, ofta i form av avverkningsuppdrag, men även genom import och genom centrala köp från andra skogsbolag (Sveaskog 2018).

Sveaskogs skogliga verksamhet delas in i två marknadsområden som i sin tur delas in i resultatområden och produktionsområden (Sveaskog 2018) (Tabell 1).

**Tabell 1.** Geografisk indelning av Sveaskogs skogliga verksamhet i Marknadsområde, Resultatområde och Produktionsområde

**Table 1.** A geographical division of the forestry function at Sveaskog into Market areas (Marknadsområden), Economic areas (Resultatområde) and Production areas (Produktionsområde)

Marknadsområde	Resultatområde	Produktionsområde
Nord	Västerbotten	Jämtland
		Lycksele
		Malå
		Skellefteå
		Storuman
	Norrbotten	Vindeln
		Åsele
		Arvidsjaur
		Boden

Marknadsområde	Resultatområde	Produktionsområde	
Syd	Götaland	Gällivare	
		Piteå	
		Tärendö	
		Älvsbyn	
		Överkalix	
		Berga	
		Jönköping	
		Osby	
		Vimmerby	
		Växjö	
		Södra Bergslagen	Askersund
			Karlskoga
		Norra Bergslagen	Norra Bergslagen
		Södra Norrland	Södra Norrland

## 1.2 Planering hos ett skogsföretag

För ett skogsföretag, likt Sveaskog, kan planeringsprocessen beskrivas genom att dela in de olika planeringsstegen i en hierarki. Detta kan göras utifrån vilken tidshorisont planeringssteget görs på. Planeringshierarkin består av tre delar: strategisk planering, taktisk planering och operativ planering (Söderholm 2002). Den strategiska planeringen innebär planering på lång sikt där den viktigaste delen handlar om att kvantifiera den långsiktiga avverkningspotentialen vilken kan ge den möjliga avverkningen för de kommande 10 åren. Denna information används sedan som ingångsvärde för den taktiska planeringen där avverkningar eller åtgärder (framöver kommer begreppet *åtgärd* användas för en avverkning, andra vanliga begrepp är trakt eller objekt) väljs ut för att sedan planeras i fält (Nilsson *et al.* 2012). Den skogliga planeringen i fält kommer att beskrivas mer i detalj längre fram. De planerade åtgärderna samlas i en traktbank motsvarande 1,5-2 års avverkningsvolym. Traktbanken används som underlag vid den operativa planeringen vilken görs av avverkningsledaren (Nilsson *et al.* 2017). Under detta planeringssteg schemaläggs när åtgärderna ska göras och av vilka avverkningslag (Nilsson *et al.* 2012). Schemaläggningen görs så att avverkningsvolymerna matchar produktionsmålen, vilka består av leveransmål till olika industrierna och den tänkta lagerförändringen (Ranvald 2019). För att matchningen av avverkningsvolym mot produktionsmål ska kunna göras så korrekt som möjligt krävs goda skattningar av utfallet från åtgärden (Malinen *et al.* 2018). I många fall har man dock uppfattningen att en god

beskrivning av skogen är för dyr och man saknar dessutom uppfattning om vilken ekonomisk förlust ett dåligt dataunderlag ger (Kangas 2010).

### 1.2.1 Skoglig planering

Den skogliga planeraren ansvarar för att skapa avverkningsåtgärden genom att avgränsa åtgärden i fält och sedan inventera skoglig data (Nilsson *et al.* 2017). Det skogliga data som inventeras består av grundyta, trädslagsblandning, grundytavägd medelhöjd (Hgv), grundytavägd medeldiameter (Dgv) och stamantal. Med hjälp av dessa variabler räknas en volym fram för åtgärden. När åtgärden avgränsas får den en areal och denna kan användas för att beräkna volym/ha. Skattningarna av de olika variablerna görs genom subjektiva inventeringsmetoder där både mätningar på subjektivt utlagda provtytor och visuella bedömningar görs (Ranvald 2019).

Tidigare studier har undersökt kvaliteten i det data som samlas in genom subjektiva metoder i fält. Kangas *et al.* (2004) visade att de variabler som var svårast att skatta var grundyta och stamantal när man såg till både standardavvikelse och bias (systematiskt fel). I studien konstaterades även att skattningarnas avvikelse till stor del påverkas av vem som utför dem.

## 1.3 Utbyteprognoser

I Sverige används huvudsakligen ett kortvirkessystem (cut-to-length) för avverkning och transport av virke (Gellerstedt & Dahlin 1999). Ett sådant kundanpassat virkesflöde kräver god information om tillgängliga avverkningar för att kunna ta beslut om vilka trakter som ska avverkas och av vilka maskinlag (Malinen *et al.* 2018). Denna information består bland annat av en prognos på det utfall, volymer per sortiment, man kan förvänta sig vid en avverkning (Malinen *et al.* 2014). För Sveaskog, som ett stort skogsägande företag, kan det vara av stort värde att ha en god beskrivning av sin skogstillgång (Kangas 2010).

### 1.3.1 Tidigare metoder för utbytesprognoser

Under 1960-talet arbetade Cernold (1971) fram tabeller för utbytesberäkningar av stående skog. Med hjälp av utbytesfunktioner har man under lång tid kunnat göra prognoser för utfallet från en avverkning. Dessa bygger ofta

på mätningar av provträd och erfarenhetstal. Ett exempel på dessa utbytesfunktioner är *Ollas utbytesfunktioner* (Ståhl & Wilhelmsson 2009).

### 1.3.2 Utbytesprognoser baserade på skördardata

Skogforsk har tillsammans med Södra, Sveaskog och SCA tagit fram ett nytt system för att göra utbytesprognoser där man kan skatta utfallet från avverkningar. Detta system bygger på att man med skogliga data, så som grundyta, höjd och brösthöjdsdiameter som ingångsvärden, kan prognostisera utfallet genom imputering. Med imputering menas att man med olika metoder fyller i saknat provytedata. I detta fall är det skördardata från tidigare avverkningar som används för att fylla luckorna. Detta innebär att den skogliga data som samlats in för avverkningen kompletteras med information om utfall per sortiment från tidigare avverkningar som gjorts i närheten. Imputeringsmetoden som används är kMSN (Most Similar Neighbour). Utgångspunkten vid imputeringen är att bestånd som har liknande beståndsegenskaper bör få ett liknande utfall. Vid imputeringen får objektet som ska avverkas samma egenskaper som tidigare avverkade objekt i närheten hade baserade på skördardata från avverkningen. Härifrån kan sedan en beräkning av vilka stockar och sortiment som kan förväntas vid en avverkning göras vilket resulterar i en utbytesprognos (Möller *et al.* 2017).

Under hösten 2018 började Sveaskog att använda sig av utbytesprognoser baserade på skördardata och en utvärdering av hur väl dessa fungerar har tidigare inte genomförts.

### 1.3.3 Tidigare studier

Tidigare studier för att utvärdera utbytesprognoserna har genomförts av Skogforsk och resultaten från dessa studier visade att det krävs en korrekt trädslagsfördelning i indata för att få prognoser med hög precision på sortimentsutfallen. Man kunde även konstatera att precisionen hos prognoserna bygger främst på kvalitén i ingående beståndsdata (Söderberg *et al.* 2017). Det krävdes även referensytor med liknande trädslagsfördelning som för beståndet som skulle beräknas för att matchningen skulle kunna göras mot ett bestånd med liknande egenskaper (Söderberg *et al.* 2018).

Genom att utföra utbytesberäkningarna utan felaktigheter i indata kunde man konstatera att det inte fanns några signifikanta systematiska fel i skattningarna (Söderberg *et al.* 2018) (Tabell 2). En viktig anledning till att det

ändå uppkom avvikande resultat mellan prognostiserat och faktiskt utfall var att den planerade arealen inte alltid överensstämde med den avverkade. Detta kunde till exempel bero på att bärigheten på trakten var sämre än förväntat och att vissa delar därför lämnades. Förutom att detta påverkade den totala volymen kan detta även förändra trädslagsfördelningen. Genom att lämna de blöta partierna som troligtvis innehöll mer gran så kommer andelen gran som avverkas vara lägre än vad man skattade vid planeringen (Söderberg *et al.* 2017).

