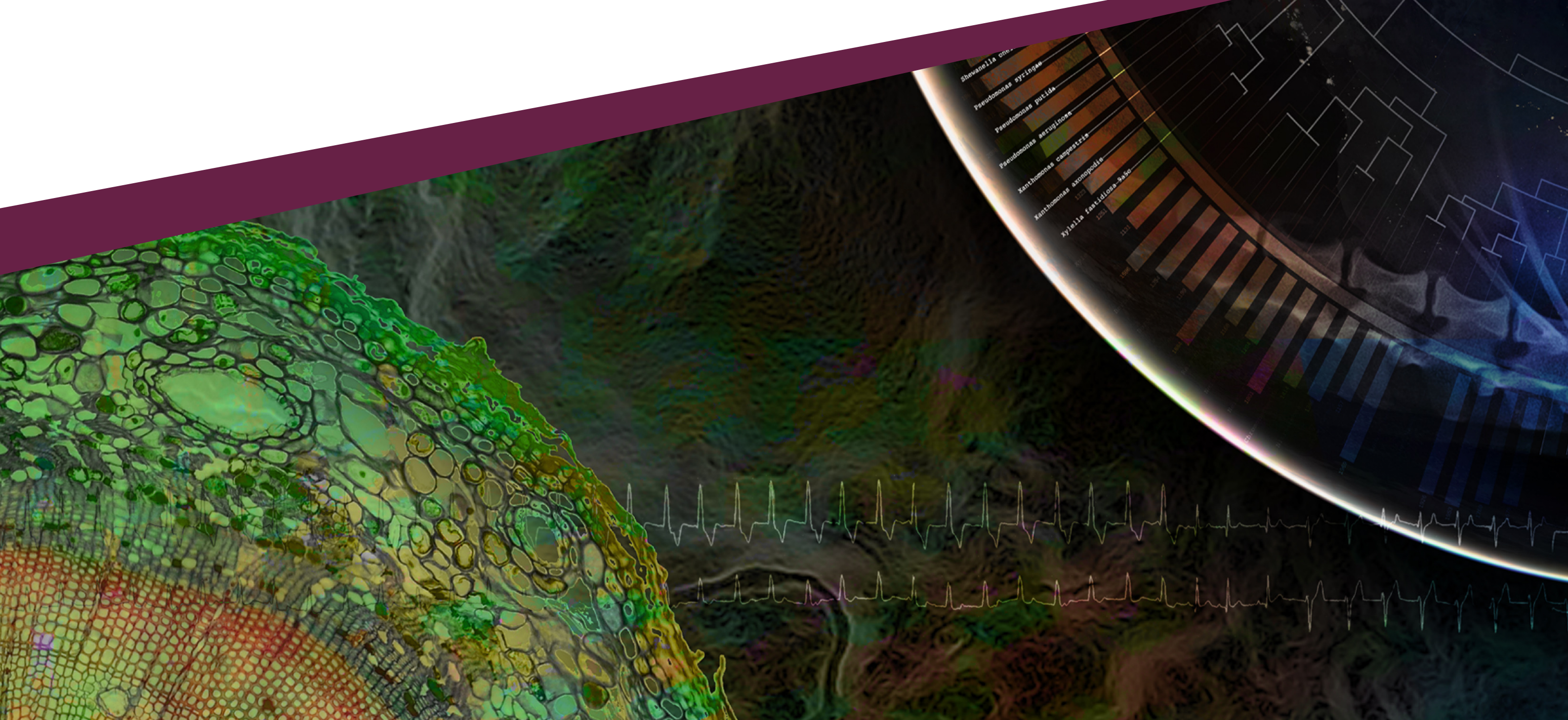
Suitable tract bank size: exploration & estimation

Lämplig traktbanks storlek: Utforskning & skattning

Anders Rowell

Examensarbete • 30 HP
Sveriges lantbruksuniversitet, SLU
Skogsfakulteten, Institutionen för skogens biomaterial och teknologi
Jägmästarprogrammet
Rapport från Institutionen för skogens biomaterial och teknologi, 2022:7
Umeå 2022



Suitable tract bank size: exploration & estimation

Lämplig traktbanks storlek: Utforskning & skattning

Anders Rowell

Supervisor:	**Dan Bergström, Swedish university of Agricultural Sciences, Department of forest biomaterials and Technology**
Assistant supervisor:	Ola Lindroos Swedish university of Agricultural Sciences, Department of forest biomaterials and Technology
Examiner:	Dag Fjeld, Swedish university of Agricultural Sciences, Department of forest biomaterials and Technology

Credits:	60
Level:	A2E
Course title:	**Masterarbete I skogsvetenskap, A2E – Skogens biomaterial och teknologi**
Course code:	**EX0954**
Programme/education:	**Jägmästarprogrammet**
Course coordinating dept:	**Department of Forest Biomaterials and Technology**
Place of publication:	**Umeå**
Year of publication:	**2022**

Keywords:	Tactical-level, strategic-level, harvest management, harvest planning, inventory theory, forestry, tract bank.

Swedish University of Agricultural Sciences

Faculty of Forest Science

Department of Forest Biomaterials and Technology

Abstract

The collection of stands planned and ready for harvest, also called the tract bank, gives the harvest manager a selection of stands to choose from while scheduling harvests. Too few and the harvest manager might be forced to make bad decisions; too many bring additional unnecessary costs. However, little research has been done on adapting the tract banks to regional and source-dependent uncertainty, such as contracted harvests or harvests on company-owned land.

This study aims to, on a strategic level, per geographic region and source, further the understanding of the tract bank and the uncertainties that effect it. By estimating suitable tract banks using inventory theory, considering outcome deviation and the seasonal differences between planned and harvested stands. The estimates are complemented by regression modelling of the utilisation of stand seasonality classifications, quantified past tract bank sizes and stand storage time.

