



Prognostisering av sortimentsutfall – korrelation mellan pri-filer och beståndsdata

*Predicting the assortment of logging – correlation between pri-
files and stand data*

Erik Persson

**Arbetsrapport 17 2014
Examensarbete 30hp A2E
Jägmästarprogrammet**

**Handledare:
Tomas Nordfjell**

Prognostisering av sortimentsutfall – korrelation mellan pri-filer och beståndsdata

*Predicting the assortment of logging – correlation between pri-
files and stand data*

Erik Persson

Nyckelord: regressionsanalys, utbytesberäkning, skördardata, beräkningsprogram

Arbetsrapport 17 2014

Jägmästarprogrammet

EX0772, A2E, Examensarbete vid institutionen för skogens biomaterial och teknologi

Handledare: Dag Fjeld, Institutionen för skogens biomaterial och teknologi

Examinator: Ola Lindroos, SLU, Institutionen för skogens biomaterial och teknologi

Sveriges lantbruksuniversitet

Institutionen för skogens biomaterial och teknologi

Utgivningsort: Umeå

Utgivningsår: 2014

Rapport från Institutionen för skogens biomaterial och teknologi

Förord

Jag vill börja med att tacka min handledare på Sveaskog Jonas Gustafsson för förtroendet och möjligheten att skriva detta arbete.

Flera personer har varit involverade och gett stöd under arbetet. Ett stort tack till min handledare Tomas Nordfjell för stödet under examensarbetet. Jag vill tacka Anton Grafström, för svaren kring de statistiska funderingar som dykt upp.

På Skogforsk vill jag tacka John Arlinger och Nazmul Bhuiyan för att de tagit sig tid och hjälpt mig, dels utifrån deras erfarenhet på området men även för tillhandahållande av de mjukvaruprogram som jag använt mig av under arbetets gång.

Utöver tidigare nämnda personer vill jag tacka Tobias Norrbom på Sveaskog för hans hjälp med data och snabba svar.

Min far, Ulf Persson, skall ha särskilt tack för råd, stöd och korrekturläsning under arbetets gång.

Umeå mars 2014

Erik Persson

Sammanfattning

Målet med detta arbete har varit att undersöka potentialen för skördardata (pri-filer) när det kommer till prognostisering av sortimentsutfall från avverkningar. Det gjordes genom regressionsanalys, där data från Sveaskogs beståndsregister utgjorde de förklarande variablerna och data från pri-filer utgjorde responsvariablerna. Arbetet hade i huvudsak tre analysmoment: undersökning av beståndsregistrets noggrannhet, samt hur väl arbetets prognoser stod sig mot facit (pri-filer) och dels mot Sveaskogs egna prognoser via VALS. Totalt utgjordes data av 42 trakter med tillhörande beståndsdata innan avverkning och de traktspecifika pri-filer som skapades under avverkningarna.

Det påvisades avvikelser, stundtals stora avvikelser, mellan beståndsregistret och pri-filer. Avvikelserna redovisades inte per hektar, p.g.a. avvikelserna i areal mellan beståndsregistret och pri-filerna. För vissa pri-filer saknades dessutom gps-koordinater för de enskilda träden, vilket omöjliggjorde arealberäkning. För regressionsanalysen tillhandahölls åtta förklarande variabler sammanställda utifrån data från 20 trakter och åtta sortimentsvisa funktioner skapades. Funktionerna validerades sedan av data från sex trakter.

Prognostisering för totala volymer per trakt gav högre noggrannhet än prognostisering för enskilda sortiment. Funktionerna för grantimmer hade lägre förklaringsgrad än tallfunktionerna, men hade högre noggrannhet vid valideringen. Variabeln stamfelsesved utnyttjades i sju funktioner och bör anses som mycket viktig vid prognostisering.

Funktionerna bör inte anses som tillförlitliga då de baserats på data som anses ha varit osäker och för liten till sin omfattning. Dock är metoden applicerbar och bör kunna generera bra prognoser om dataunderlaget är större och har högre noggrannhet. Arbetet visar också att smarta systemlösningar och tydliga förhållningsätt i tillvaratagandet av skördardata är nyckelfaktorer inför framtida studier inom området.

Nyckelord: regressionsanalys, utbytesberäkning, skördardata, beräkningsprogram

Summary

The aim of this study was to examine the potential for harvester data (pri-files) when it comes to forecasting the assortments outcomes from harvesting. It was examined by regression analysis, in which data from stands constituted the explanatory variables and data from pri-files constituted response variables. The work had three main analysis steps: the accuracy of stand data, and how well the created predictive functions corresponded with pri-files (actual assortment outcome from harvesting) and with Sveaskog's own forecasts from VALS, respectively.

Total data consisted of 42 stands with associated stand data before harvesting and the stand-specific pri-files created during harvesting. It was demonstrated abnormalities, and sometimes large discrepancies, between stand data and pri-files. The deviations were not presented per hectare, due to deviations in the area between stand data and pri-files. Additionally for certain pri-files there were missing gps-coordinates for individual trees, which made it impossible to calculate area. The regression analyses were provided with eight explanatory variables, compiled from data from 20 stands, and eight assortments specific functions were created. These functions were then validated with data from six other stands.

Prediction of total volumes per stand was more accurate than predictions per assortment. The functions for spruce-timber explained less of the observed variation than pine-timber-functions, but gave higher accuracy during validation. The variable stemwood damage was utilized in seven functions and should be considered as a very important variable in forecasting. These functions should not be considered as reliable as they are based on data that were too uncertain and too small in scope. However, the method is applicable and should be able to generate good predictions if the data set is larger and has higher accuracy. The work also shows that smart system solutions and clear approach at the procurement of harvest data are key factors for future studies in the area.

Keywords: regression analysis, yield calculation, harvest data, calculation software

Innehållsförteckning

1. Introduktion	2
1.1 Prognostisering av sortiment inom skogsbruket.....	2
1.2 Pri-fil - vanligt filformat för skördardata.....	3
1.3 Sveaskog.....	4
2. Syfte.....	6
2.1 Mål.....	6
2.2 Arbetets avgränsning	6
3. Material och metoder.....	7
3.1 Generell överblick av metodens struktur.....	7
3.1.1 Överblick över utnyttjade program och filformat.....	9
3.2 Bearbetning av beståndsdata	9
3.3 Bearbetning av pri-filer kopplade till beståndsdata.....	11
3.3.1 Bearbetning av stockdata inför regressionsanalysen.....	11
3.3.2 Bearbetning av stamdata för analys av beståndsregistrets noggrannhet.....	12
3.3.3 Bearbetning av pri-filernas GPS-positioner för skattning av beståndsareal.....	12
3.4 Framställandet av regressionsfunktioner	12
3.5 Analyser.....	12
3.5.1 Analys 1 - Jämförelse mellan beståndsregistret och verkligt utfall (pri-filer).....	12
3.5.2 Analys 2 & 3 - Skillnad mellan VALS-prognoser och arbetets prognoser samt skillnad mellan arbetets prognostiserade volymer och verkligt utfall (pri-filer).....	13
3.5.3 - Allmän felsökning kring avvikelser	13
4. Resultat	14
4.1 Analys 1- Hur väl stämde beståndsregistret med utfallet från avverkningen?	14
4.2 Analys 2 – Hur skiljde sig Sveaskogs prognoser med VALS gentemot arbetets prognoser?	15
4.3 Analys 3 – Hur skiljer sig utbytesprognoserna mot verkligt utfall?.....	17
5. Diskussion	19
Referenslista	22