**Tabell 2.** Standardavvikelse (SD), medelfel (RMSE) och systematiska fel (Bias) har beräknats på utbytesprognoser gjorda på fem beräkningsytor i slutavverkning och för olika produktgrupper. Beräkningen har gjorts på skillnaden mellan prognos och det skördarmätta utfallet i volym (m<sup>3</sup>fub/ha). Resultatet är en utvärdering genomförd av Söderberg *et al.* (2018)  
**Table 2.** Standard deviation (SD), root mean square error (RMSE) and Bias has been calculated for pre-harvest forecasts done at five homogeneous areas in final harvest and for different assortment groups (Produktgrupp). The calculations were made on the difference between the forecast and the harvester measured volume (m<sup>3</sup>fub/ha). The results are an evaluation done by Söderberg *et al.* (2018)

Produktgrupp	SD	RMSE	Bias
Normaltimmer	17,9	17,9	1,4
Kubb/klentimmer	7,9	7,9	0,3
Massaved	12,6	12,6	-1,3
Massaved inkl. brännved	14,1	14,2	-1,5

De tidigare utvärderingar som har genomförts har framförallt gjort utvärderingen på prognosyta (en mindre, homogen, yta som tillsammans med andra prognosytor utgör ett avverkningsobjekt). Denna studie ska istället göras åtgärdsnivå.

Metoden att använda imputering för att göra utbytesprognoser är inte unik för Sverige. Även i Finland har man jobbat med att göra prognoser med denna metod. I en studie gjord av Malinen *et al.* (2018) har man gjort en utvärdering på utbytesprognoser på sortimentsnivå, dvs. man har kopplat prognoserna till en stambank och en apteringsinstruktion, därigenom har man kunnat göra en teoretisk aptering. Utvärderingen har sedan gjorts på ett simulerat bestånd genom att jämföra det prognostiserade värdet för avverkningsobjektet och det uppmätta värdet. Utvärderingen resulterade i ett RMSE på 132,84 € och bias på -16,56 € för tallbestånd och för granbestånd ett

RMSE på 54,90 € och bias på -7,67 €. Vilket kan jämföras med att medelvärdet för tallbestånden var 1500,73 € och för granbestånden 2564,38 €.

Hur utbytesprognoserna kan göras har beskrivits i Möller et al. (2017) men hur arbetsprocessen ser ut då den implementeras hos ett skogsbolag kan skilja något. Dessutom bör man kartlägga hur man i Sveaskogs verksamhet jobbar med utbytesprognoserna för att få full förståelse för resultatet av utvärderingen av precisionen hos utbytesprognoserna.

Att kunna göra utbytesprognoser med hög noggrannhet och precision kan vara värdefullt både för den operativa planeringen av avverkningar men även för senare steg i försörjningskedjan (Holappa Jonsson 2018).

En utbytesprognos som görs för de övergripande sortimenten normaltimmer, massaved, energived och klintimmer kommer framöver benämnas direktutbyte

## 1.4 Syfte

Syftet med detta arbete var att utvärdera noggrannheten och precisionen av utbytesprognoserna baserade på skördardata hos Sveaskog. Vilket kan delas in i följande delsyften;

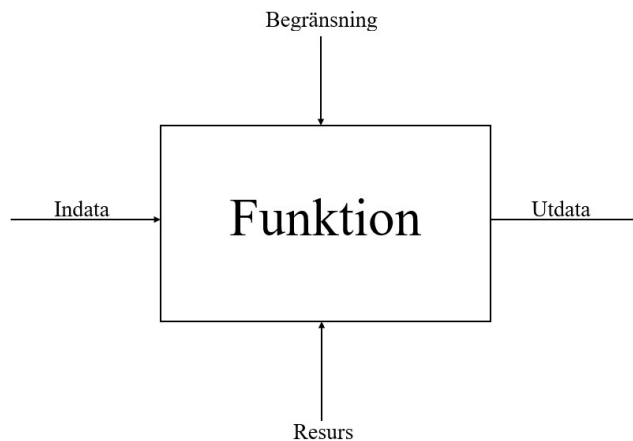
1. Kartlägga arbetsprocessen där skördardata används för utbytesprognoser med fokus på hur dessa används i Sveaskogs verksamhet.
2. Kvantifiera noggrannheten och precisionen för det skogliga ingångsdata för direktutbytet samt undersöka om det finns något samband mellan beståndsdata och avvikelser i volym.
3. Kvantifiera noggrannheten och precisionen för direktutbytet.

## 2 Material och metoder

### 2.1 Material

#### 2.1.1 Kartläggning av prognosmetoden hos Sveaskog

Till att börja med gjordes en kartläggning av arbetsprocessen där utbytesprognoser tas fram med hjälp av skördardata. Denna kartläggning gjordes dels genom att studera litteratur som beskriver utbytesprognoserna samt genom att anordna en workshop där insatta personer från både Skogforsk och Sveaskog deltog. Resultatet kunde sedan sammanställas i en processkarta enligt IDEF0-modellen. Med hjälp av IDEF0-metoden kan man på ett enkelt och visuellt sätt beskriva de funktioner och det informationsflöde som finns inom en process. (Ang *et al.* 1997). Modellen innehåller ett antal funktioner (boxar) och dessa sammankopplas genom olika pilar som symboliserar det informationsutbyte som finns mellan funktionerna (Figur 1). Att beskriva arbetsprocessen med hjälp av denna modell valdes för att kunna visa vilket informationsflöde som finns mellan de olika funktionerna samt kunna visa hur informationen används i den nästkommande funktionen. För att begränsa kartläggningen bestämdes det att den endast skulle begränsas till just den del av processen som behandlar utbytesprognoser. På grund av denna begränsning har en viss modifikation av IDEF0-modell behövts göras då en del av informationen som mottas och skickas kommer från delar som inte tagits med i kartläggningen. Processkartläggningen genomfördes för att få en grundförståelse för utbytesprognoserna, innan de kommande analyserna genomfördes. Resultatet från kartläggningen ska även kunna användas som ett exempel för hur utbytesprognoser baserade på skördardata kan implementeras i en verksamhet.



**Figur 1.** Figuren visar en box (funktion) med tillhörande pilar som bygger på IDEF0-modellen. Funktionen illustrerar en del av en arbetsprocess och pilarna illustrerar hur information kan rör sig mellan funktioner. Översatt från Ang et al. (1997).

**Figure 1.** The figure shows a box (function) with arrows based on the IDEF0-model. The function illustrates a part of a work process and the arrows illustrates how information can flow between different functions. Translated from Ang et al. (1997).

### 2.1.2 Datamaterial

För denna studie användes två separata datamaterial, båda bestående av två delar. Datamaterialen hämtades ut Sveaskogs datorsystem. Det första datamaterialet var en beskrivning av de skogliga variablerna för förnygringsavverkningar genomförda på Sveaskogs mark mellan oktober 2018 och maj 2019. Den första delen av detta material var de skattningar som den skogliga planeraren gjort för variablerna grundyta, Dgv, Hgv, stamantal, areal och volym. Den andra delen av detta material innehöll skördarmätta värden för samma variabler och för samma åtgärder. I materialet fanns även trädslagsfördelning samt koordinater för åtgärden. Under rubrik 2.1.3 beskrivs det utförligare hur båda dessa delar omarbetades för att sedan kunna jämföras.

I det andra datamaterialet var den första delen uppgifter om volymsutfall per produkt. Värdena för volymsutfall i denna del bestod av direktutbytet som prognostiserats innan avverkning för ett antal åtgärder. Denna prognostisering gjordes innan värdena hämtades ur datasystemen och enligt den process som beskrivs under rubrik 3.1 i resultatdelen. Den andra delen i detta datamaterial bestod av det skördarmätta volymsutfallet för samma åtgärder



som finns med i den första delen. Detta datamaterial innehöll både föryngringsavverkningar och gallringar utförda mellan oktober 2018 och maj 2019. Under rubrik 2.1.4 beskrivs det utförligare hur båda dessa delar omarbetades för att sedan kunna jämföras.

### 2.1.3 Datasammanställning - skogliga variabler

För att kunna genomföra analyser av datamaterialet krävdes en del omarbetningar. Innan några omarbetningar gjordes innehöll datamaterialet för de skogliga skattningarna 4 270 rader och 1 709 åtgärder. Datamaterialet från skördarmätningen innehöll 3 617 rader och 913 åtgärder. Eftersom de skogliga skattningarna var en total för beståndet och inte ett förväntat uttag så sorterades endast föryngringsavverkningar ut från detta material. Gallringar kunde inte tas med då skördaren endast mätt de träd som tas ut. Först sammanställdes variablerna som en total per åtgärd. För volym, grundyta och stamantal kunde detta göras genom att endast summera. Detta gjordes i Microsoft Access. Arealen fanns redan redovisad som en total per åtgärd. För att kunna ta fram Dgv och Hgv för åtgärden räknades det först fram en trädslagsblandning baserad på grundyta. Denna trädslagsblandning kunde sedan användas för att ta fram ett viktat medelvärde för både Dgv och Hgv per åtgärd. För dessa kompletterande beräkningar användes Microsoft Excel.