The study finds that forestry is more seasonally constrained northwards in geography compared to the south and in company-owned stands. It also finds that estimated stand volume is systematically underestimated too varying degrees. Moreover, it contributes with quantified regional storage time and regional mean coverage time (months) of past tract banks of a Swedish forest company: North 13,1; Middle 9,5; and South 5,8. Comparably the suitable tract bank sizes estimated in this study are: North 9,5; Middle 5,9; and South 7,1. These results are relevant to those that work with the tract bank, harvest management and harvest planning and could be used in further research of the tract bank.

Keywords: Tactical-level, strategic-level, harvest management, harvest planning, inventory theory, forestry, tract bank.

Sammanfattning

Trakter som är färdigplanerade och redo för avverkning, också kallade traktbanken, ger produktionsledaren ett urval av trakter att välja bland vid schemaläggning av avverkningsresurser. För få trakter kan leda till tvingande dåliga beslut, men för många kan leda till onödiga kostnader. Trots detta finns det ont om forskning som tittar närmare på hur traktbanksstorleken bör anpassas efter geografi och ursprung – ursprung som i köpta avverkningsrätter och egen skog. Den här studien ämnade att på en strategisk nivå, per geografi och ursprung, utveckla vad vi vet om traktbanken och dess varierande osäkerheter.

I studien skattades lämplig traktbanks storlek från årlig variation i månatlig utfallsavvikelse, och variation i utnyttjandet av trakter med viss säsongsklassifikation, med hjälp av lagerteori. De skattade traktbankerna kompletterades med regressionsmodellering av traktutnyttjandet med viss säsongsklassifikation och kvantifierade tidigare traktbanks storlekar samt trakters tid i traktbanken.

Studien fann att skogsbruket är mer begränsat av årstid i den egna skogen och i nordligare geografi jämfört med sydlig geografi. Den fann även att den planerade volymen generellt underskattas jämfört med den industriinmätta, vilket varierar något per geografi och ursprung. Därtill bidrar studien med kvantifierade täckningstider: Nord 13,1 månader; Mitt 9,5 månader; och Syd 5,8 månader. Dessutom jämförbara siffror på lämplig traktbaks storlek per region som skattat i den här studien: Nord 9,5 månader; Mitt 5,9 månader; och Syd 7,1 månader. Dessa resultat är av nytta till de som arbetar med traktbanken, avverkningsplanering, avverkningsledning och kan användas i framtida studier gällande traktbanker.

Preface

This 60-credit master thesis was done as my final project in the forestry program at Swedish university of agriculture in Umeå, in collaboration with Holmen skog. The target group of this study is fellow forestry academics that have an interest in the tract bank, forest planning and forest operations. The 60 credits in the course have enabled me to seek out knowledge, skills and methods that otherwise would have been barred in the courses of the program. Knowledge, skills, and methods that that made this study possible and opened doors after.

The study was written in collaboration with Holmen skog and I would like to thank Miriam Nordh and Jonas Eriksson for giving me that opportunity. I would also like to thank you for your ideas and insights.

Thank you Mattias Gustafsson (Holmen skog) for making the access to data go smoothly.

Thank you, the Swedish Meteorological and Hydrological Institute, and thank you, the Food and Agriculture Organization for providing data under the creative commons.

Thank you, Hilda Edlund, for sharing your algebraic and statistical knowledge.

Thank you, Anabelle for your knowledgeability in meteorology, climatology, and mathematics. But also, for normalising my appreciation of academia and supporting me in this work.

Thank you, Dag Fjeld, for guidance in what paths to tread and not to tread upon.

Thank you, Ola Lindroos, for your quick eye for detail, expertise, and supervision.

I would especially like to thank you Dan Bergström, for without you I would not have had the will to write a 60-credit master thesis and make the most of it. Thank you for great mentorship, your time, helpful discussions, letting me explore ideas and encouraging the joy of learning.

I hope you enjoy reading, Anders Rowell, Uppsala 2022 July 5th

Table of contents

List of tables

List of figures

Abbreviations

TB	Tract bank
CTD	Coverage time difference
r	Region
s	Source
c	Seasonality classification
m	month
y	Year
i	The i stand
CT	The coverage time of a tract bank

1. Introduction

The forestry sector has a central role in the bioeconomy in many regards. Providing raw material for established wood-based processes and new value chains, but also in the aspect of bioeconomy where sustainability, biodiversity and other ecosystem services are valued (Bugge et al. 2016).

To conduct sustainable and profitable forestry, the large Swedish forest companies do extensive planning in all branches of their value chains, from planning forest operations to planning the sales of refined wood products in both the short and long term. The goals, plans and operations are structured into three general hierarchical levels, strategic, tactical, and operational (Söderholm 2002; Church 2007). The hierarchical model results from the wood supply chains complexity, where the levels become a simplification of a complex system into more manageable levels (Church 2007).

The hierarchies within the forest company can generally be divided into two parallel hierarchies, the forest planning and the wood supply (fig. 1). Forest planning generally works within a longer time frame of strategically 100 years to operationally day-to-day tasks, while the wood supply works within a shorter timeframe of 5 years strategically to operational tasks. The hierarchical relationship between the two can also be described as a tug of war, a push and pull. Where the push would mean that the forest planning has a supply of wood that is pushed to the industry, and the pull is a demand of wood where the wood supply demands wood and certain assortments from the forest planning (Fjeld & Dahlin 2017; Church 2007; Nilsson et al. 2013; Holappa Jonsson 2018; Carlsson & Rönnqvist 2005).