1. Introduktion

1.1 Prognostisering av sortiment inom skogsbruket

Det är viktigt att företag inom skogsbranschen har ett bra planeringsunderlag och god uppfattning om vilka sortiment, dimensioner, kvalitéer, trädslag och värden som en avverkning genererar. Med ett bra planeringsunderlag och en god uppfattning möjliggörs en effektiv och korrekt planering av företagsverksamheten (för berörda företag inom skogsbranschen) över kort- och lång tidshorisont (Ollas, 1980). För att skapa ett bra planeringsunderlag har det varit brukligt att utnyttja utbytesfunktioner för träd och trakter (Ståhl & Wilhelmsson, 1994). Hur väl utbytesfunktionernas skattningar stämmer med verkligt utfall varierar, men högst överensstämmelse erhålls när skattningarna utförs på enskilda träd (Ståhl & Wilhelmsson, 1994). En försvårande omständighet med utbytesberäkningar är att ett träd kan generera olika sortimentsutfall beroende på hur apteringen utförs. Apteringen och således sortimentsutfallet styrs av flera olika faktorer, som trädets avsmalning, kvalitetsgränser och prislister (Ståhl & Wilhelmsson, 1994). Ollas (1980) nämner också att den minsta tillåtna toppdiametern på sågtimret i hög grad påverkar timmerandelen av stamvolymen. Även trädslag, stockkvalité, ålder och definitionen av timmer mellan företag påverkar timmerandelen i trädet (Ollas, 1980). Utöver de faktorer som påverkar apteringen och sortimentsutfallet finns det en skillnad mellan teoretiskt optimal apteringen och vad som är praktiskt utförbart. Bedömningen har gjorts att 98 % apteringsgrad (dvs 98 % av vad som optimalt kan uppnås) är maximum vid praktiskt utförande och 95 % anses som bra, med ett minimum på 75 % (Ståhl & Wilhelmsson, 1994).

Det finns flera personer som har skrivit om och bidragit till ämnet utbytesberäkningar. Åke Cernhold (1981) gjorde utbytestabeller för rotstående skog. För att utnyttja dessa tabeller behövs träddata över höjd och brösthöjdsdiameter och via tabellerna kan råvaruvolymer avläsas och sedan omvandlas till diverse måttenheter. Exempelvis kan volymen gagnvirke fördelas över timmer och massaandel (Cernhold, 1981). Via tabellerna kan även utbytet av biomassa skattas i form av grenar, toppar och barr samt stubbar (Cernhold, 1981). Även Rune Ollas (1980) har bidragit till ämnet med sina utbytesfunktioner. Ollas funktioner används även i dagens skogsbruk av t.ex. SCA, om än modifierade (Samuelsson, 2005). I likhet med Cernholds utbytesberäkningar är trädhöjd och brösthöjdsdiameter nödvändigt indata för Ollas utbytesberäkningar.

Samuelsson (2005) fastslår att utbytesberäkningar med hjälp av datorprogrammen Bass Best/upr från Berget Systemdesign och Timan-Aptan från Skogforsk skiljer sig åt. Mer specifikt konstaterades avvikelser för enskilda sortiment mellan prognostiserade volymer och referensvolymerna med ett spann mellan en underskattning på 10 % (Bass Best/upr) till en överskattning på 47 % (Timan-Aptan). Samuelsson (2005) konstaterar också att stamfelsandelen i hög grad påverkar sortimentsutfallet.

Prognostisering via regressionsanalys är också genomförbart. Eriksson (2010) utnyttjade regressionsanalys och avgränsade sig till en upplösning för prognoserna i storleksordningen

”en månads sortimentsutfall för ett större geografiskt område.” och fick goda resultat. Regressionsfunktionerna skapades utifrån olika indata, som beståndsregistret och fältplanerade trakter, men samma responsvariabel (volymen inmätta vid industri). Regressionsfunktionerna som nyttjade indata från fältplanerade trakter genererade bäst prognoser och som anledning angavs den högre noggrannheten (Eriksson, 2010).

Även Barth et al. (2013) visade att noggranna indata är viktiga för att få bra utbytesberäkningar. I studien undersöktes hur ett bättre underlag av indata påverkar utbytesberäkningar. Utbytesberäkningarna baserades på laserdata samt traditionell subjektiv skattning (fältinventering) och jämfördes med det verkliga utfallet i skördardata (pri-filer). Resultatet blev att de utbytesprognoser som baserats på laserdata genererade bättre prognoser jämfört med fältinventeringen. Prognoserna blev mer noggranna till följd av att laserdata beskrev trädens diameterfördelning och längder mer noggrant. Skadefrekvensen som genererar nedklassningar av stockar från timmerkvalité till massakvalité erhöles genom skattningar från tidigare pri-filer i området vilket gav ett bra underlag för beräkning av andelar nedklassat virke inom diameterklasser och trädslag (Barth et al. 2013). Det konstaterades att välskötta d.v.s. homogena skogar var lättare att utföra utbytesberäkningar på, samt att korrekt uppfattning av trädform och skadefrekvens i skogarna är av betydelse för att generera bra utbytesberäkningar (Barth et al. 2013).

Det finns stor potential i skördardata, som ofta finns tillgängligt i filformatet pri. Potentialen ligger framförallt i utförligare beskrivningar av skogsråvaran på volymsnivå, men även egenskapsnivå skapar förutsättningar för en bättre integration mellan skogsbruket och industri (Nordström et al. 2010). Konsekvenserna av en förbättrad integration och mindre osäkerhet kring inkommande råvara kan resultera i ett höjt värde på industrins produkter och lägre kostnader. De positiva konsekvenserna är till följd av mer tillförlitliga leveranser som möjliggör ett mindre behov av säkerhetslager, färre vrak och allmänt mindre spill av råvara (Nordström et al. 2010). Exempel på en förbättrad integration kan vara att sågverken delges specifikationer på inkommande råvara såsom kärnvedsandel och kvistgrovlek men även att sågverk kan lägga specifika beställningar (förutsatt utbytesberäkningar med mycket god precision). Massaindustrin kan erhålla mer detaljerad information om t.ex. torrsvikt och massavedens dimensioner. Utifrån skördardata kan energisektorn få utförlig information kring råvarans fukthalt och energiinnehåll (Nordström et al. 2010).

För att dessa förbättringar skall bli realitet behövs systemlösningar som tar till vara på och sammanställer den viktiga och stora informationskälla som skördardata utgör (Nordström et al. 2010).

1.2 Pri-fil - vanligt filformat för skördardata

Pri-filer (produktion, individuell) skapas i samband med avverkning och innehåller gps-koordinater (oftast) och detaljdata över träd och stockar från avverkningen, exempelvis sortiment (SSTE, där SS står för sortiment, T för trädslag och E för egenskaper), trädslag och stockarnas längd (Arlinger et al. 2003). Datastrukturen i pri-filen bygger på "Standard for Forestry Data and Communication" (StandForD) (Skogforsk 2013). Anledningen till detta är

att maskintillverkarna och skogsbruket har enats om att underlätta informationsutbytet med skogsmaskinernas datorer och följaktligen enats om en informationsstruktur (Skogforsk 2013). Information rörande stockarna lagras i variabeln LOGDATA som i sin tur är uppbyggd av NUMLOGCOD och NUMLOGS (Arlinger et al. 2003). Trädldata lagras i variablerna NUMSTEMS, NUMTRECOD, TREECODE och TREEDATA (Arlinger et al. 2003). Pri-filen är uppbyggd så att all data, rörande stammar och stockar, lagras i olika variabler beroende på datats relevans mot variablerna, exempelvis data angående stockar lagras i variabel 257 (LOGDATA) medan trädldata lagras i variabel 267 (TREEDATA).

För att undvika att stora datamängder måste skickas från skördaren till SDC (Skogsnärings IT-företag), bör pri-filen skickas in dagligen och pri-filen bör behandlas som en icke kumulativ fil (Arlinger et al. 2003).