Sammanställningarna gjordes på samma sätt för både de skogliga skattningarna och det skördarmätta referensmaterialet och de båda datamaterialen kunde sedan sammanfogas på åtgärds id. När sammanfogningen gjordes togs endast åtgärder som fanns med i båda dataseten med. Detta resulterade i att 376 unika åtgärder fanns kvar för vidare studier. Då hade även de åtgärder som saknade ett unikt åtgärds id tagits bort.

För varje åtgärd fanns även koordinater angivna som kunde läggas till i ett kartobjekt över Sveaskogs geografiska verksamhetsindelningar. I ArcMap kunde åtgärderna sammanfogas med de geografiska områdena och på så vis tilldelades varje åtgärd en geografisk region. Detta gjordes för att kunna studera om det finns några samband mellan noggrannheten och precisionen samt regionstillhörighet (Tabell 3). I detta fall användes Resultatområde som

den geografiska indelning som gjordes av åtgärderna. Resultatområden valdes för att få en så noggrann indelning som möjligt men ändå få ett betydande antal åtgärder per område.

**Tabell 3.** Antalet åtgärder i studien per respektive resultatområde inom Sveaskog

*Table 3. The number of harvests in the study distributed to the economic areas (resultatområde) at Sveaskog*

Resultatområde	Antal åtgärder
Götaland	89
Norra Bergslagen	60
Norrbottn	41
Södra Bergslagen	56
Södra Norrland	36
Västerbotten	94

Trädslagsblandningen kunde även användas för att dela in åtgärderna i olika beståndstyper (Tabell 4).

**Tabell 4.** I studien har avverkningsåtgärder delats in i olika beståndstyper utifrån vilken trädslagsfördelning åtgärden har haft. Trädslagsfördelningen är framräknad utifrån hur stor andel av den totala grundytan respektive trädslag utgör, baserat på skördarmätt grundytta. Beståndstyperna är definierade enligt Riksskogstaxeringen (2019) definitioner

*Table 4. In the study the harvesting operations have been divided into different forest types based on the species distribution of the operation. The species distribution is based on what portion of the total basal area the species has, calculated from the harvester measured basal area. The forest types are based on Riksskogstaxeringen (2019) definition*

Beståndstyp	Trädslagsfördelning	Antal åtgärder
Talldominerad	Tallandel >65%	118
Grandominerad	Granandel >65%	111
Björkdominerad	Björkandel >65%	1
Lövdominerad	Övrigt lövandel >65%	0
Contortadominerad	Contortaandel >65%	1
Blandbestånd	Samtliga trädslag <65%	145

När sammanställningen var klar räknades avvikelserna mellan de skattade och skördarmätta värdena fram genom att beräkna skillnaden mellan de skattade värdena och det skördarmätta utfallet. En relativ avvikelse beräknades med ekvation 1. Detta gjordes för variablerna grundyta, grundytavägd medeldiameter (Dgv), grundytavägd medelhöjd (Hgv), areal, stamantal och volym.

#### 2.1.4 Datasammanställning - direktutbyte

För att kunna genomföra analyser av resultaten från direktutbytet krävdes vissa omarbetningar av datamaterialet. Innan omarbetningarna gjordes bestod materialet av 42 164 rader och 7 948 åtgärder. I ursprungsmaterialet fanns volymer uttryckt i m<sup>3</sup> per produkt och åtgärd. Det första som gjordes var att indela produkterna i produktgrupper (Tabell 5) och även här togs åtgärder som saknade ett unikt åtgärds id bort. Tidigare utvärderingar av utbytesprognoserna visar att skillnad mellan prognostiserat utfall och skördat utfall kan härledas till att arealen som avverkas inte är den samma som den planerade. Därför togs den planerade arealen med och volymen per hektar räknades fram för samtliga produktgrupper. Samma sammanställning gjordes för motsvarande datamaterial innehållande de skördarmätta volymerna. Här användes den avverkade arealen för att räkna fram volym per hektar och dessa båda datamaterial kunde sedan sammanfogas i Microsoft Access för att jämförelser och analyser skulle kunna genomföras. Sammanfogningen gjordes så att endast de åtgärder som fanns med i båda datamaterialen togs med och de åtgärder som inte hade ett unikt åtgärds id togs bort, vilket resulterade i att det slutgiltiga datamaterialet bestod av 1 703 unika åtgärder (tabell 5).

**Tabell 5.** De produktgrupper som använts i studien är en gruppering av olika produkter (sortiment) utifrån de definitioner som används av Sveaskog och antalet åtgärder per produktgrupp

*Table 5. A grouping of the assortments based on the groups used by Sveaskog has been done and the number of harvests in every assortment group*

Produktgrupp	Produkter	Antal åtgärder
Energived	Energived	691
Klentimmer	Klentimmer tall	1528
	Klentimmer gran	
Massaved	Granmassaved	1703
	Barrmassaved	
	Lövmassaved	
Normaltimmer	Lövtimmer	1669
	Lärktimmer	
	Grantimmer	
	Talltimmer	

## 2.2 Metoder

Analysen av datamaterialet gjordes dels genom att beräkna relativ avvikelse (ekvation 1), medelfel (RMSE) (ekvation 2), absolut medelfel (MAE) (ekvation 3), standardavvikelse (SD) (ekvation 4) och systematiska fel (bias) (ekvation 5). Relativ avvikelse räknas fram för att kunna jämföra variabler med olika enhet med varandra samt för att kunna jämföra produktgrupper trots eventuella skillnader i volym. RMSE är standardavvikelsen för residualerna, dvs. skillnaden mellan det skattade värdet och det faktiska värdet och ger därför ett mått på noggrannheten för prognosmetoden. Det samma gäller för MAE som ger medelvärdet för den absoluta avvikelsen mellan prognosen och utfallet. SD är istället ett mått på hur mycket värdena avviker från medelvärdet och ger även det ett mått på spridningen och därmed precisionen. Om de olika värdena ligger nära medelvärdet (liten spridning) blir standardavvikelsen låg.

Bias ger avvikelseernas medelvärde och ger även det ett mått på precisionen hos skattningen. Ett medianvärde för skillnaden mellan prognos (imputering eller skogligt data) och utfallet togs också fram för att kunna se prognosernas noggrannhet. Medianvärden togs med då datamaterialet innehöll en del åtgärder med väldigt stora avvikelser och genom att använda medianvärdet får dessa extremvärden inte för stor vikt.

### 2.2.1 Formler

$$\text{Relativ avvikelse} = \frac{(x_i - y_i)}{y_i} \quad (1.)$$

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

$$\text{MAE} = \frac{\sum_{i=1}^n |x_i - y_i|}{n} \quad (3.)$$

$$\text{SD} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}} \quad (4.)$$

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

Där  $x_i$  är det prognostiserade värdet för åtgärd  $i$ ,  $y_i$  är det observerade värdet för åtgärd  $i$ ,  $\bar{x}$  är medelvärdet för de prognostiserade värdena och  $n$  är antalet observationer

### 2.2.2 Regressionsanalys

En regressionsanalys gjordes för att undersöka om det finns något samband mellan avvikelser i volym och beståndsdata. Detta gjordes i Minitab 18. Det som undersöktes var om det fanns något samband mellan responsen, relativ avvikelsevolym, och de förklarande variablerna grundyta, Dgv, Hgv, stamantal, areal och volym per hektar samt de kategoriska variablerna Beståndstyp och Resultatområde. I detta fall används de skördarmätta värdena för variablerna. Vid undersökningen var nollhypotesen att det inte finns något samband mellan den relativa avvikelsevolymen och de olika förklarande variablerna. Att ta med samtliga variabler i regressionsanalysen kan skapa problem. Det ena problemet är att modellen kan bli onödigt komplicerad och det andra problemet är multikollinearitet, då två eller flera variabler förklarar ungefär samma sak, dvs. är korrelerade. Vid hög korrelation mellan variabler bör endast en av dem tas med då multikollinearitet ökar osäkerheten i skattningarna (Körner & Wahlgren 2005). Det första som undersöktes var om det fanns någon korrelation mellan de olika förklarande variablerna och en korrelationsmatris togs fram. På grund av att det fanns starka korrelationer mellan flera av de ingående variablerna gjordes regressionsanalysen som en stegvis regression (Stepwise regression) i Minitab (Tabell 6). Vid stegvis regression i Minitab låter man programmet själv lägga till och plocka bort

variabler för att få fram en signifikant modell och detta görs utifrån ett förprogrammerat alfa-värde. Alfa-värdet är det värde som modellen matchar p-värdet med. Modellen stannar när p-värdet för alla ingående variabler är lägre än alfa-värdet (Minitab 2019). För denna analys användes alfa-värdet 0,10.