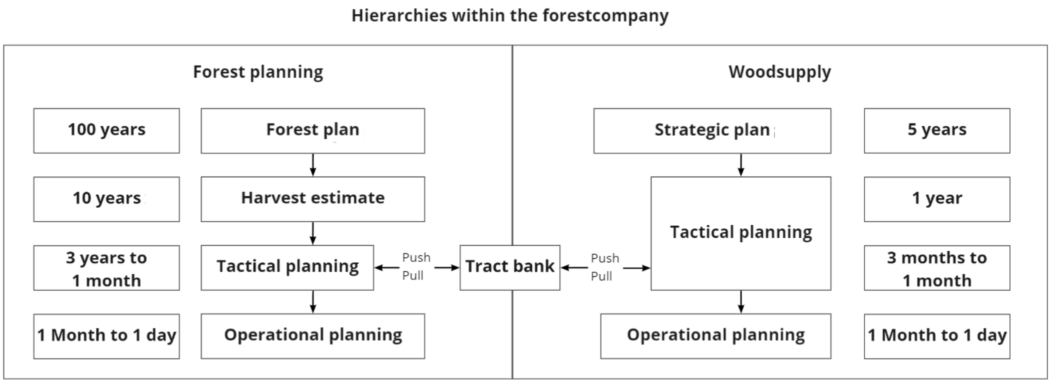


Figure 1. The chart shows a general hierarchical structure of a forest company, where the arrow visualises a higher hierarchical level and the years/months a general timeframe for the hierarchical levels. When the arrow points both ways, the hierarchical relationship is dynamic, not one-way.

The strategic level determines goals, develops resources, and provides a framework for the operational level. The strategic forest planning level has a long-term planning horizon of approximately 100 years for all stands in all age classes. The 100-year planning horizon can be likened to the rotational period of a stand in Swedish forests, which is also approximately 100 years. The long-term forest plan is implemented through the harvest estimate, which estimates wood volumes for the coming 10-year period. As the ten years are aggregated data over large areas, it cannot be directly implemented for operational use without further planning on lower hierarchical levels.

The operational level is executed locally in the short term, which is in the time frames days, weeks, to months – it is especially stark contrast to the 10-year long-term framework of the harvest estimate. The operational level seeks to solve short-term problems and goals, trusting the higher hierarchical levels bigger picture. Examples of the operational level would be the operational planning, execution and control of the transportation or production plans. These plans specify wood volumes to be harvested and transported to industry within the month.

The tactical level translates the strategic level to be operational by prioritising and clustering it into a midterm time horizon. The midterm plans can then be planned, executed, and controlled on an operational level to meet the strategic goals. In the case of forest planning, the output of tactical forest planning is the tract bank (TB). It results from the tactical level prioritising and clustering the harvest estimate into operationally executable stands. In other words, the TB is the collection of stands where harvest planning is finished but has yet to be harvested. The TB is also the point where the two hierarchies formally interact. The stands in the TB limit the decisions and demands defined by wood supply in its higher hierarchical levels. However, the wood supply also circumvents the limitations of the TB to some degree by scheduling stands not in the TB for harvest (Fjeld & Dahlin 2017; Church 2007; Nilsson et al. 2013).

Since the TB is the focus of this study, it is vital to understand how it is designed and utilised. Essentially harvest planning and harvest management are work processes directly connected through the TB in the wood supply chain - in how they generate and use the TB. However, they also interact with other work processes, enabling, constraining, and setting performance goals. For example, the harvest manager works closely with transport managers, who manage how and when the roundwood should be moved from the forest roadside storage to the industry (fig. 2). In addition, however, the harvest manager is responsible for providing the agreed-upon wood volume on an assortment level at the roadside.

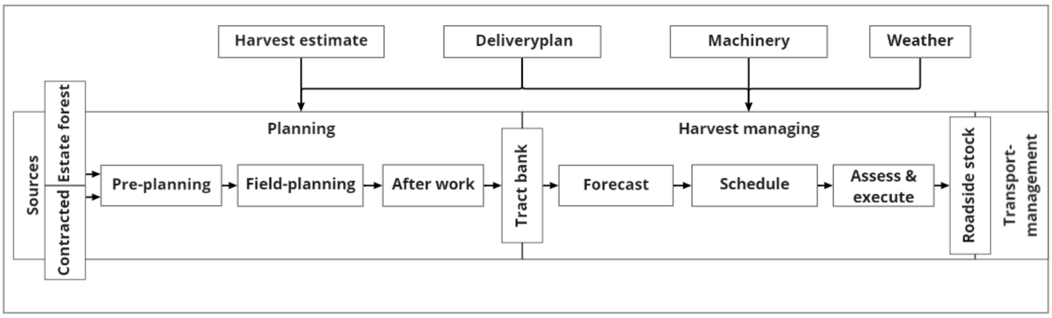


Figure 2. The flowchart showcases the work processes harvest planning and harvest managing. The flowchart also visualises how stands from the two sources move through the wood supply chain and a selection of external influences that affect the work processes. The flowchart is a compilation of information from Nilsson et al. 2013, Wildmark 2014 and Willén et al. 2021.

The planning process determines how much and what qualifies for harvesting from its two sources, contracted and estate forest stands (fig. 2). It should be said that not all forestry companies, forest owners' associations or wood supply organizations have a considerable part of their harvest from their own estate. However, for those that do, the estate is a significant asset as it enables agility in the supply. As synthesised by Shuva Gautam (2013) agility in wood procurement is determined by the three capabilities responsiveness, flexibility, and timeliness: responding to external change, changing with the situation, and still delivering on time. These capabilities are then enabled by flexible logistics, integrated planning, collaboration, information technology and flexible supply. The estate gives a flexible supply and enables control over the wood flow from the forest to the industry and security of supply (Andreasson 2018), which leads to less required flexibility (Erlandsson 2016). On the other hand, when contracting a stand, the same level of security is lacking as other parties are often interested in contracting that stand. That competition does not directly limit harvests from the company's estate;

instead, the long-term harvest estimate aims to keep the wood flow steady and can even out the uncertainty associated with contracting stands.