Stockdata som måste registreras i pri-filer vid avverkning är de som följer (Arlinger et al. 2003):

- * Prismatris nummer
- * Topdiameter på bark respektive under bark
- * Längd
- * Volym enligt prismatris (var161)
- * Volym m³ på respektive under bark
- * Stamnummer
- * Stocknummer

Utöver de ovan nämnda stockdata måste även nedanstående trädldata registreras (Arlinger et al. 2003):

- * Trädslag
- * Brösthöjdsdiameter, mm

1.3 Sveaskog

Examensarbetet är utfört på uppdrag av Sveaskog, Sveriges största skogsägare. Av deras leveranser till kund kommer över 50 % från det egna skogsinnehavet. Sveaskogs omsättning uppgår till cirka 6,3 miljarder kronor och deras arbetsstyrka är i storleksordning 700 anställda. Personalen är fördelad över de tre marknadsområdena Nord, Mitt och Syd och ansvarar för de cirka 4,1 miljoner hektar mark som Sveaskog äger (Sveaskog, 2013).

I nuläget är Sveaskogs utbytesberäkningar, på operativ nivå, inte helt tillfredsställande. Därför vill Sveaskog utveckla och skapa prognosverktyg med högre noggrannhet och med fler specifikationer över råvaran. Detta eftersträvas för att möjliggöra en effektivare planering med bättre framförhållning, för Sveaskog och deras kunder. Enligt Nordström (2010) finns det potential att förbättra och utveckla sortimentsprognoser. Skördardata, mer specifikt pri-filer, möjliggör stor inblick i sortimentsutfallet från varje avverkning. Pri-filer innehåller detaljerad information över varje enskilt träd och stock. Denna information skapar en potential för

utformning av bättre sortimentsprognoser och skördardata utgör på så vis ett intressant kunskapsområde att undersöka. Ovanstående problematik utgör grunden för detta examensarbete, d.v.s. att undersöka pri-filers potential med avseende på sortimentsprognoser.

2. Syfte

Syftet med arbetet är att belysa möjligheten att koppla samman beståndsdata för ännu ej avverkade slutavverkningstrakter, med information från skördarens pri-filer från redan avverkade trakter av likartad beskaffenhet, för att på det sättet skapa funktioner över sortimentsutfallen på traktnivå.

2.1 Mål

- Undersöka pri-filers potential för sortimentsprognoser
- Skapa sortimentsprognoser, trädslagvis för sortimenten normaltimmer, klentimmer, massaved och brännved
- Belysa sortimentsprognosernas noggrannhet
- Belysa avvikelser mellan beståndsdata och pri-filer
- Belysa sortimentsprognosernas skillnader från Sveaskogs befintliga prognoser

2.2 Arbetets avgränsning

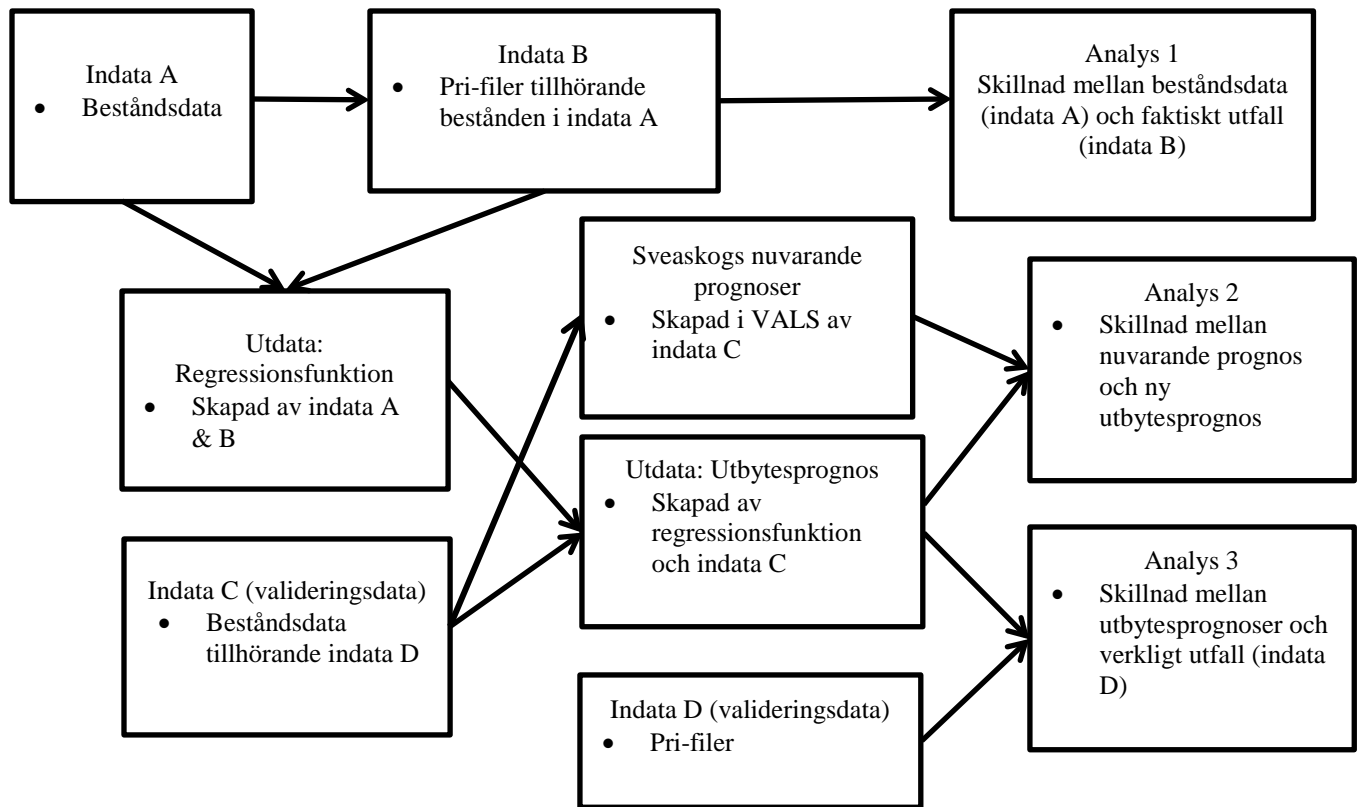
Arbetet avgränsas till slutavverkning som domineras av tall och gran. Arbetet avgränsas geografiskt till resultat område "Norra Bergslagen" i Sveaskogs marknadsområde Mitt.

3. Material och metoder

3.1 Generell överblick av metodens struktur

Indata A och indata B utgjorde grunden för konstruktion av funktionerna medan indata C och indata D användes för validering av funktionerna (figur 1 & 2). Indata A utgjordes av beståndsdata från beståndsregistret över de trakter som genererade indata B när de avverkades. Indata B bestod av pri-filer och förädlades så till vida att varje pri-fil kopplades samman med tillhörande avverkningsobjekt och att volymerna per stock i pri-filerna sammanställdes i volymer per sortiment. Under denna bearbetning noterades andelen timmerstockar som blivit nedklassade till massaved, så kallad stamfelsved. En stock är stamfelsved om topdiameterna under bark är > 15 cm och sortimentskoden (SSTE) börjar med 1 (massaved) eller 5 (röta/bränsle). Indata D utgjordes av sammanställda sortimentsvolymer från pri-filerna tillhörande bestånden i indata C (figur 1 & 2). Indata D utgjorde på så vis facit och referens för de prognostiserade sortimentsvolymer som arbetets utbytesprognoser genererade.

Avgränsningen mellan klintimmer och normaltimmer gjordes vid 18 centimeter och klintimrets nedre gräns var 11 centimeter toppmått. Brännved definierades utifrån sortimentskod (SSTE 5090) och likaså gjordes för de olika massasortimenten.



Figur 1. Generell överblick över metodiken och dess tre analysmoment (läs från vänster till höger). Indata A och B utgör data för skapandet av funktionerna och tillhör samma bestånd medan indata C och D utgjorde data för validering av funktionerna och baserades således inte på samma bestånd som indata A & B. Tre analysmoment; 1) Hur väl stämde indata A med indata B? 2) Hur väl stämde Sveaskogs prognoser med arbetets prognoser? och 3) Hur väl stämde arbetets prognoser med faktiskt utfall?