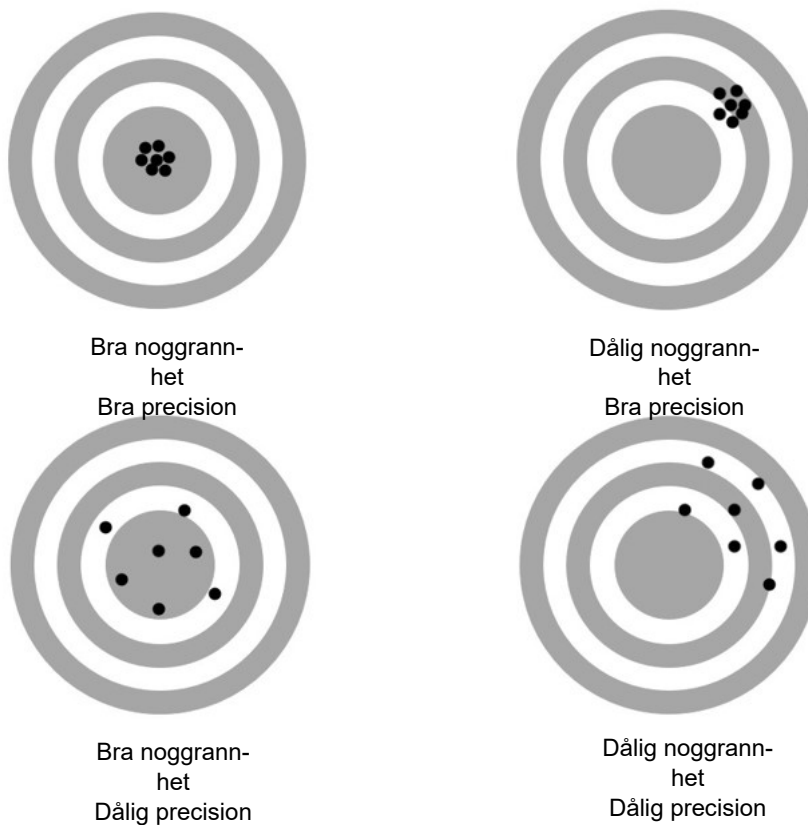
**Tabell 6.** Resultatet av ett test av korrelationen mellan alla de ingående kontinuerliga variablerna, grundyta, grundytavägd medeldiameter (Dgv), grundytavägd medelhöjd (Hgv), areal, stamantal och volym. Det övre värdet är Pearsons korrelationskoefficient och det under värdet är korrelationskoefficientens p-värde

**Table 6.** The results from test of correlation of the different continuous variables, basal area, basal area weighted mean diameter, basal area weighted mean height, area, number of stems and volume. The upper value is Pearsons correlation coefficient and the value underneath is the p-value of the correlation coefficient

	Grundyta	Dgv	Hgv	Areal	Stamantal
<b>Dgv</b>	0,620 <0,001				
<b>Hgv</b>	0,744 <0,001	0,916 <0,001			
<b>Areal</b>	-0,260 <0,001	-0,185 <0,00	-0,268 <0,00		
<b>Stamantal</b>	0,102 0,049	-0,639 <0,001	-0,487 <0,001	0,015 0,768	
<b>Volym</b>	0,967 <0,001	0,732 <0,001	0,868 <0,001	-0,271 <0,001	-0,091 0,078

### 2.2.3 Precision och noggrannhet

För både imputeringen och de skogliga skattningarna ska både precision och noggrannhet studerats (Figur 2). Precisionen har i detta arbete definierats som skattningarnas tillförlitlighet, spridningen i mätresultat. Noggrannhet innebär skillnaden mellan det skattade värdet och referensvärdet.



**Figur 2.** Figuren är en illustration över hur precision och noggrannhet har definierats i denna studie.

**Figure 2.** The figure is an illustration on how precision and accuracy has been defined in this study.

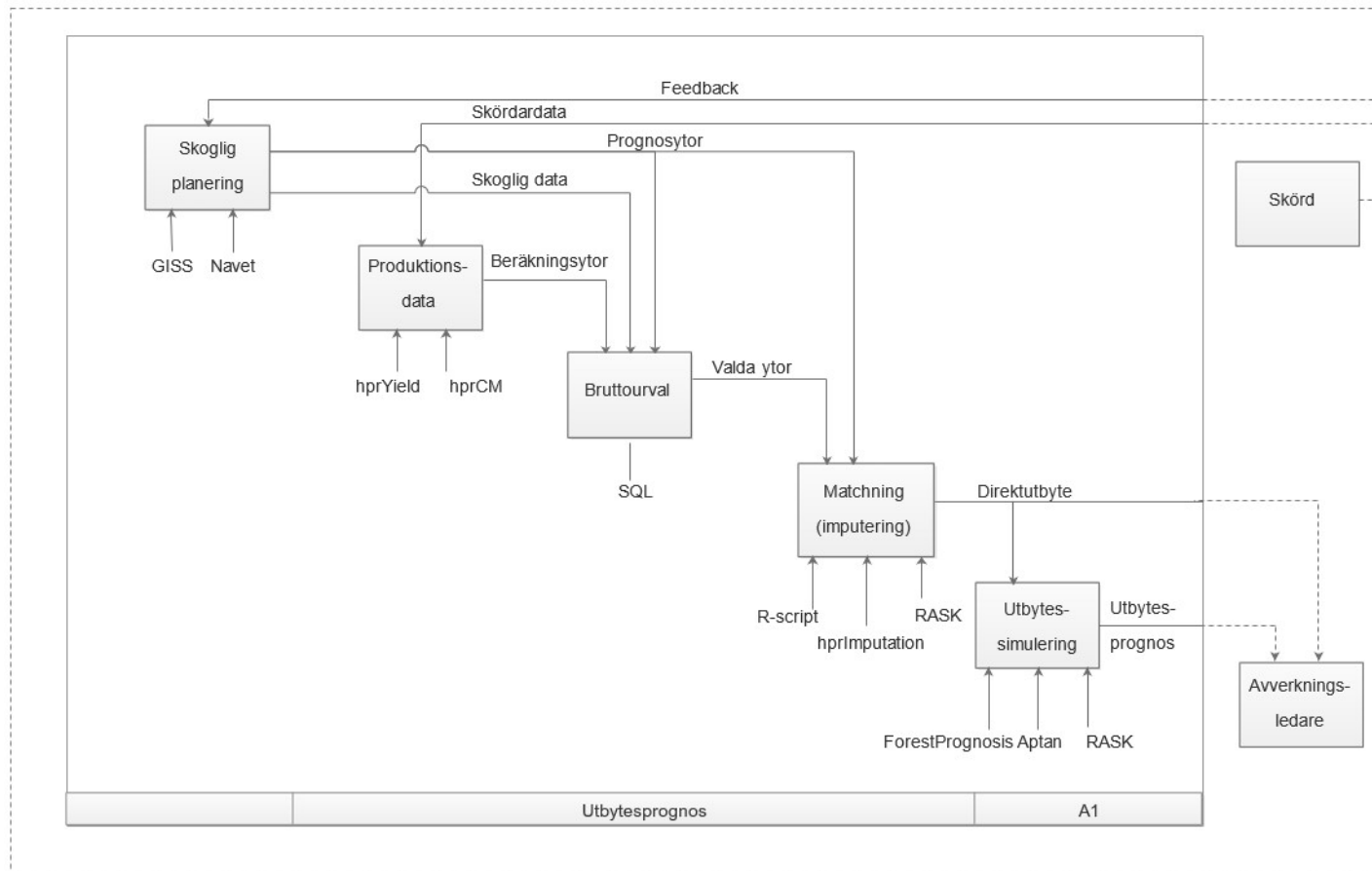
## 3 Resultat

### 3.1 Kartläggning av arbetsprocessen för utbytesprognoser hos Sveaskog

Kartläggningen av arbetsprocessen för utbytesprognoser hos Sveaskog resulterade i att en indelning i fem olika funktioner kunde göras enligt Figur 1. För de flesta funktioner sker informationsutbytet inom den huvudfunktion som kallas för "Utbytesprognos" men ett visst informationsutbyte görs även med funktioner som ligger utanför den del som kartlagts. "Direktutbyte" och "Skördardata" är exempel på sådan information (Figur 3). Mottagaren eller avsändaren av denna information har illustrerats i Figur 3 utan att specificera vilken övriga information som når dessa funktioner eller vilken plats de i övrigt har i processen. Vissa av funktionerna görs automatiskt i datasystem som används av Sveaskog, andra funktioner kräver att någon aktivt sätter igång funktionen (Tabell 7).