Furthermore, the harvest planning process consists of three main subprocesses pre-planning, field-planning, and finishing the plan (fig. 2) (Willén et al. 2021). The purpose of pre-planning is to make field planning more effective; in this subprocess, different types of information are interpreted and analysed in a geographical information system (GIS). During the field planning, stand information is assessed, and the planner uses physical marking bands to convey borders and consideration areas to the harvest team. When the plan is finished, marked areas and points of interest are digitised in a GIS, and the stand is marked in the company's internal system as "planning finished". When the stand has been marked as "planning finished", it becomes available for the harvest manager to schedule into the operational plan (Willén et al. 2021) and is now part of the TB.

The harvest manager is responsible for the harvesting teams' quality of work and produced wood volume. The harvest management of the harvesting teams can be condensed into three subprocesses, forecasting, scheduling, and creating the necessary documents for harvest (fig. 2) (Epstein 2007; Wildmark 2014). There is a control of the harvesting teams during the forecasting, and the delivery plan is analysed to compare with previous forecasts. When reality and the original plan have come too far apart, then that must be considered in the next step. In the scheduling, the harvest manager seeks out stands in the TB that could fulfil the delivery plan and does a preliminary routing for what order these stands are to be harvested and by which harvesting team. Finally, when creating documents necessary for harvest, the preliminary plan from the scheduling is assessed (Wildmark 2014). The details vary to some extent and can be described through a gradient scale of centralised to decentralised harvest management. The decentralised has more focus on the movement of the harvester teams, while the centralised has a greater focus on fulfilling the delivery plan (Wildmark 2014).

Furthermore, the harvest planner and harvest manager also consider factors other than their inputs (sources and TB). For example, they consider that the harvesting teams must be provided with the agreed-upon wood volumes clustered spatially and temporally for them to be sufficiently satisfied (fig. 2) (Erlandsson & Fjeld 2017). They also must consider suboptimal routes to ensure all stands in the long-term plan are harvested and that the stand composition of the TB can meet demand throughout the year. Generally, the harvest planner seems to follow the classical hierarchy, prioritising the long-term forest plan over the delivery plan. On the other hand, the harvest manager deviates from the classical hierarchy and prioritises the delivery plan over the forest plan. Occasionally, this leads to the harvest manager scheduling harvests not in the TB to fulfil the delivery plan, deviating from the original long-term planning. However, it is unclear to what extent this is done or to what extent this is accounted for on a higher management level (Nilsson et al. 2013).

Furthermore, it has been shown that the demand is more likely to be met with a large TB. Also, the available options when scheduling increases with more stands in the TB and possibly lower costs. Industry personnel (Holmen, SCA and Sveaskog) have also commented that it’s easier to perform high-quality work with a large TB, and that most harvest managers (76%) have been shown to want a larger TB than they currently have (Nilsson et. al 2013). However, a survey of Holmen in 2003 found that for Holmen-specific respondents (16 districts/respondents), the average tract bank (available for harvest managers) was 17 and 7 months for north and south, respectively (33 % and 74 % private contracts) (Larsson 2003). Again, the interesting result was that 80 and 36 % of respondents considered the tract bank coverage time sufficient in the south and north, respectively.

On the other hand, a small TB can function just as well as the large one, but then there are greater requirements for the composition to be close to what the industry demands. But some bottlenecks are hard to account for, for example, not having enough suitable stands for spring or fall harvesting. The harvesting manager knows when a stand is suitable for harvest, as all stands planned for harvest have their terrain and road accessibility classified by seasonality. So, with dialogue and planning, a good composition for each season is possible, but there are uncertainties as to how much is needed, when seasons change and what weather will occur (Staland 2001; Nilsson et al. 2013; Jacobsson 2005; Renström 2008).

The challenge of adapting the stand utilisation seasonally in a district was introduced by Renström (2008). However, the question remains of how that seasonality varies between regions, as Sweden have different climates depending on geography. The south of Sweden has a warm humid continental climate, while the north has a subarctic climate (climatedata.org 2022). The two climates have in common that they are both cold temperate climates (coniferous). This means that the year's coldest month has a normal mean temperature below -3°C, and the warmest month has a normal mean temperature above 10 °C and precipitation year-round (SMHI 2022b). However, they differ because the subarctic climate has less than four months, with normal mean temperatures above 10 degrees Celsius, and the warm continental climate has more than four (SMHI 2021).

Explaining the seasonal and yearly variations is a complex task, but Machine learning (ML) is increasingly used in ecology and forestry science. ML is for example, used in hazard assessment and prediction, and soil trafficability forecasting (Jones & Arp 2019; Liu et al 2018). The complex nature of explaining the decision-making regarding seasonal utilisation might make ML a good option in describing the variation. There are more factors than the weather that effect how we utilise stands, such as the wood market, that might make us more inclined to procure a certain assortment—potentially harvesting many more stands of a suboptimal trafficability classification to meet that demand.

The TB is seen as the stock between the planning and harvest managing processes, it is described in how many months or years it would take to harvest the entire TB with the current harvesting resources. However, when referring to the TB, it is usually separated into two parts, where the stands source separates them. The contracted stands are a part of the contract bank, whilst the stands from the forest estate are a part of the stand bank. They are calculated separately by their respective wood volume and yearly or monthly harvesting wood volumes. The industry TB norms are to have two years coverage time of harvest in the stand bank and no more than eight months coverage time in the contract bank. It is unclear exactly where the two-year stand bank norm comes from, but it is established in the industry, and it is, for example, mentioned in the interviews conducted by Nilsson et. al 2013. The contracted volumes are limited not to exceed eight months because private forest owners have shown displeasure in previous studies when the harvesting of their stands are on hold longer than this (Roth 2010).