Figure 1. General overview of the methodology (read from left to right). Indata A and B are not based on the same stands as indata C and D are based on. The functions created were based on data from indata A and B and the validations of these functions were done with indata C and D. Three analysis stages, Analysis 1: How well did indata A match with indata B? Analysis 2: How well did the Sveaskog's forecast of assortments match this rapport forecast of assortment? Analysis 3: How well did the rapport forecasts match with actual assortments from logging?

Steg efter bearbetning av indata utgjordes av regressionsanalys med de förklarande variablerna ålder, höjd, stamfelsved, bonitet, stamtäthet, grundyta, volym m³fub och medelstam. De fyra sistnämnda variablerna analyserades per trädslag och de övriga var generella data för trakterna som helhet.

Ovan nämnda variabler ingick i dataseten som nyttjades vid beräkningarna med datorprogrammet Minitab (Minitab 2013). Det skapades en funktion per sortiment och trädslag (formel 1). Arbetets sista steg var att utföra de tre analyser som visas i figur 1.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7 + \beta_8 X_8 + \varepsilon_i \quad [1]$$

Formel 1. Potentiellt utseende för utbytesfunktionen då alla åtta förklarandevariabler (X) används. Y representerar en sortimentsvolym per trädslag och de förklarande variablerna representerar bonitet, ålder, höjd, stamantal, grundyta, volym, medelstamsvolym och stamfelsesved

Equation 1 . Potential look for assortment function when all eight explanatory variables (X) are used. Y represents a volume of an assortment and the explanatory variables representing site quality, age, height, number of stems, basal area, volume and proportion of downgraded logs

3.1.1 Överblick över utnyttjade program och filformat

Flera mjukvaruprogram och filformat användes under arbetets gång för att skapa och bearbeta data. Programmen hade olika grafiska gränssnitt och varierad upplösning, vilket medförde att de olika programmen utnyttjades i varierad omfattning beroende på ändamålet t.ex. felsökning och verifiering eller skapa data i bulk.

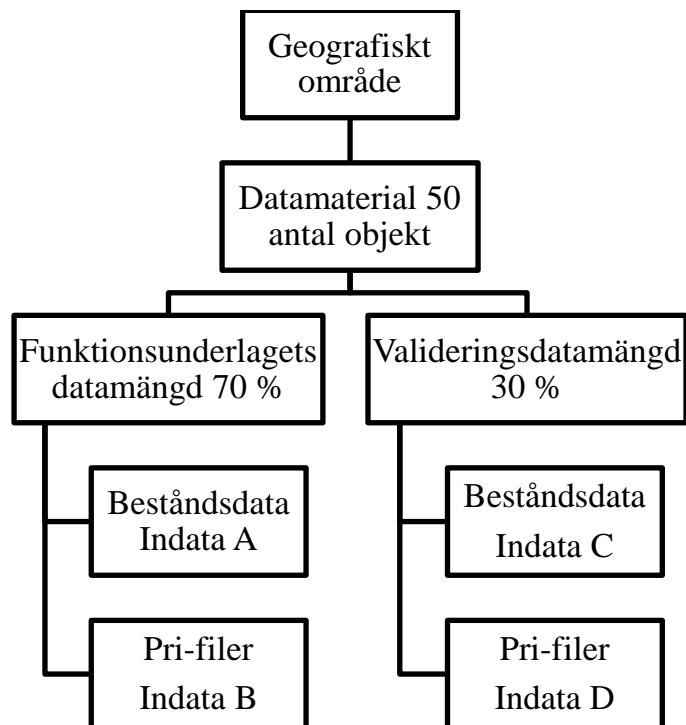
Tabell 1. Sammanfattning av de filformat och program som använts under arbetet
Table 1. Summary of file formats and software used during the work

Program	Förklaring
Biobränsle	Ett beräkningsprogram för pri-filer. Tillhandahållet av Skogforsk.
Hpr-analys	Ett beräkningsprogram för hpr-filer. Tillhandahållet av Skogforsk.
Klippsk	Ett program för att öppna pri-filer med ett enklare grafisk gränssnitt. Tillhandahållet av Skogforsk.
TatukGIS	Ett freeware GIS-program för hantering av lägesbunden information
Prototypprogram av Hpr-analys	Ett program som kan läsa in .shp, .jpeg och hpr-filer. Tillhandahållet av Skogforsk.
Excel	Ett kalkylprogram skapat av Microsoft
Filformat	
Pri	Filformat för skördardata som innehåller information om stammar och stockar från avverkning
Hpr	Filformatet som är pri-formatets efterföljare
Txt	En enkel textfil som innehåller enkel text utan någon formatering
Shp	Ett vanligt GIS-filformat för vektorer
Jpeg	Ett filformat för bilder

3.2 Bearbetning av beståndsdata

Datamaterialet för regressionsanalyserna kom från samma geografiska område ”Norra Bergslagen” för att skogen i hög grad skulle varit utsatt för likadana abiotiska och biotiska förutsättningar. För att kunna koppla pri-filer till respektive trakt utnyttjades VALSID (identifikationsnummer i Sveaskogs dataprogram VALS). För att vara säker på att pri-filerna som användes endast tillhörde en trakt valdes endast de VALSID som var kopplat till en enda

trakt. Det är vanligt förekommande att flera trakter hamnar under samma VALSID, vilket försvårar sammankopplingen mellan pri-filer och dess tillhörande trakt. Datamaterialet utgjordes ursprungligen av 50 avdelningar och till dessa begärdes tillhörande pri-filer. Materialet delades i två delar, indata A och indata C. Indata A användes för att utveckla regressionsfunktioner och indata C användes för validering av funktionerna (figur 1). Fördelningen av materialet var 70 % på indata A och 30 % på indata C (figur 2). Motivet bakom fördelningen var att det ansågs viktigare att skapa funktioner baserade på ett större material än att göra en mer omfattande validering. Urvalet genomfördes i Excel genom att datamaterialet stratifierades i grupper utifrån vanligast förekommande bonitet. Utifrån gruppens storlek i förhållande till det totala antalet trakter varierade det procentuella urvalet från varje grupp. Efter genomförd stratifiering bearbetades trakterna inom varje grupp med en slumpgenerator i Excel som genererar värden mellan 0 och 1. Slumptalen sorterades i fallande ordning och den valda fördelningen, 70/30, avgjorde antalet trakter som fördelades till indata A respektive indata C. Detta gjordes för att de mer vanligt förekommande trakterna skulle väga tyngre vid utformningen av funktionerna än de mindre vanligt förekommande trakterna. Fördelningen av bonitet inom indata A och C var procentuellt lika, fränsett i vissa fall då avrundningen gjordes till indata A:s fördel. Anledningen till den procentuellt lika fördelningen var för att hindra funktionerna från att baseras på en uppsättning av trakter (indata A) och valideras av en annan uppsättning trakter (indata C). Resonemanget bakom detta var att andelen av utbytet per sortiment antogs skilja sig åt mellan trakter med olika bonitet. Till de ursprungliga 50 avverkningstrakterna lokaliserades pri-filer till 42 trakter, vilket innebar att åtta trakter (som saknade fullständiga pri-filer) togs bort från urvalet och påverkade inte nämnvärt den procentuella fördelningen av bonitet mellan indata A och C.



Figur 2. Överblick över det ursprungliga datamaterialets struktur och omfattning
Figure 2. Overview of the structure and amount for the original data material

Det innebar att fördelningen mellan indata A och indata C förändrades från 35 respektive 15 trakter till 30 respektive 12 trakter.