Funktionerna kopplas samman genom olika varianter av informationsutbyte samt de system som används för funktionerna. Systemen som används är både Sveaskogs egna, exempelvis GISS, Navet och RASK, men även olika beräkningsmoduler och programmeringsspråk används (Tabell 8).





**Figur 3.** Flödesschema över hur arbetsprocessen för att ta fram utbytesprognoser och direktutbyte ser ut hos Sveaskog. Lådor symboliserar funktioner och pilar symboliserar informationsutbyte mellan funktionerna. Den yttre, heldragna rutan symboliserar den övergripande funktionen "Utbytesprognos" som kartläggningen har genomförts på. Den yttre streckade linjen fångar upp de funktioner som ligger utanför "Utbytesprognoser" men som ändå har ett informationsutbyte med de ingående funktionerna. Även de streckade pilarna symboliserar det informationsutbyte som sker utanför "Utbytesprognos". Förklaring av ingående funktioner och pilar i tabell 7 & 8.

**Figure 3.** A flowchart of the work process to make pre-harvest forecast at Sveaskog. The boxes are illustrating different functions and the arrows are illustrating how information can flow between different functions. The outer, solid, box represents the overall forecasting function which the mapping has focused on. The outer, dotted, line captures the functions that are not included in the mapping but still have an exchange of information with functions within the mapping. Also, the dotted arrows represent exchange of information outside the forecasting box. Further explanations in table 7 & 8.

**Tabell 7.** Sammanställning över de funktioner som identifierats vid processkartläggningen av utbytesprognoser baserade på skördardata hos Sveaskog

**Table 7.** *Compilation of the functions identified during the mapping of pre-harvest forecasts based on harvester data at Sveaskog*

<b>Funktion</b>	<b>Beskrivning</b>	<b>Sker automa- tiskt?</b>
Skoglig planering	Planering av avverkningsåtgärd. Åtgärden märks ut i fält och hänsynsytor märks ut. Skoglig data samlas in. Åtgärden ritas in i GISS (GIS-system), skoglig data förs också in i GISS.	Nej
Produktionsdata	Skördardata från tidigare avverkningar sammanställs i en databas. Från skördardata kan beräkningsytor skapas.	Ja
Bruttourval	Ett urval av beräkningsytor görs. De 1000 geografiskt närmaste beräkningsytorna väljs ut.	Ja
Matchning (imputering)	De ytor som väljs ut vid bruttourvalet matchas sedan med prognosytorna för den aktuella åtgärden. Matchningen görs utifrån trädslagsblandning, grundyta, grundytvägd medelhöjd, grundytvägd medeldiameter och koordinat. Resultatet blir ett direktutbyte.	Ja
Utbytessimulering	Detta steg görs då användaren aktivt väljer det. Här kopplas direktutbytet mot en stambank och en teoretisk aptering görs. Detta resulterar i en utbytesprognos	Nej
Skörd	Har ej ingått i kartläggningen. Placeringen av lådan behöver inte stämma överens med hur informationsflödet ser ut i verkligheten	
Avverkningsledare	Har ej ingått i kartläggningen. Placeringen av lådan behöver inte stämma överens med hur informationsflödet ser ut i verkligheten	

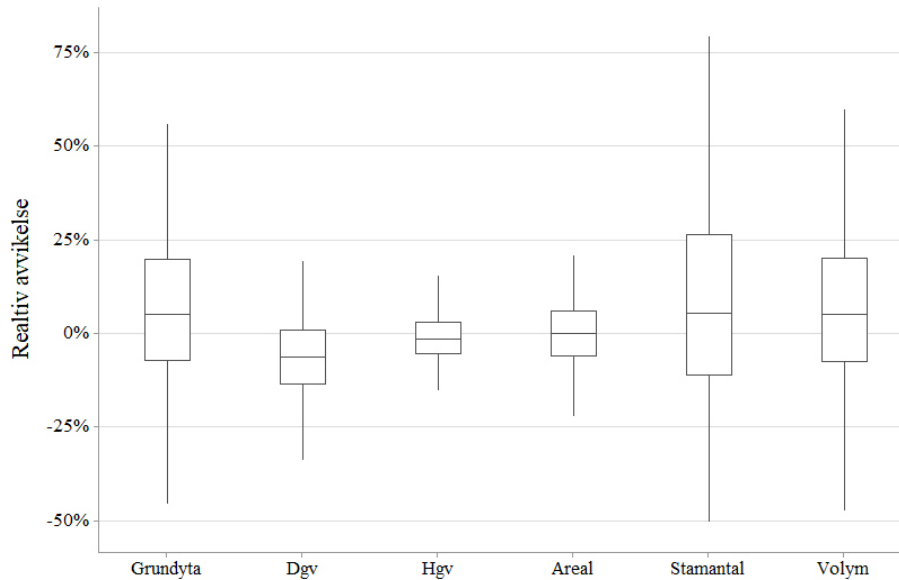
**Tabell 8.** Sammanställningen över det informationsutbyte som identifierats vid processkartläggningen av utbytesprognoser baserade på skördardata hos Sveaskog

**Table 8.** Compilation of the information exchange that could be identified during the mapping of pre-harvest forecasts based on harvest data at Sveaskog

<b>Namn</b>	<b>Beskrivning</b>
Skördardata	Data på trädnivå som samlas in för samtliga träd som skördas
Prognosytor	En åtgärd byggs upp av en eller flera prognosytor. En prognosyta är en homogen yta för vilken utbytet ska beräknas
Skoglig data	Information om skogliga parametrar, oftast insamlat i fält men kan också baseras på laserdata. Parametrarna är trädslagsblandning, grundyta, hgv, dgv och koordinat
Feedback (finns inte ännu)	Återkoppling om hur väl den skogliga planeringen stämmer med resultaten efter avverkning. Ska finnas med i framtiden.
Beräkningsytor	En yta på 0,7-2,0 ha som med hjälp av hprYield skapas utifrån skördardata
Valda ytor	De 1000 geografiskt närmaste ytorna som valts ut bland beräkningsytorna
Direktutbyte	Åtgärdens eller arbetsorderns beräknade produktutfall
Utbytesprognos	Ett resultat av en teoretisk aptering utifrån en befintlig stambank
GISS	Sveaskogs GIS-system som innehåller data om den egna skogen och där den skogliga planeraren bland annat kan skapa en åtgärd
Navet	En databas som hanterar alla Sveaskogs skogliga åtgärder och som kopplar samman information från olika system som används inom verksamheten
hprYield	En beräkningsmodul som används för att analysera och dela in skördardata i homogena beräkningsytor
hprCM	En beräkningsmodul som kan användas för att beräkna trädhöjder
hprImputation	En R-script instruktion för att skapa datamodeller och som i statistikpaketet R anropar imputeringsfunktioner
SQL	Ett programmeringsspråk som används i relationsdatabaser
R-script	Det programmeringsspråk som används vid imputering
RASK	Ett verktyg för virkesflöden på Sveaskog. Kan användas för planering, utförande, lager och uppföljning av flödet. I RASK visas den planerade och imputerade volymen för en åtgärd för användaren
ForestPrognosis	En modul som genererar en stambank på information från hprYield
Aptan	Ett program för apteringssimulering från Skogforsk

### 3.2 Utvärdering av skogligt data

Här kan man se att det finns stora variationer mellan de olika variablerna. Där framförallt grunddyta, stamantal och volym utmärker sig med både sämre precision och noggrannhet än de övriga (Figur 4).



**Figur 4.** Lådagram över de relativa avvikelserna för grunddyta [m<sup>2</sup>], grundtytvägd medeldiameter (Dgv) [cm], grundtytvägd medelhöjd (Hgv) [m], areal [ha], stamantal [st] och volym [m<sup>3</sup>sk/ha] där strecket i lådan är medianvärdet.

**Figure 4.** A boxplot that shows the relative deviations for basal area [m<sup>2</sup>], basal area weighted mean diameter (Dgv) [cm], basal area weighted mean height (Hgv) [m], area [ha], number of stems and volume [m<sup>3</sup>sk/ha]. The median is shown by the line inside the box.