The interviews by Nilsson also indicate that some stands in the TB are not ready for harvest, making the absolute TB size as it is referred to by Nilsson et al. (2013), smaller than it seems. Consequently, the TB should be considered smaller than the actual TB as some stands are prioritised over others, entail more work, and can be neglected entirely. For example, remotely located stands are less likely to be harvested, contracted stands are generally prioritised, and some stands are in the TB for many more years than the rest. These stands that spend a long time in the TB risk getting planned for harvest more than once. These stands risk possibly diverting the operational plan away from what has been planned on higher hierarchal levels, such as the tactical and by extension strategic (Nilsson et al. 2013).

As the TB is the stock between the planning and harvest management, inventory theory should be able to estimate a suitable TB size by region. If developing a model for optimal decision support for dimensioning the TB, it should probably be done on the tactical level with a district or harvest manager perspective, as it is on this level that the decisions are made. For that, a dynamic programming approach could be advisable to in steps model for a single production unit and period at a time. By looking at the cost in meeting demand going into a period with different stock levels, consumption and acquisition, examples of such approaches would be Nilsson (2008), Aggeryd (2009), Staland (2001) & Jacobsson (2005).

Suitably stock size determined by inventory theory can be divided into the cycle and the security stock. The cycle stock is the wood volume that is expected to be needed throughout the year, and the security stock (SL) is added to handle the variation of what is required (Lumsden et al. 2019). This can be expressed in a more mathematical sense, as is shown in equations 1 and 2.

$$\text{stock} = \text{cycle stock} + \text{security stock} \tag{1}$$

$$\text{security stock} = Z\sqrt{\mathrm{LT} \cdot \sigma_{\mathrm{D}}^{2} + (\sigma_{\mathrm{LT}} \cdot \mathrm{D})^{2} + \sigma_{\mathrm{Reg}}^{2}} \tag{2}$$

Where D is the demand and LT is the lead time. σ_{D}^{2} is the variance of the demand and σ_{LT} is the standard deviation of the LT. The $\sigma_{\mathrm{Reg}}^{2}$ is the variance of what is thought to be in the stock and the actual stock. Z is the service level, or the number of security stocks needed to reach desired delivery precision.

This study aimed to, on a strategic level, per geographic region and source, further the understanding of the tract bank and its uncertainties. The study quantified past tract bank dimensioning, storage time, outcome deviation, stand utilisation, and prerequisites of a Swedish forest company. Moreover, it estimated a suitable tract bank dimensioning using inventory theory from three perspectives: seasonality, region and by source. Firstly, by per source quantifying past tract bank dimensioning, outcome deviation, stand storage time, and the utilisation of stands with different seasonality classifications; secondly, by developing a model for estimating suitable TB size for two sources; Thirdly, by presenting the effects of uncertainty depending on geography on the dimensioning of the TB. The following research questions were answered to reach the aim of the study:

- How large has the TB been in the observed years, and how does it differ between regions?
- For how long are stands in the TB?
- How much do the estimated stand volume and the stand volume measured in the industry differ monthly?
- What patterns and variations are there in the seasonal utilisation of stands?
- How large would a suitable TB be based on the seasonal variations and uncertainty?
- What data can be used to predict the utilisation of stands with specific classifications of seasonality, and how does that predictability differ between region, source, and seasonality?

This study assumes: A constant industry demand, yearly harvesting capacity, yearly harvest planning capacity and composition of the tract bank.

2. Materials & methods

2.1. Industry case with Holmen Skog

Holmen is one of Sweden's largest forest owners, with 1.3 million hectares of forest land, of which 1 million are productive. Within the Holmen concern, Holmen Skog's main tasks are managing the estate and procurement from private forest owners and trading with the wood supply. Consequently, Holmen skog is responsible for contracting the right to harvest other forest owners' lands and supply biomass to their own and the industries they are committed to (Holmen 2020).

Holmen Skog is structured in three regions, North, Middle, and South (fig 3).

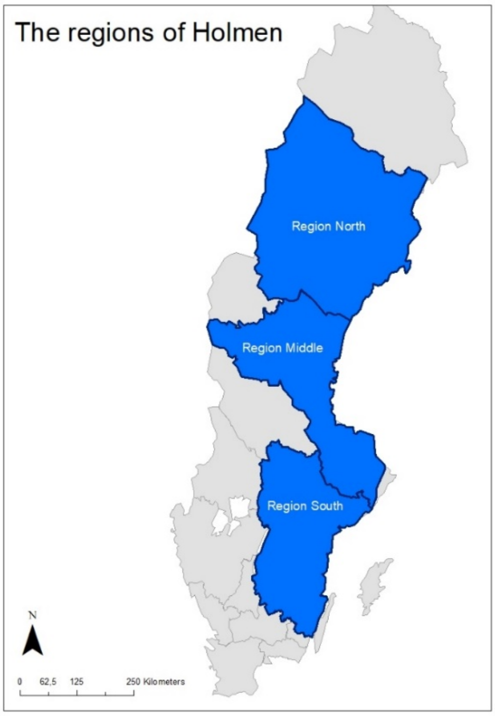


Figure 3. The map shows Sweden in grey and the three regions of Holmen as blue.

Holmen has a larger part of its forest holdings in the north, which handles fewer assortments. In the South, they have a smaller share of forest estate and handle a larger number of assortments. The middle region, has roughly equal shares of estate and contract harvesting. The stands that are part of the estate are characterised by long-term planning, more qualitative data, and control such that the stand and road history is known. On the contrary, contracted harvests tend to be characterised by short-term planning, less precise data, and stand-specific constraints. Constraints that include the forest owner contracting with demands on what season or harvesting team should perform the harvest.

2.2. Workflow overview

The following 7 steps were used in exploring and estimating uncertainty in different regions TB (also see steps 1-9 in flowchart figure 4):

The first step was acquiring and pre-processing the data. The data came from three sources, the host company, SMHI's open data (SMHI 2022a), and FAO's (FAO 2022) documented wood timber prices. The pre-processing consisted of formatting the data so all sources could be analysed as a single data frame in the software environment R (R Core Team 2021). By classifying each stand by current administrative boundaries, determining what date the stand entered the TB and left it (if it has left) and merging road and terrain classifications into a seasonality classification.