Sedan skedde ytterligare en minskning i antalet trakter för dessa två dataset. Minskningen berodde på avvikelser i variabeln stamantal mellan beståndsregistret (indata A respektive indata C) och sammanställningen av pri-filer (indata B respektive indata D) tillhörande dessa bestånd. När denna avvikelse steg över 40 % av totala antalet enligt sammanställningen av pri-filer så antogs felet vara av sådan karaktär att skulle påverka funktionerna allt för negativt och trakten togs bort. Det resulterade i att det slutgiltiga antalet trakter för indata A var 20 respektive sex trakter för indata C. För de kvarvarande trakterna sammanställdes vardera åtta dataset med de åtta förklarande variablerna som användes vid skapandet av regressionsfunktionerna. Sammanställningen av åtta dataset per trakt gjordes p.g.a. det var ett dataset per sortimentsfunktion.

Tabell 2. Slutgiltig fördelning av data över indata A och indata C, således data som funktionerna baserats respektive validerats på

Table 2. Final distribution of data over the indata A and indata C, thus data functions were based and validated on, respectively

Trakternas bonitet	Trakter		Indata A		Indata C	
	Antal (n)	Andel (%)	Antal (n)	Andel (%)	Antal (n)	Andel (%)
Grupp 1 Bonitet 20-23	2	7	2	10	0	0
Grupp 2 Bonitet 24-26	15	58	12	60	3	50
Grupp 3 Bonitet 27-30	9	35	6	30	3	50
Samtliga boniteter	26	100	20	77	6	23

3.3 Bearbetning av pri-filer kopplade till beståndsdata

Första steget vid bearbetningen av pri-filerna var att öppna alla traktvis, i programmet Biobränsle. Det kunde vara allt från en till 18 pri-filer per trakt. Utifrån VALSID var pri-filerna kopplade till varje trakt. I Biobränsleprogrammet kunde stock respektive stamdata utläsas per pri-fil, och programmet skattade även trädhöjd per stam. För varje enskild pri-fil exporterades .txt-filer över stock- och stamdata som sedan öppnades i Excel för bearbetning och sammanställning.

3.3.1 Bearbetning av stockdata inför regressionsanalysen

Stockdata som exporterades och sedan öppnades i Excel bestod av fyra kolumner med rubrikerna trädslag, sortimentskod (SSTE), diameterTUB och stockvolym m³fub. De fyra kolumnerna från alla de enskilda pri-filerna från en trakt sammanställdes i ett Excelark och skapade på så vis fyra kolumner med stockdata som representerade hela den avverkade trakten. När samtliga stockdata för en trakt var sammanställd i fyra kolumner vidareförädlades stockdata via pivot-tabeller i sortimentsvolym för de åtta sortimenten kopplade till respektive funktion. Dessa steg gjordes för samtliga trakter.

3.3.2 Bearbetning av stamdata för analys av beståndsregistrets noggrannhet

Stamdata som exporterades och sedan öppnades i Excel bestod av fyra kolumner med rubrikerna trädslag, dbh, höjd och stamvolym m^3 fub. De fyra kolumnerna från alla de enskilda pri-filerna sammanställdes i ett Excelark och skapade på så vis fyra kolumner med stamdata som representerade hela den avverkade trakten. De fyra kolumnerna vidareförädlades till fyra av de fem variabler som användes vid analysen av beståndsregistret noggrannhet. De fyra variablerna var hgv (grundtyevägd medelhöjd), stamantal per trädslag, grundyta per trädslag och volym i m^3 fub per trädslag. Dessa steg gjordes för samtliga trakter.

3.3.3 Bearbetning av pri-filernas GPS-positioner för skattning av beståndsareal

Den femte variabeln som användes vid analysen av beståndsregistrets noggrannhet var areal. För att skatta arealen utnyttjades GPS-positionerna för varje träd som fanns lagrat i pri-filerna. För några pri-filer saknades dock GPS-positionerna och i dessa fall skattades ingen areal för trakten. I de fall koordinaterna fanns användes programmet Hpr-analys för att konvertera pri-filerna till hpr-filer och areal kunde sedan avläsas per hpr-fil, med standardinställningarna. För att få hela traktens areal summerades de enskilda arealerna från varje hpr-fil.

3.4 Framställandet av regressionsfunktioner

Åtta funktioner skapades, tre vardera för gran respektive tall, samt en funktion för björkmassa och en för brännved. De tre funktionerna för gran och tall utformades likadant, med en funktion för massaved och två för timmer. Timmerfunktionerna delades in i normaltimmer och klentimmer med avgränsningen 18 centimeter toppmått under bark, med en nedre gräns på 11 centimeter för "klen". Det innebar t.ex. för tall att alla stockar med sortimentskoder (SSTE) som inte är brännved (5090) eller massaved (1000) delades in i klen- respektive normaltimmer utifrån diameter.

Vid skapandet av varje funktion sammanställdes ett dataset från de 20 trakter (indata A), över de åtta förklarade variablerna samt tillhörande sortimentsvolym (t.ex. volymen klentimmer för tall som fallit ut vid avverkningen) för var och ett av de 20 trakterna. När detta var gjort lästes varje dataset in i Minitab och med hjälp av verktyget "Best subset regression" valdes den kombination av variabler som bäst lyckades förklara variationen inom responsvariabeln. Efter detta urval utfördes en regression med just dessa variabler och en regressionsfunktion skapades. Denna process upprepades för samtliga åtta sortiment.

3.5 Analyser

Det utfördes tre analyser och en allmän felsökning kring avvikelser i datamaterialet.

3.5.1 Analys 1 - Jämförelse mellan beståndsregistret och verkligt utfall (pri-filer)

Denna jämförelse gjordes på det ursprungliga datasetet om 30 trakter gentemot deras respektive sammanställningar av tillhörande pri-filer (kap 3.3.2 & 3.3.3). Avvikelsen beräknades genom att subtrahera värdet för varje enskild variabel per VALSID (trakt) från pri-filerna, från motsvarande värde för variablerna i beståndsregistret (tabell 3). För avvikelserna per variabel genomfördes ytterligare beräkningar för att sammanställa avvikelserna.

Beräkningarna som gjordes för varje kolumn var kalkylering av medelvärde, standardavvikelse, max- och minvärde, antal observationer och antalet över- respektive underskattningar.

Tabell 3. Exempel över metodik för att beräkna avvikelse mellan beståndsregistret och verklig utfall från pri-filen. Avvikelserna är uttryckta i enheten som hör till varje specifik variabel, i detta exempel stamantal

Table 3. Example of the method used for calculating deviations between stand data and the actual outcome from the pri-files. The deviations are expressed in the unit which correspond to each variable, in this example number of stems

VALSID	Stamantal		Avvikelse (st)
	Beståndsregister	Pri -fil	
10010	100	90	10
12010	90	100	-10
Medelvärde			0
Standardavvikelse			-
Max			10
Min			-10
Antal observationer			2

3.5.2 Analys 2 & 3 - Skillnad mellan VALS-prognoser och arbetets prognoser samt skillnad mellan arbetets prognostiserade volymer och verkligt utfall (pri-filer)

Det var sex VALSID som utgjorde valideringsdata och metoden för dessa analyser liknar den som användes i analys 1, men beräkningarna var färre och avvikelserna beräknades sortimentsvis och på totala volymer. För att belysa skillnader mellan arbetets sortimentsprognoser, VALS-prognoser och verkligt utfall beräknades medelavvikelse i absoluta tal i % per sortiment (medelavvikelse i absoluta tal innebär att avvikelserna enbart är uttryckta i positiva tal för att hindra negativa och positiva avvikelserna från att ta ut varandra) och avvikelse baserat på totala volymer. Hur väl funktionerna prognostiserade utifrån de olika dataseten indata A och C (funktionsdata och valideringsdata) studerades med hjälp av standardavvikelse, justerad förklaringsgrad och förklaringsgrad R^2 (Chatterjee & Hadi, 2006).