Bias, som i detta fall också skulle kunna beskrivas som medelvärdet för avvikelserna, visar att skattningarna för flera parametrar ligger nära noll. Stamantal och volym visar dock på något större avvikelser. Dessa beror till viss del på att extremvärden förekommer, vilket kan ses genom att jämföra värdet för bias med medianvärdet (Tabell 9).

**Tabell 9.** Medelfel (RMSE), absolut medelfel (MAE), standardavvikelse (SD) och systematiskt fel (Bias) beräknat på de skogliga skattningarna. Medianvärdet är beräknat på skillnaden mellan de skogliga skattningarna och det skördarmätta utfallet vid avverkning

**Table 9.** Root mean square error (RMSE), mean absolute error (MAE), standard deviation (SD) and Bias based on forest estimates and the result measured by the harvester

	RMSE	MAE	SD	Bias	Median
<b>Grundyta [m<sup>2</sup>]</b>	7,43	4,65	10,04	1,90	1,27
<b>Hgv [m]</b>	1,38	1,07	3,41	-0,25	-0,30
<b>Dgv [cm]</b>	3,24	2,61	4,77	-1,64	-1,62
<b>Areal [ha]</b>	1,33	0,48	6,57	-0,01	-0,01
<b>Stamantal [st]</b>	251,56	169,59	280,16	61,21	38,00
<b>Volym [m<sup>3</sup>sk/ha]</b>	79,09	47,50	124,62	20,40	9,12

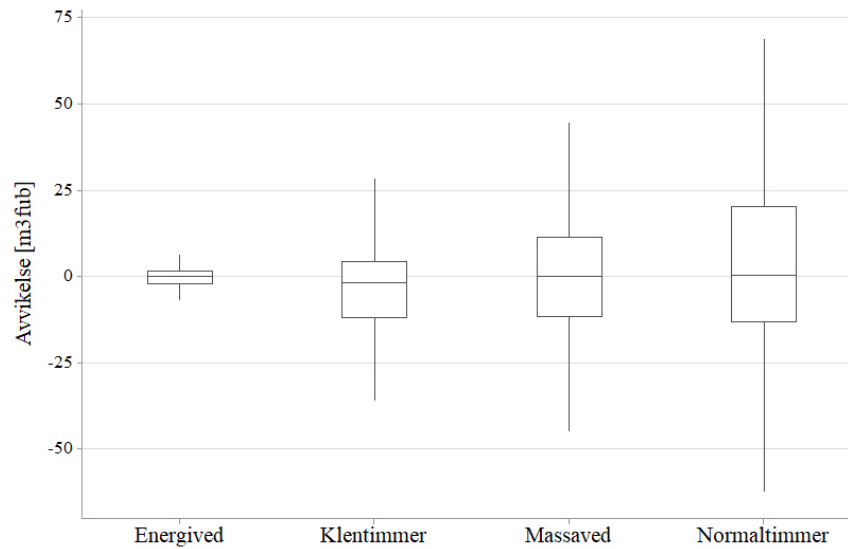
Resultatet av den stegvisa regressionen var att endast den kontinuerliga variabeln volym och den kategoriska variabeln resultatområde togs med i regressionen. Detta gav en förklaringsgrad för regressionen på 14,23 % och p-värde på <0,05 för samtliga variabler, förutom för Norra Bergslagen som har ett p-värde på 0,647. Variablerna grundyta, Dgv, Hgv, stamantal, areal och beståndstyp togs bort av den stegvisa regressionen. Resultaten visar att åtgärdens volym per ha och till vilket resultatområde åtgärden tillhör har ett samband med de relativa volymsavvikelserna.

De linjära ekvationerna för resultatområdena presenteras nedan:

- Götaland: relativ avvikelse volym =  $0,4840 - 0,001265 * \text{volym (m3sk/ha)}$
- Norra Bergslagen: relativ avvikelse volym =  $0,5058 - 0,001265 * \text{volym (m3sk/ha)}$
- Norrbotten: relativ avvikelse volym =  $0,2888 - 0,001265 * \text{volym (m3sk/ha)}$
- Södra Bergslagen: relativ avvikelse volym =  $0,6095 - 0,001265 * \text{volym (m3sk/ha)}$
- Södra Norrland: relativ avvikelse volym =  $0,2263 - 0,001265 * \text{volym (m3sk/ha)}$
- Västerbotten: relativ avvikelse volym =  $0,2141 - 0,001265 * \text{volym (m3sk/ha)}$

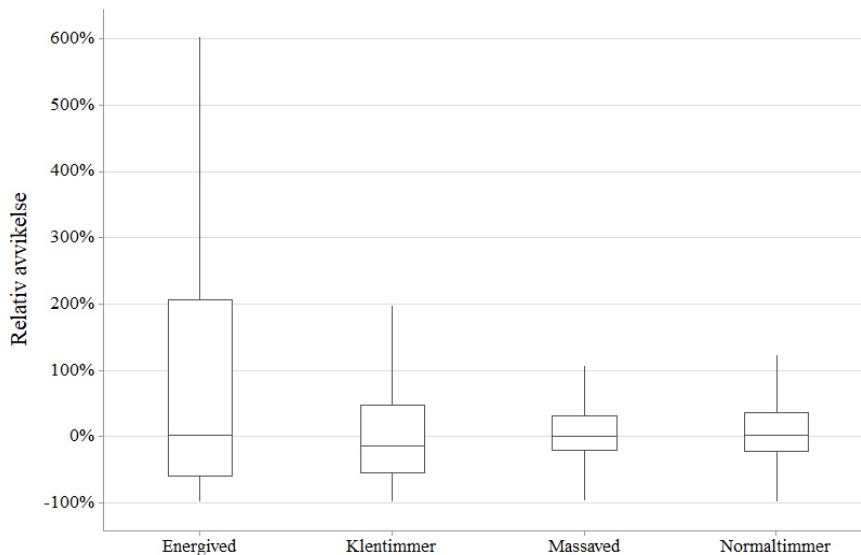
### 3.3 Utvärdering av direktutbyte

Precisionen och noggrannheten för imputeringen presenteras som både en faktisk avvikelse och en relativ avvikelse. Här kan stora variationer ses mellan både om man studerar den faktiska och den relativa avvikelsen för de olika produktgrupperna. Avvikelsen ligger nära 0 för alla produktgrupper, förutom kluentimmer, vilket visar på en hög noggrannhet för dessa produktgrupper och en något sämre noggrannhet för kluentimmer (Figur 5 & 6).



**Figur 5.** Lådagram som illustrerar avvikelserna mellan volym beräknat genom imputering och den skördarmätta volymen för fyra produktgrupper. Strecket i lådan visar avvikelsernas medianvärde.

**Figure 5.** The boxplot illustrates the difference between volume calculated with imputation and the volume measured with the harvester for four groups of assortments. The median is illustrated with a line inside the box.



**Figur 6.** Lådagram som visar de relativa volymavvikelserna mellan volym beräknat genom imputering och den skördarmätta volymen för fyra produktgrupper. Strecket i lådan visar medianvärdet för den relativa volymavvikelsen.

**Figure 6.** The boxplot illustrates the relative difference between volume calculated with imputation and the volume measured with the harvester for four groups of assortments. The median is illustrated with a line inside the box.

Samtliga produktgrupper har medianvärden som ligger när noll, vilket visar på god precision och noggrannhet för prognosmetoden. Jämför man dessa värden med exempelvis bias visar resultaten att extremvärden förekommer och då framförallt extrema överskattningar. Resultaten visar att det framförallt är normaltimmer som sticker ut, genom att uppvisa höga värden för samtliga statistiska mått förutom bias (Tabell 10).