The second step was calculating past monthly TB for the observed period, 2013-2020.

The third step was calculating the stand storage time in the TB in months for the observed period.

The fourth step was calculating the monthly outcome deviations for the observed period, 2013-2020.

The fifth step was to estimate seasonal indexes for harvesting and planning per seasonality classification and region.

The sixth step was estimating a suitable tract bank size. That was done in two parts; first, the cycle stock was estimated from the mean seasonal indexes. Then a security stock was estimated based on the variance of the seasonal indexes and the monthly outcome deviation.

The seventh step was training randomForest models per geography, source and seasonality classification to predict the seasonal index variation.

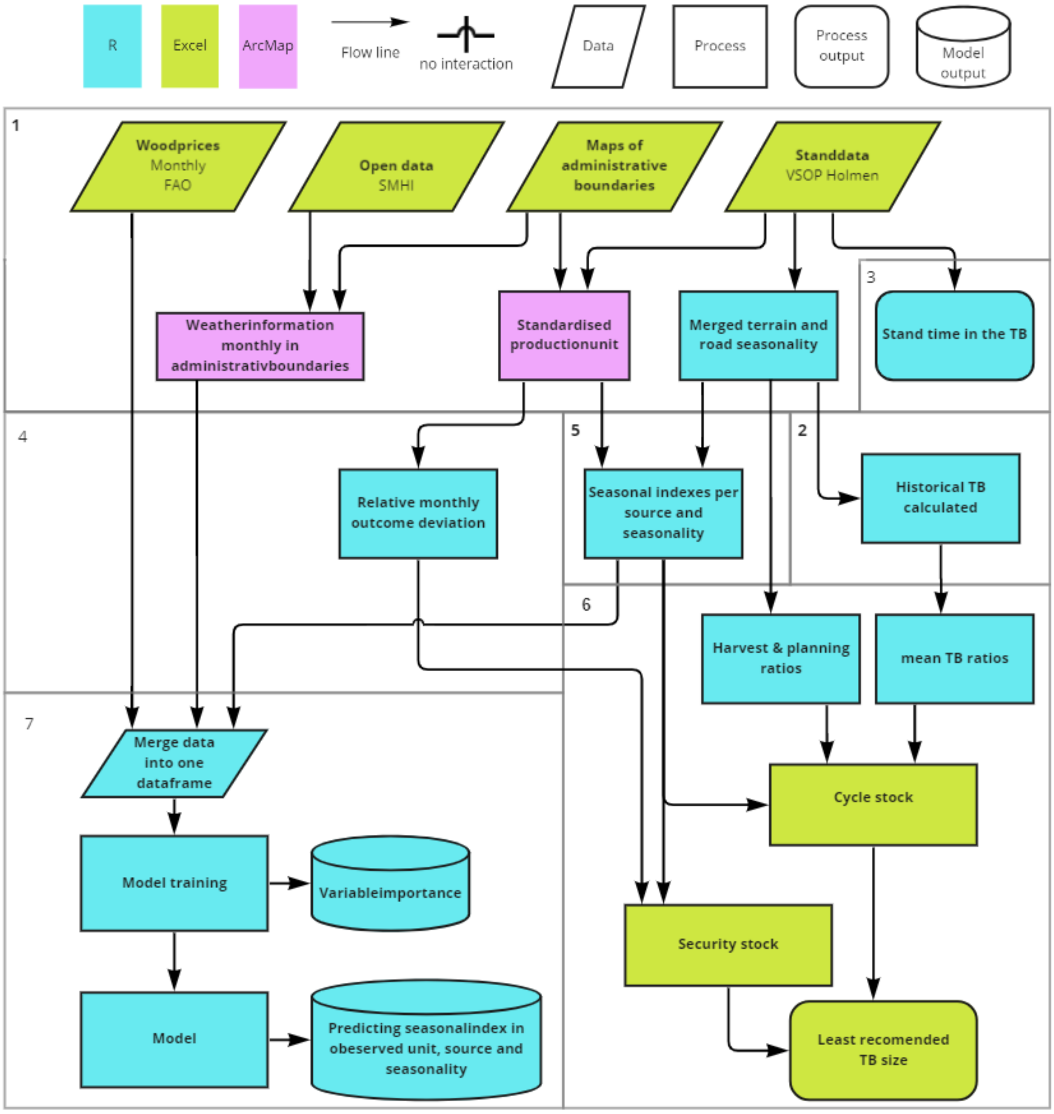


Figure 4. The flowchart shows an overview of the workflow's data input, processes, and outputs. The processes are square boxes coloured after the software is used, pink is ArcMap (GIS), blue is R, and green is Excel. The different shapes indicate what it is, a quadrate is a process output, rhomb's are data frames, and the circular box is a model output. All shapes are also numbered after what step in the process they were performed. 1) acquiring the data three three sources, Holmen, SMHI, and FAO. Then cleaning that data for further processing. 2) pre-processing the data, by standardising location, dates and merging road and terrain into one seasonality classification. 3) Calculating the monthly TB's for 2013-2020. 4) Calculating stand storage time in the TB in months for 2013-2020. 5) Calculating uncertainty of the monthly outcome for the observed period. 6) estimate seasonal indexes per seasonality classification, spatiality and harvested and planned volume source. 7) Calculating the difference between the mean seasonal indexes of planned and harvested volumes. 8) Estimating a suitable TB size based on variation in seasonal indexes and outcome, as well as the asymmetry estimated in previous steps. 9) training models per spatiality and source to predict the seasonal index variation.