3.5.3 - Allmän felsökning kring avvikelser

Programmet Klippsk har använts för att specialstudera vissa pri-filer som uppvisat uppseendeväckande avvikelser. Även TatukGIS har utnyttjats för att kontrollera avvikelser. Utöver de ovan nämnda program användes ett prototypprogram likt Hpr-analys för att läsa in .shp, .jpeg och .hpr-filer för att studera beståndsgränser gentemot avverkningsgränser mer specifikt för att undersöka arealavvikelser mellan beståndsregistret och pri-filer.

4. Resultat

Ursprunglig uppdelning av materialet på indata A & C utifrån bonitet har antagligen inte varit utslagsivande då variabel ”bonitet” användes en gång i en funktion och sambandet mellan ”bonitet” och sortimentsutfall var således inte starkt.

4.1 Analys 1- Hur väl stämde beståndsregistret med utfallet från avverkningen?

Avvikelse förekom för alla fem variabler men avvikelseernas storlek skiljde sig mellan variablerna. Värdena är uttryckta i totala avvikelser per variabel således inte per hektar då detta inte var möjligt. I ett fall förekom en exakt matchning mellan beståndsregistret och pri-filerna och det var för areal. För variablerna höjd, grundyta och volym var överskattningar i beståndsregistret vanligare än underskattningar (tabell 4). För de två variablerna stamantal och areal var underskattningar vanligare än överskattningar. Variabel areal fick ett lägre antal observerade avvikelser och datasetets storlek var mindre i jämförelse med de övriga fyra variablerna. Antalet överskattningar var störst för variabel höjd, en förekomst i 70 % av fallen medan underskattningar var vanligast för variabel areal där den procentuella förekomsten av underskattningar uppgick till 75 %.

Tabell 4. Sammanställning av avvikelserna för fem variabler, mellan beståndsregistret och data från pri-filer. Avvikelseerna beräknades genom att subtraherar värden i beståndsregistret med värden från pri-filen. Negativa värden är underskattningar och positiva värden är överskattningar i beståndsregistret.

Table 4. Summary of deviations of five variables between stand data and data from the pri-files. The deviations were calculated by subtracting the values of the stand data with the values from the pri-files. Negative values are underestimations and positive values are overestimates of the stand data

	Höjd (M)	Stamantal (st)	Grundyta (m ²)	Volym (m ³ fub)	Areal (Ha)
Medelvärde	0,41	-434,27	-5,57	-57,36	-1,64
Standardavvikelse	1,83	1015,79	54,16	480,41	2,61
Max	3,91	2052,95	142,28	633,82	0,75
Min	-3,86	-2761,48	-151,91	-1415,88	-8,62
Antal jämförelser	30	30	30	30	24
Överskattningar i (%)	70	33	57	53	21
Underskattningar i (%)	30	67	43	47	75

4.2 Analys 2 – Hur skiljde sig Sveaskogs prognoser med VALS gentemot arbetets prognoser?

Genomgående var medelavvikelsen för sortimenten grantimmer och talltimmer i absoluta tal mindre (se kap 3.5.2 för förklaring av absoluta tal) för VALS-prognoserna än för arbetets prognoser. Skillnaden i medelavvikelse mellan de två prognoserna var större för talltimmer än för grantimmer. Vid jämförelse av avvikelser på totalvolym kunde en mindre skillnad mellan de två prognoserna för sortimentet talltimmer utläsas medan skillnaden ökade för grantimmer (tabell 5). Båda prognoserna för grantimmer underskattade den faktiska volymen som föll ut. Alla fyra prognoser hade högre noggrannhet sett till prognoser för totala volymer än för VALSID och sortimentsspecifika prognoser. För talltimmer blev avvikelsen på totalvolym för prognos- volymen låg (underskattning på 4 %).

Tabell 5. Skillnaderna mellan prognostiserade volymer med VALS och arbetets prognostiserade volymer gentemot verkligt utfall (pri-filer) för sortimenten gran- och talltimmer. Volymerna från VALS var en summering av alla sortimentskoder (SSTE) som kan sorteras in i grantimmer och talltimmer, medan volymerna från arbetets prognoser är summeringar av prognostiserade volymer för klen- och normaltimmer för gran och tall

Table 5. The differences between the forecasted volumes with VALS and the reports forecasted volumes against actual outcome (pri-files) for assortments of spruce and pine timber. Volumes from VALS were a summation of all Assortments codes (SSTE) that can be sorted into the spruce and pine logs. While volumes from the rapport predictions are summaries of predicted volumes for small diameter timber and normal timber of spruce and pine

VALSID	Grantimmer (m ³ fub)			Talltimmer (m ³ fub)		
	VALS	Pri	Prognos	VALS	Pri	Prognos
22144	431	549	412	1286	1214	1139
23051	217	199	121	1281	1170	587
25482	660	620	609	82	94	163
25483	550	371	479	454	368	418
25576	655	1141	470	3647	3158	3487
29288	492	482	555			
Total volym	3005	3362	2646	6750	6005	5794
Medelavvikelse för VALSID i absoluta tal % (n=6)	22		28	13		31
Avvikelse för totalvolym % (n=1)	-11		-21	12		-4

Sett till absoluta tal var skillnaden i medelavvikelserna på totalt prognostiserad volym för VALSID respektive prognosverktygen mindre än 3 procentenheter (16,4 % respektive 18,4 %). Medelavvikelsen var mindre för de prognostiserade volymerna från VALS än prognostiserade volymer från arbetets prognosfunktioner. Vid jämförelse mellan tabell 6 och tabell 5 observerades en genomgående högre medelavvikelse i absoluta tal i tabell 5 med undantag för prognostiserad volym talltimmer med VALS, som var mindre. Funktionerna gav

följaktligen bättre prognoser för totala volymer per VALSID än sortimentsspecifika prognoser, undantaget för talltimmer.

Tabell 6. Skillnader mellan VALS-prognoser och arbetets prognoser med avseende på total predikterad volym m³fub per VALSID, gentemot total volym, utifrån pri-filer, per VALSID
Table 6. Differences between VALS forecasts and the rapports forecasts for total predicted volume per VALSID against the total volume per VALSID based on pri-files

VALSID	Total pri-volym	Total prognosvolym	Total VALS-volym
22144	2101	2189	2122
23051	1701	1175	2104
25482	989	1028	1019
25483	926	1234	1292
25576	5402	5512	4905
29288	609	847	742
Medelavvikelse på total volym i absoluta tal % (n=6)		18,9	16,4

4.3 Analys 3 – Hur skiljer sig utbytesprognoserna mot verkligt utfall?

Det justerade R^2 -värdet för varje enskild funktion baserade på indata A & B var genomgående högt, undantaget funktionerna för sortimenten massaved för björk och brännved. Funktionerna för brännved och massaved för löv erhöll låga respektive orimliga värden för R^2 , baserade på valideringsdata (indata C & D) (tabell 7).

Tabell 7. Funktionernas förklaringsgrad utifrån funktionsdataset (indata A) och valideringsdataset (indata C)

Table 7. The functions degree of explanation based on indata A and indata C

	Bränn Gran	Massa Gran	Klen Gran	Normal Gran	Massa Tall	Klen Tall	Normal Tall	Massa Björk
R^2 Valideringsdata	0,07	0,09	0,35	0,20	0,80	0,92	0,93	1,00
R^2 Funktionsdata	0,70	0,93	0,96	0,89	0,99	0,97	0,96	0,72
Justerad R^2								
Funktionsdata	0,59	0,91	0,95	0,86	0,99	0,96	0,96	0,57

Ovan nämnda funktioner erhöll även stora medelavvikelser i absoluta tal. De största standardavvikelserna återfanns i funktionerna NormalGran, NormalTall och MassaTall (tabell 8). Utifrån R^2 -värdet för valideringsdata hade timmerfunktionerna för tall högst värde medan medelavvikelsen var lägst för timmerfunktionerna för gran (tabell 7 & 8).