**Tabell 10.** Medelfel (RMSE), absolut medelfel (MAE), standardavvikelse (SD), systematiskt fel (Bias) beräknat på de fyra produktgrupperna. Medianvärdet är beräknat på skillnaden mellan det prognostiserade värdet och det skördarmätta värdet

**Table 10.** Root mean square error (RMSE), mean absolute error (MAE), standard deviation (SD), Bias for four groups of assortments. The median is calculated on the difference between prognosed and the value measured with harvester

	RMSE	MAE	SD	Bias	Median
<b>Energived [m<sup>3</sup>fub/ha]</b>	13,08	6,18	6,57	4,09	0,01
<b>Klentimmer[m<sup>3</sup>fub/ha]</b>	144,17	23,10	71,85	22,47	-1,83
<b>Massaved [m<sup>3</sup>fub/ha]</b>	233,76	27,00	93,63	23,25	-0,04
<b>Normaltimmer [m<sup>3</sup>fub/ha]</b>	987,99	66,54	225,6	52,76	0,50



## 4 Diskussion

### 4.1 Kartläggning av arbetsprocessen för utbytesprognoser hos Sveaskog

Denna metod att göra utbytesprognoser har beskrivits en gång tidigare i Möller *et al.* (2017). I rapporten beskrevs processen relativt detaljerat och en skiss över det tänkta arbetsflödet presenterades också. Med detta som grund gjordes i detta arbete en kompletterande kartläggning av arbetsprocessen för utbytesprognoser hos Sveaskog för att få en bild av hur metoden har kunnat implementeras i en befintlig verksamhet. Resultaten visar på att det finns flera både likheter och skillnader mellan arbetsprocessen som beskrevs i Möller *et al.* (2017). Till att börja med finns det skillnader mellan de begrepp som använts för att beskriva de olika ingående funktionerna, vilket kan vara viktigt att vara medveten om. Bortsett från detta så är de skillnader som framkommit att det i beskrivningen av arbetsprocessen i Möller *et al.* (2017) fanns fyra förslag på ingångsdata för prognosen, manuell inmatning av beståndsdata, import från beståndsregister, skogliga grunddata från nationell laserskanning och beståndsbeskrivning från hprYield. Hos Sveaskog används i stort sett endast en typ av ingångsdata som bygger på de skattningar i fält om den skogliga planeraren gör. Nästa steg i arbetsprocessen har i detta arbete delats in i två funktioner, "Produktionsdata" och "Bruttourval", detta kallades endast för "Bruttourval" vid framtagandet av arbetsprocessen i Möller *et al.* (2017). Sveaskog har till stora delar anammat denna metod och det finns inga större skillnader mellan dessa att lyfta. De två sista stegen som finns i arbetsprocessen, "Matchning" och "Utbytessimulering" är även de i stora drag väldigt liknande de funktioner som presenterats i Möller *et al.* (2017). De har även vävts in i de system för produktionsplanering som finns hos Sveaskog för att de personer som kan ha användning för utbytesprognosen på ett enkelt sätt ska kunna ta del av dessa. Att det finns stora

likheter mellan de båda beskrivningar av arbetsprocessen var väntat, dels för att representanter från Sveaskog har varit delaktig vid framtagande av ursprungsbeskrivningen och för att representanter från både Sveaskog och Skogforsk har varit delaktig vid framtagandet av beskrivningen i detta arbete.

Något som inte togs upp i den tidigare processbeskrivningen var möjligheten att även skicka feedback på utfallet till den skogliga planeraren. Något som vid framtagande av denna beskrivning togs upp som en stor potential med detta arbetssätt och som i framtiden förhoppningsvis finns med i processen.

Att kunna skicka feedback på utfallet från avverkningen till den skogliga planeraren bör ses som ett viktigt informationsutbyte i försörjningskedjan då den analys av skogliga data som gjorts visar på att det finns en del systematiska fel i skattningarna. Skattningarna på Hgv och areal har hög precision och noggrannhet, jämfört med övriga parametrar. Bland de övriga parametrarna finns det en tendens till att överskatta grundyta och stamantal samt att underskatta Dgv. Detta resulterar även i att volymen som helhet tenderar att överskattas. Resultaten av denna studie stämmer väl överens med de resultat som presenterades i Kangas *et al.* (2004) där det konstaterades att den största avvikelser i de subjektiva skattningarna var för stamantal och grundyta. Detta gjordes genom att låta 19 personer göra subjektiva bedömningar av 18 provytor och sedan jämföra deras skattningar med referensvärden, noggrant uppmätta värden för provytorna.

## 4.2 Utvärdering av skogligt data

Resultaten av regressionsanalysen visar också på att det finns ett visst samband mellan avvikelser i volym och resultatområde. Skillnaderna mellan resultatområden skulle kunna förklaras av både de skillnader som finns mellan skogar i olika geografier (trädslagsblandning och åtgärdens areal), men även av att det inom olika resultatområden kan finnas olika arbetssätt och såklart olika individer som bedömer skogen olika. Någon vidare studie på detta har inte rymts inom denna studie och för att klargöra vad skillnaderna beror bör dataunderlaget kompletteras med en kartläggning av hur man arbetar med skoglig planering i respektive resultatområde. På detta vis skulle man då kunna påvisa om det finns några skillnader i arbetssätt och om dessa i sådana fall har något samband med skillnader mellan skattning och utfall.

Beståndstyp (trädslagsblandning) togs med som en variabel vid regressionsanalysen men resultaten visade att denna variabel inte hade någon signifikant påverkan. Vilket till viss del bör tala för att skillnaden mellan olika resultatområden inte enbart kan förklaras av de skillnader som finns mellan skogar i olika geografier. Dock kan det även poängteras att den indelning olika beståndstyper som gjordes var en ganska grov indelning och det är möjligt att trädslagsfördelningen skulle haft en större inverkan på resultaten om en mer noggrann indelning hade kunnat göras. Orsaken till att det finns skillnader mellan de olika resultatområdena har inte kunnat undersökas mer noggrant.

### 4.3 Utvärdering av direktutbyte

Analysen av utbytesprognoserna visar att det finns en del avvikelser för de olika produktgrupperna. Den relativa avvikelsen för produktgrupperna visar att både precisionen och noggrannheten är störst för massaved och normaltimmer. För energived och klintimmer är noggrannheten hög men precisionen något lägre, sett till den relativa avvikelsen (Figur 6). Ser man till den faktiska avvikelsen är precisionen bättre för energived och klintimmer än massaved och normaltimmer, vilket beror på att volymerna för dessa produktgrupper generellt sett är betydligt lägre (Tabell 10). För vissa åtgärder har prognoserna gett väldigt höga volymsskattningar, de har i vissa fall varit på några tusen m<sup>3</sup>fub/ha, något som bör vara orealistiskt. Dessa värden har inte tagits bort från datamaterialet och därför även påverkat resultaten och bör kunna förklara att RMSE är högt.

En liknande analys av utbytesprognoserna har genomförts i Söderberg *et al.* (2018). Några av resultaten från den rapporten har redovisats i tabell 2. Det finns tre stora skillnader mellan studien gjord av Söderberg *et al.* (2018) och denna studie. Deras studie genomfördes med ett ingående datamaterial som man visste inte innehåller några felaktigheter, analysen gjordes på en mindre nivå (beskrivningsenhet istället för åtgärd) och datamaterialet var mindre. Jämför man resultaten från denna studie med de som Söderberg *et al.* presenterade ser man att de kunnat visa på en betydligt högre precision och noggrannhet än vad denna studie visar. Till exempel var RMSE för normaltimmer 17,9 i studien av Söderberg *et al.* (2018) och i denna studie var RMSE för normaltimmer 987,99. Bias för samma produktgrupp var i (Söderberg *et al.* 2018) 1,4 och 52,76 i denna studie. Samma tendenser kunde även ses för övriga produktgrupper. Detta bör visa att de felaktigheter som finns i det skogliga data som denna studie kunnat visa ger en mindre

tillförlitlig prognos, vilket även kan kopplas till storleken på ytan som imputeringen görs på.

#### 4.4 Styrkor och svagheter i studien

Metoden som valdes för att genomföra och redovisa kartläggningen var IDEF0-modellen. För att använda denna modell på rätt sett bör kartläggningen då ha innehållit hela försörjningskedjan för att på ett korrekt sätt koppla samman samtliga pilar och funktioner. På grund av arbetets avgränsningar valdes det dock att endast göra en modell för arbetsprocessen att ta fram utbytesprognoser vilket dessvärre resulterar i att några pilar inte har någon koppling till en funktion. Detta har dock lösts genom att förklara dem lite mer i detalj i den tillhörande tabellen.