2.3. Data

The basis for the analysis in this study is harvest data exported from Holmen Skog's system VSOP from 2013 to 2020. VSOP contains most of the relevant information relating to the harvests and the processes connected to them, such as planning and the transport management following it (fig 2). The data, on stand level, that is used for analysis in this study is: Stand id, or as it is used in this study the ith stand, industry measured volume (m^3fub), planned volume by their quality (species and dimension) (m^3fub), the date the stand was created in the system year-month-day (y-m-d), the date that the stand was finished planning for harvest year-month-day (y-m-d), the date harvesting started year-month-day (y-m-d), the date the harvesting was finished year-month-day (y-m-d), the date administrative tasks associated with the stands harvest were completed year-month-day (y-m-d), type of harvest (thinning, final felling, other), stand coordinates in SWEREF 99 (x & y), classification of road accessibility (where 1 is the high and 4 is low {1-4}), classification of terrain accessibility (where 1 is the high and 5 is the low{1-5}).

In this study, only two dates per stand were needed, one date for the creation of the stand in the TB and one date for when harvest occurred to represent when the stand left the TB. Many of the harvested stands never were marked with "planning finished" in the system, so the date for when the stand was created in the system was used instead. In the case of when the stand left the TB, the earliest date associated with harvesting was used. If the start date was missing, the harvest end date was used, and if that also was missing, the harvest complete date was used.

2.3.1. Administrative area classifications

The administrative boundaries naturally change over time; for example, a stand in 2013 and the stand next to it in 2016 might belong to different production units making consistent analysis of a geographic area impossible based on the documented unit or district. Stands were classified according to current administrative boundaries using the geographic information system (GIS) software ArcMap to get around the problem. If stands were located inside the administrative boundary or outside all of the boundaries, then to the closest one. This process gave stands geographically standardised classification throughout the analysis (fig 5): standardised regions {North, Middle, South}, standardised districts {Northern North, Southern North, Middle, South}, standardised production units $\{\mathrm{unit}_1, \ldots, \mathrm{unit}_n\}$.

The same methodology was used for weather data from the Swedish Meteorological and Hydrological Institute (SMHI). SMHI provides open weather data from their weather stations, including their coordinates. The weather data used for analysis in this study was the monthly precipitation and temperature. The weather in the administrative boundaries was analysed by classifying the weather stations using the administrative boundaries set for the stands (fig. 5). Only the weather stations that are inside an administrative boundary were classified. The new classifications and weather information for each stand were then connected to the original data using the stand id of each stand.

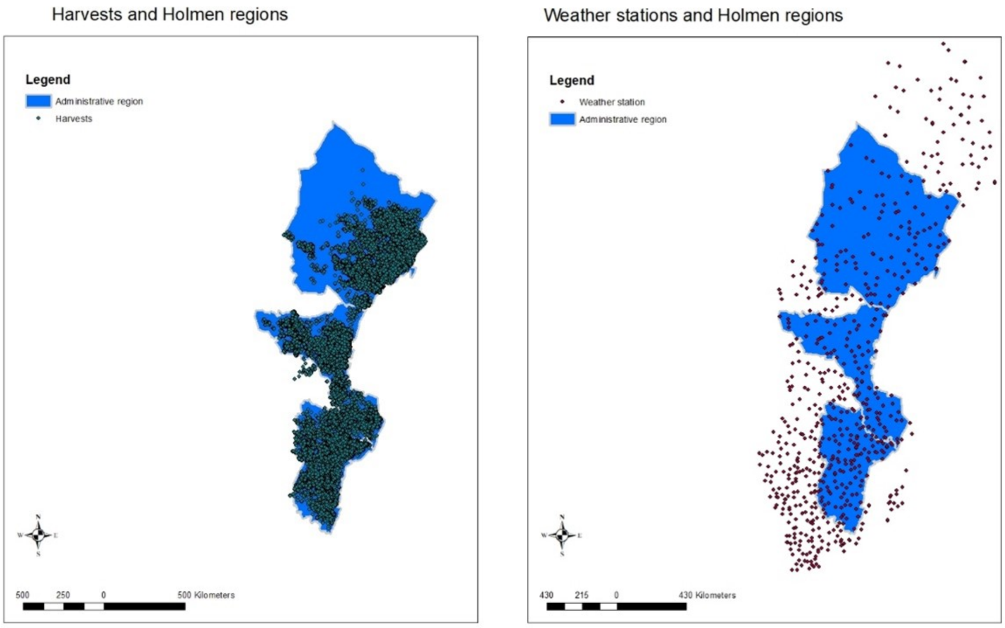


Figure 5. The map on the left shows the harvests analysed in this study with green dots, and the right map shows the weather stations whose monthly data was used to extrapolate weather information in past periods. The blue polygons are the regions of Holmen, North, Middle and South; these regions were used both to classify harvests and weather stations. The classification was also done in smaller district polygons and even smaller than that production unit polygons.

2.3.2. Seasonality classification

Each stand has its road and terrain conditions classified prior to harvesting. Spring is often seen as the most critical period in wood supply management; as the frost thaws in spring, roads and terrain easily break down. In autumn, the drying effect of plant transpiration is lacking, and often, periods of precipitation occur, which also requires more durable road conditions. Summer is a period with conditions that are neither especially high nor low trafficability. Dry summer stands can be harvested in summer if the terrain and road conditions are better than usual. Winter stands need cold conditions, i.e., frozen ground.

The road and terrain classifications were merged into a single variable named seasonality classification, where the highest value dictates the classification. The classification values in the original classifications are on a scale of 1 is high and can be utilised in any weather condition, and 5 is low, only to be utilised in frozen or dry conditions. The new classification was split into five groups, Spring, Autumn, Summer, Dry summer, and Winter (Table 1).