Tabell 8. Prognosavvikelsernas andel av pri-volym (facit) uttryck i %, utifrån valideringsdataset (indata D), per sortiment. Funktionernas medelavvikelse, i absoluta tal uttryck i %, från pri-volymer per sortiment. Negativa värden är underskattningar och positiva värden är överskattningar av prognoserna gentemot valideringsdata (indata D)

Table 8. The proportion of the deviations, from the pri-volume expressed in %, based on indata D, per assortment. Each functions mean deviation, in absolute terms, expressed in %, from the pri-files per assortment. Negative values are underestimations and positive values are overestimates of the functions predictions

VALSID	Bränn Gran	Massa Gran	Klen Gran	Normal Gran	Massa Tall	Klen Tall	Normal Tall	Massa Björk
22144	550	28	-41	-16	113	-3	-7,4	2389,2
23051	220	-27	-59	-28	36	29	-60,9	27,9
25482	-44	23	7	-4	77	58	74,4	-
25483	86	29	-5	43	208	88	-1,7	-
25576	68	-40	-45	-62	281	68	-0,5	-
29288	23	42	5	18	-	-	-	-
Standard- avvikelse m³fub	26,5	47,3	43,2	252,4	418,7	144,2	281,0	21,2
Medelavvikelse i absoluta tal i %	165	31	26,8	28,4	143	49,1	28,9	1208

Antalet förklarande variabler som utnyttjades vid skapandet av funktionerna varierade mellan tre och fem (tabell 9). Variabeln volym användes inte i funktionerna för NormaltimmerTall eller MassaTall och variabeln stamfelsved utnyttjades inte i funktionen för MassaBjörk.

Tabell 9. De åtta funktionernas utformning utifrån regression i Minitab
Table 9. The eight functions that were created from the regression in Minitab

Funktioner	
Brännved Gran	$80,8 - 0,0225 \text{ StamantalGran} + 0,188 \text{ VolymFUBGran} - 1,12 \text{ GrundytaGran} + 466 \text{ Stamfelsved} - 209 \text{ Medelstamsvolym}$
MassaGran	$32,5 + 0,110 \text{ StamantalGran} + 0,101 \text{ VolymFUBGran} - 1,90 \text{ GrundytaGran}$
Klentimmer Gran	$35,9 + 0,104 \text{ StamantalGran} + 0,201 \text{ VolymFUBGran} - 2,66 \text{ GrundytaGran} - 404 \text{ Stamfelsved}$
Normaltimmer Gran	$- 1165 + 0,692 \text{ VolymFUBGran} - 2,60 \text{ GrundytaGran} + 50,0 \text{ Bonitet} - 1096 \text{ Stamfelsved}$
MassaTall	$59,2 - 4,64 \text{ Höjd} - 0,121 \text{ StamantalTall} + 0,399 \text{ Ålder} + 3,06 \text{ GrundytaTall} + 300 \text{ Stamfelsved}$
Klentimmer Tall	$137 + 0,185 \text{ VolymFUBTall} - 0,865 \text{ Ålder} - 709 \text{ Stamfelsved}$
Normaltimmer Tall	$- 26 - 0,710 \text{ StamantalTall} + 15,7 \text{ GrundytaTall} + 1468 \text{ Stamfelsved}$
MassaBjörk	$530 + 0,136 \text{ StamantalBjörk} + 0,526 \text{ VolymFUBBjörk} - 2,10 \text{ Ålder} - 13,5 \text{ Bonitet}$

De förklarande variabler som användes i flest funktionerna var stamantal, grundyta, volym och stamfelsved. Variablerna höjd och medelstamsvolym förekom en gång i funktionerna och var således de variabler som utnyttjades minst. Anmärkningsvärt var att variabeln volym inte var med i samtliga funktioner och att inte starkare samband mellan bonitet och sortimentsutfall kunde hittas.

Tabell 10. Antalet gånger varje förklarande variabel förekom i någon av de åtta funktioner som skapades under arbetet
Table 10. The number of times each explanatory variable was present in any of the eight functions that were created during the report

	Höjd	Stam antal	Volym Fub	Ålder	Grund yta	Bonitet	Medel stamsvolym	Stamfel sved
Antal förekomster i olika funktioner	1	6	6	3	6	2	1	7

5. Diskussion

Arbetets syfte och de uppsatta målen anses utifrån redovisade resultat vara uppfyllda. Storleken på datasetet blev mindre än vad som ursprungligen planerats, från 50 bestånd till 26 stycken, vilket innebär allmänt större osäkerhet kring resultaten. Det finns flera orsaker till den mindre omfattningen av datamaterialet. Bland annat tekniska aspekter såsom avsaknad av gps-koordinater i pri-filerna men även stora skillnader mellan beståndsregistret och pri-filerna men ingen av orsakerna hade kunnat förutspås och kan endast ses som en lärdom inför framtida arbeten. Metoden bör ses som tillämpbar i framtida studier, dock är några modifieringar befogade för att möjliggöra bearbetning av ett avsevärt större datamaterial på ett mer tidseffektivt sätt.

Det har antagits under arbetets gång att dessa slutavverkningar, som resulterat i en mängd pri-filer, utförts på ett sådant sätt att alla stammar som stått på trakterna har avverkats. Det har alltså inte tagits i beaktande att miljöhänsyn lämnas vid slutavverkning. Miljöhänsyn i detta avseende innebär att alla stammar inte har avverkats och arealen som påverkats inte är densamma som beståndsarealen i beståndsregistret. Antagandet var medvetet och görs på grund av att det anses vara för svårt och osäkert att skatta vad denna hänsyn skulle innebära i värden av beståndsegenskaper.

En felsökning kring förekomsten av orimligt stora avvikelser utfördes och i dessa fall antas förklaringen, efter diskussion med Sveaskog, vara att det efter avverkningen skett en ajourhållning i GISS (Sveaskogs GIS) av trakterna och det resulterat i arealavdrag från trakterna i beståndsregistret på grund av avvikande ståndorter. Ajourhållningen förklarar varför beståndsregistret till exempel för en trakt visade en areal på 1,64 ha och pri-filerna visade en avverkningsareal på över 8 hektar. Dessa stora avvikelser togs bort ur datasetet i och med att bearbetningen av beståndsdata, beskrivet i metoden, kap 3.2.

I resultatet presenteras avvikelserna mellan beståndsregistret och pri-filerna men det är svårt att relatera till avvikelsernas storlek då det inte är uttryckt i exempelvis ”per hektar” eller i ”procent av totalen”. Orsaken till att avvikelserna inte är uttryckt per hektar är till följd av avvikelser i areal mellan beståndsregistret och pri-filerna och ibland helt avsaknad av areal från pri-filen då gps-koordinater inte var registrerade. På grund av dessa avvikelser ansågs det vara tydligare att göra beräkningar på totala summor och redovisa avvikelser i totala mängder för de enskilda variablerna.

Arbetes mål att skapa åtta funktioner lyckades men vissa funktioner lyckades inte så väl förklara variationen i responsvariabel och erhöll således ett lågt värde på förklaringsgraden (R^2). Enligt regressionsteori skall det justerade R^2 -värdet användas vid jämförelse mellan funktioner och det kan även användas för att jämföra en funktions resultat utifrån olika dataset. I detta arbete var en minskning i förklaringsgrad väntad när funktionerna användes på valideringsdata eftersom det är ett mindre dataset och att funktionerna inte är baserade på datasetet. Dock var den mycket stora minskningen för samtliga granfunktioner inte väntad, men förklaringen ligger i det låga antalet trakter som utgör valideringsdatasetet. I och med det

låga antalet får varje enskild avvikelse stor påverkan på förklaringsgraden. Något som styrker detta är att trots den låga förklaringsgraden så är medelavvikelsen för funktionerna klen- och normalgran mindre än för klen- och normaltall trots det omvända förhållandet med avseende på förklaringsgrad (tabell 7 & 8). Även förklaringsgraden för MassaBjörk exemplifierar problematiken med ett lågt antal observationer då funktionen erhåller R^2 -värdet 1 för valideringsdata medan medelavvikelsen i absoluta tal är ca 1200 %. Funktionernas användningsområde är snävt eller obefintligt till följd av det låga antalet trakter i funktionsdata men visar att det finns potential för metoden.