För att analysera det skogliga data gjordes bland annat en stegvis regressionsanalys. Denna metod valdes då det fanns flera förklarande variabler som även hade starka korrelationer mellan varandra. Den stegvisa regressionsanalysen genomfördes i en modell i Minitab 18 och att använda denna metod är förknippat med viss osäkerhet. Bland annat finns det risk för att det inte är den modell med högst förklaringsgrad som i slutändan väljs av programmet. När det finns flera starkt korrelerande variabler, likt detta data-material, finns det risk för att endast en tas med men flera kunde ha varit viktiga för modellen (Minitab 2019). Analysen hade kunnat förbättrats genom att manuellt göra den stegvisa regressionsanalysen. På detta vis hade man med säkerhet kunnat säga att modellen som tas fram är den som bäst beskriver datamaterialet och att samtliga viktiga variabler finns med. Den manuella metoden hade dock varit väldigt tidskrävande och valdes därför bort. För analysen av både de skogliga skattningarna och imputeringen gjordes inget urval av åtgärder utan alla åtgärder användes. Det var endast åtgärder som saknade väsentliga uppgifter som togs bort från analysen. Att analysen tog hänsyn till samtliga åtgärder under en given tidsperiod bör ses som en styrka i studien. Vid analysen av imputeringen visade resultaten att datamaterialet innehöll flera orealistiska observationer. Dessa kunde vara upp mot 2000 m<sup>3</sup>fub/ha för exempelvis normaltimmer. Att ta med dessa observationer påverkade även resultaten till att visa stor spridning för prognoserna, men skulle dessa observationer plockas bort skulle man även behöva sätta en gräns för vad som är rimligt eller ej. Detta hade även kunnat lösas genom att sätta upp en begränsning på maximal avvikelse mellan prognos och utfall. Dessa avvikande värden valdes att inte plockas bort på grund av

arbetets tidsbegränsning. Att dessa värden fanns kvar i datamaterialet vid analysen bidrog till en större osäkerhet kring resultaten.

#### 4.5 Implementering och framtida studier

För att få en högre precision och noggrannhet på direktutbytet skulle eventuellt en validering av data kunna göras. Som processen med att ta fram direktutbyte och utbytesprognoser ser ut idag så används skogliga data som insamlats av planeraren till att göra det bruttourval som prognoserna baseras på (Figur 3). Innan detta steg skulle man då kunna lägga till någon form av validering som exempelvis skulle kunna innebära en korrigering för systematiska fel. En sådan justering skulle kunna bygga på en modell per resultatområde som tas fram genom att jämföra planerat data med data från åtgärdens utfall. Justeringen skulle också kunna göras utifrån data från laserskanning för den specifika yta som åtgärden ska utföras på, detta skulle då kunna vägas samman med det som skattats i fält.

Studien som gjordes av Malinen *et al.* (2018) visar hur väl utbytesprognoser kan skatta värdet på en avverkning. Detta skulle vara en intressant att göra även för utbytesprognoserna hos Sveaskog om man i framtiden även vill implementera dessa i verksamheten för virkesköp.

Resultaten från denna studie visar på vikten av att ha ett korrekt dataunderlag för att kunna dra nytta av potentialen av att kunna göra utbytesberäkningar baserade på skördardata. Genom att jobba vidare med att förbättra ingångsdata kan metoden bli väldigt användbar för skogsföretag som är beroende av goda skattningar av sortimentsutfall.

I detta arbete har det skogliga datamaterialet och datamaterialet för direktutbyte studerats separat. Den ursprungliga planen var att dessa både skulle kunna kopplas samman för att kunna studera hur väl direktutbyte eller utbytesprognoser fungerar för olika typer av åtgärder. Detta kunde dock inte göras på grund av arbetets tidsbegränsning. En självklar utveckling och fortsättning på detta arbete skulle därför vara att i en ny studie koppla samman dessa. Det skulle också vara intressant att studera ett större datamaterial som framförallt sträcker sig över en längre tidshorisont för att fånga upp fler påverkande faktorer så som olika avverkningssäsonger och så vidare. Med ett större datamaterial skulle man även studera i vilken utsträckning den enskilda individen som utför planeringen påverkar avvikelserna i de skogliga skattningarna. Förutom fler åtgärder skulle även ett större datamaterial kunna innehålla flera olika avverkningstyper än just förnygringsavverkningar som har studerats här.

#### 4.6 Slutsatser

- Det finns systematiska avvikelser i de skogliga skattningarna, där framförallt grundyta, stamantal och volym tenderar att överskattas samt skattas med en relativt stor spridning.
- Åtgärdens volym per ha och vilket resultatområde åtgärden tillhör påverkar felaktigheter i volymsskattningar vid skogliga inventeringar.
- Resultaten visar en bra noggrannhet men dålig precision för beräkningarna av direktutbyte.
- Resultaten från denna studie kan bidra till den fortsatta utvecklingen av utbytesprognoser baserade på skördardata genom att jobba med de brister som lyfts i denna rapport. Kan man även koppla samman det skogliga datamaterialet med direktutbyte så kan man även studera när avvikelserna uppkommer mer i detalj.

## Referenslista

- Ang, C.L., Luo, M., Khoo, L.P. & Gay, R.K.L. (1997). A knowledge-based approach to the generation of IDEF0 models. *International Journal of Production Research*, 35:5, ss. 1386-1412.
- Cernold, Å. (1971). *Utbytestabeller för rotstående skog*. 3 uppl. Falun: Sågverksintressenter.
- Gellerstedt, S. & Dahlin, B. (1999). Cut-To-Length: The Next Decade. *Journal of Forest Engineering*, 10:2, ss. 17-24.
- Holappa Jonsson, S. (2018). *Evaluation of the increased pre-harvest forecasting precision of sawlog supply by use of historical harvester data and wood properties models - A case study on Scots pine in northern Sweden*. Umeå: Institutionen för skogens biomaterial och teknologi.
- Kangas, A. (2010). Value of forest information. *European Journal of Forest Research*, 129, ss. 863-874.
- Kangas, A., Heikkinen, E. & Maltamo, M. (2004). Accuracy of partially visually assessed stand characteristics: a case study of Finnish forest inventory by compartments. *Canadian Journal of Forest Research*, 34, ss. 916-930.
- Körner, S. & Wahlgren, L. (2005). *Statistiska metoder*. 2 uppl. Lund: Studentlitteratur.
- Malinen, J., Kilpeläinen, H. & Verkasalo, E. (2018). Validating the predicted saw log and pulpwood proportions and gross value of Scots pine and Norway spruce harvest at stand level by Most Similar Neighbour analyses and a stem quality database. *Silva Fennica*, 52:2, ss. 1-21.
- Malinen, J., Kilpeläinen, H. & Ylisiniö, K. (2014). Description and evaluation of Prehas software for pre-harvest assessment of timber assortments. *International Journal of Forest Engineering*, 25:1, ss. 66-74.
- Minitab *Basics of stepwise regression*. Tillgänglig: <https://support.minitab.com/en-us/minitab/18/help-and-how-to/modeling-statistics/regression/supporting-topics/basics/basics-of-stepwise-regression/> [2019-11-27].
- Möller, J.J., Arlinger, J., Bhuiyan, N., Eriksson, I. & Söderberg, J. (2017). *Utbytesprognoser baserade på skogs- och skördardata - Modell- och systembeskrivning för*

*skapande av stambanker och imputerat utbyte.* (Arbetsrapport 961. Uppsala: Skogforsk.

Nilsson, M., Staal Wästerlund, D., Wahlberg, O. & Eriksson, L.O. (2012). Forest Planning in a Swedish Company - a Knowledge Management Analysis of Forest Information. *Silva Fennica*, 46(5), ss. 717-731.

Nilsson, M., Staal Wästerlund, D., Wahlberg, O. & Eriksson, L.O. (2017). The use of forest information in timber sales planning - a case study in a Swedish forest-owning company. *Scandinavian Journal of Forest Research*, 32(4), ss. 320-326.

Ranvald, C. (2019). Projektledare, Sveaskog

Riksskogstaxeringen, S. (2019). *Skogsdata 2019*. Umeå.

Ståhl, G. & Wilhelmsson, E. (2009). *Planering av skogsbruk*. Umeå: Sveriges lantbruksuniversitet, Institutionen för skoglig resurshushållning.

Sveaskog (2018). *Års- och hållbarhetsredovisning*. Stockholm.

Söderberg, J., Möller, J.J. & Willén, E. (2018). *Utvärdering av utbytesprognoser med skördardata*. (Arbetsrapport 981. Uppsala: Skogforsk.

Söderberg, J., Willén, E., Möller, J.J., Arlinger, J. & Bhuiyan, N. (2017). *Utvärdering av utbytesprognoser med skogliga laserskattningar och skördardata - resultat från tre fallstudier*. (Arbetsrapport 937. Uppsala: Skogforsk.

Söderholm, J. (2002). *De svenska skogsbolagens system för skoglig planering*. (Arbetsrapport 98. Sveriges Lantbruksuniversitet: Institutionen för skoglig resurshushållning och geomatik. Umeå.

## **Muntliga källor**

Ranvald, C. (2019). Projektledare, Sveaskog