Vid jämförelse med Sveaskogs nuvarande prognoser skapade via VALS var arbetets funktioner sämre (tabell 5 & 6) men sett till totala volymer var arbetes prognostiserade volymer endast 2,5 procentenheter sämre (tabell 6). Vid en djupare analys av timmervolymer (tabell 8) uppstår en avsevärt större skillnad mellan de olika prognosernas medelavvikelser. VALS var mer noggrann i sina prognoser för tall vilket kan ses som ett tecken på att talltimmer är lättare att prognostisera än grantimmer. Det är även viktigt att poängtera att funktionerna är baserade på dataset som har specifika spridningsområden för de enskilda variablerna och funktionerna får således ett snävt användningsområde och värden utanför spridningsområdet på funktionsvariablerna bör användas med stor försiktighet.

Variabeln bonitet som inledningsvis bedömdes vara en viktig variabel och hade en stor plats i fördelningen av trakter mellan funktionsdata och valideringsdata visade sig vara mindre viktig. Funktionernas utformning visar tydligt att det inte fanns starka samband mellan bonitet och de sortiment som fallit ut (tabell 10).

För att lyckas med regressionsanalys som metod för prognostisering av sortiment krävs förklarande variabler som har ett starkt samband till responsvariabeln. De variabler som förklarade variationen hos responsvariabeln bäst var stamfelsved, stamantal, volym, och grundyta (tabell 10). Stamfelsved är en variabel som generellt sett inte finns med i beståndsregister och det är för att den beräknas utifrån avverkade stockar. Den visade sig vara det mest utnyttjade variabeln, då den användes i sju fall av åtta möjliga. Stamfelsvedsvariabeln beskriver skadefrekvens i trakter och beskriver således en variabel som enligt Barth et al. (2013) och Samuelsson (2005) är mycket viktigt vid prognostisering av sortiment. Utöver stark korrelation mellan förklarande variabler och responsvariabel krävs det bra noggrannhet i indata och att omfattningen av data är tillräcklig. Det är även ett rimligt antagande att anta att vissa variabler påverkar mer än andra och inte nödvändigtvis linjärt utan exponentiellt. Vilket inte har utforskats inom denna rapport. Dessa fyra viktiga aspekter har delvis varit bristfälliga i detta arbete.

Alla skördare registrerade inte gps-koordinater för träden och det förekom även avvikelser inom rapporteringsstrukturen för pri-filerna. Exempel på detta är när numreringen för stamantalet började om för varje pri-fil. Utöver ovan nämnda problem förekom det avvikelser kring när pri-filer rapporterades in, t.ex. en fil per dag eller flera. Alla dessa avvikelser försvårar databearbetningen. En tydligare standard kring dessa vore till gagn för Sveaskog vid framtida studier.

Inför framtida studier på området, specifikt för Sveaskog, är det viktigt att poängtera att det finns vissa aspekter med det nuvarande systemet över hur pri-filer är kopplade till deras ursprungliga trakter (med VALSID) som kan förbättras. Med dagens system kan flera trakter vara kopplade till ett VALSID och således kan flera pri-filer vara kopplade till ett VALSID. I samtliga fall där det är fler än en trakt kopplad till ett VALSID, innebär det ett merarbete att isolera vilka pri-filer som tillhör varje specifik trakt. Vidare bör förfarandet vid lokalisering av gamla pri-filer och historisk beståndsdata (data över trakter innan de avverkades, data som pri-filerna bygger på) standardiseras för på så vis tidseffektivisera datainsamling.

I början på diskussionen förklaras antagandet kring avsaknad av miljöhänsyn vid avverkningarna inom arbetet och dess konsekvenser. Detta är en felkälla, för studiens prognoser men att belysa just den punkten inom prognostisering av sortiment är viktig och är ett område som det krävs mer arbete inom. Vad blir konsekvensen av miljöhänsyn inom slutavverkning i form av förlorad gagnvolym? Hur mycket påverkar den sortimentsprognoser? Hur skall detta införlivas i prognostisering på ett bra sätt?

Under arbetets gång användes programmen Klippsk, Biobränsle, Hpr-analys, TatukGIS och ett prototypprogram som kan läsa in .shp, .hpr och .jpeg-filer. Programmen har varit outhålliga, men det finns fortfarande förbättringsmöjligheter. Nordström et al. (2010) framhäver att vidareutveckling av systemlösningar är nödvändiga inför framtidens arbete med skördardata och erfarenheterna från detta arbete stödjer deras påstående.

Slutsatser

- Pri-filer och hpr-filer har potential för att skapa bra och tillförlitliga utbytesprognoser men det krävs mer forskning på området
- Smarta systemlösningar och tydliga förhållningsätt i tillvaratagandet av skördardata är nyckelfaktorer inför framtida studier
- De skapade funktionerna i detta arbete bör inte användas skarpt då de är baserade på ett för litet datamaterial och de nuvarande prognoserna skapade av VALS är bättre
- Studien är ytterligare ett bevis på att noggrann indata är avgörande för att skapa bra prognoser och fortsätter vara en stor utmaning för skogsbranschen

Referenslista

Artiklar och böcker

Arlinger, J., Möller, J. & Sondell, J. 2003. A description of pri-files. SkogForsk. Arbetsrapport nr 599. Uppsala.

Barth, A., Möller, J., Wilhelmsson, L. & Arlinger, J. 2013. Bättre utbytesberäkningar med laserdata. Skogforsk. Resultat nr 1. Uppsala

Cernold, Å. 1981. Utbytestabeller för rotstående skog. Norra Sverige. Centrala sågverksföreningen. Falun

Chatterjee, S. & Hadi, A. 2006. Regression analysis by exempel. 4. uppl. John Wiley & Sons. New Jersey

Eriksson, M. 2010. Prognostisering av sortimentsutfall från stående skog med hjälp av befintlig data. Anpassat till Sveaskog Norrbottens planeringsprocess. Umeå. Institutionen för skoglig resurshushållning, Sveriges lantbruksuniversitet. Arbetsrapport 270

Nordström, M., Wilhelmsson, L., Arlinger, J., & Möller, J. 2010. Skördardata kan ge industrin viktig förhandsinformation från skogen. Skogforsk. Resultat nr 21. Uppsala.

Ollas, Rune. 1980. ”Nya utbytesfunktioner för träd och bestånd”, Ekonomi nr. 5. Forskningsstiftelsen Skogsarbeten, Stockholm

Samuelsson, J. 2005. En jämförelse mellan två datorprogram för utbytesberäkningar. Umeå. Institutionen för skoglig resurshushållning, Sveriges lantbruksuniversitet. Arbetsrapport 152

Ståhl, G. & Wilhelmsson, E. 1994. Planering av skogsbruk. Kurslitteratur. Sveriges lantbruksuniversitet, Institutionen för skoglig resurshushållning och geomatik. Umeå.

Webbsidor

Sveaskog. (2013). Sveaskog i korthet 2012 <http://www.sveaskog.se/sv/om-sveaskog/bestall-informationsmaterial/> [24/9-2013]

Minitab. (2013) <http://www.minitab.com/> [29/10-2013]

Skogforsk. (2013) <http://www.skogforsk.se> [4/11-2013]

SDC. (2014) <http://ny.sdc.se/default.asp?id=1007> [19/2-2014]