Table 1. The table shows how seasonality classification was determined based on the stands terrain and road classification. Columns 1-5 are the stands in field assessed terrain availability, where 1 is year-round (Spring), 2 not in thaw (Autumn), 3 summer, 4 longer dry periods (Dry summer), and 5 frozen ground (winter). Rows 1-4 are the stands in field assessed road availability where 1 is all year round, 2 not in thaw, 3 not in thaw or rain period, 4 when dry or frozen

Road \ Terrain	1	2	3	4	5
1	Spring	Autumn	Summer	Dry summer	Winter
2	Autumn	Autumn	Summer	Dry summer	Winter
3	Summer	Summer	Summer	Dry summer	Winter
4	Dry summer	Dry summer	Dry summer	Dry summer	Winter

2.4. Historical tract banks

The historical TBs were calculated for each combination of the administrative unit, source, year, and month in the observed period (TB_{ryms}). It was done by grouping the original data of the host company in the mentioned constellations. Planned volume in a certain region, year and month were defined as the total estimated volume (m^3fub) in the stands of that region, which had been planned for harvest and harvest had yet not started. For example, from the day the stand began to be planned for harvest, the total volume of that stand would be counted towards the volume of that month's TB. The stand would then be a part of the TB in the following months until the month a date was first documented for harvest.

$$TB_{ryms} = \sum_{i} P_{rymsci} \, [m^3 fub] \tag{3}$$

P_{rymsci} is the estimated volume (m^3fub) of the ith stand where rymsci refer to a setting of conditions: where r is the administrative region of observation in Holmen Skog, y represents a year from 2013 to 2020, m is the month of the year, s is the source of harvest (contracted or on company land), c is the seasonality classification and i the ith stand. The TB was estimated for each observed unit, year, month, source, and seasonality classification. A TB was transferred into a subset from the original data, as shown in figure 6.

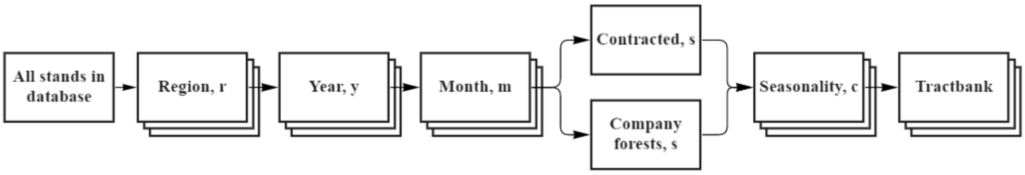


Figure 6. The flowchart visualises how the historical TBs were grouped in the calculations. Firstly grouping stands by region, then grouping the region into years, then grouping the years into months, then grouping the months by source, then grouping the sources into seasonality and then calculating that groups TB. The letters associated with each box refer to the letters $rymsci$ *that are used in defining the subsets in the equations of this study.*

To make the regions comparable with each other, the coverage time (CT_{rsym}), which shows how many months of harvest or coverage time were in that TB_{rsym}, were calculated as

$$\mathrm{CT_{rsym}} = \frac{\mathrm{TB_{rsym}}}{\overline{\mathrm{H}}_{\mathrm{ryms}}} \qquad \text{[Months]} \tag{4}$$

$\overline{H}_{ryms}$ is the mean volume of all the monthly harvests in a region and year (m^3fub/month) and were calculated as

$$\overline{\mathrm{H}}_{\mathrm{rsym}} = \frac{\sum_{\mathrm{i}} \mathrm{H_{rsymi}}}{12}. \tag{5}$$

Where H_{rsymi} is a stand that left the TB in a certain month, and then had its volume measured at the industry.

2.5. Storage time in the tract bank

The storage time or the time the ith stand *(T_i)* spent in the TB, is in this study defined as the number of months between the date it was created in the system and the first registered date of harvest or if not harvested the date the data was acquired (2021-10-11). The weighted mean (4) and standard deviation (5) of the ith stand's storage time was calculated with the weight of the ith stand's estimated wood volume ($\mathrm{P_{rsi}}$) (m^3fub).

$$\mu\mathrm{T_{rsi}} = \frac{\sum_{\mathrm{i}} \mathrm{T_{rsi}} \cdot \mathrm{P_{rsi}}}{\sum_{\mathrm{i}} \mathrm{P_{rsi}}} \qquad \text{[Months]} \tag{6}$$

$$\sigma_{T_{rs}} = \sqrt{\frac{\sum_i P_{rsi} \cdot (T_{rsi} - \overline{T}_{rsi})^2}{\frac{(n-1)}{n}\sum_i^N P_{rsi}}} \qquad \text{[Months]} \tag{7}$$

T is the storage time (months) a stand has spent in the TB, n is the number of nonzero P_{rsi} (weights), and N is the number of observations.

The share (%) of stands that spent more than 60 months (5 years), 36 months (3 years) or less than 8 months in the TB were calculated as the sum of the volume of the stands in a region and source divided by the total volume of that region and source.

$$\%_{\text{longerthan3years}} = \frac{\sum_i I_{36,i} \cdot P_{r,s,i}}{\sum_i P_{rsi}} \cdot 100 \tag{8}$$

$$I_{36} = \begin{Bmatrix} 1, if\ T > 36 \\ 0, otherwise \end{Bmatrix}$$

$$\%_{\text{longerthan5years}} = \frac{\sum_i I_{60,i} \cdot P_{r,s,i}}{\sum_i P_{rsi}} \cdot 100 \tag{9}$$

$$I_{60} = \begin{Bmatrix} 1, if\ T > 60 \\ 0, otherwise \end{Bmatrix}$$

$$\%_{\text{lessthan8months}} = \frac{\sum_i I_{8,i} \cdot P_{r,s,i}}{\sum_i P_{rsi}} \cdot 100 \tag{10}$$

$$I_{8} = \begin{Bmatrix} 1, if\ T < 8 \\ 0, otherwise \end{Bmatrix}$$

2.6. Estimating monthly outcome deviation

There is consistently some degree of error when estimating the volume of a stand in the planning process. The uncertainty in how much the actual outcome volume is compared to the planned volume in the TB is also quite substantial (fig. 7). Many stands are consistent between planned and industry-measured volume (see fig. 7 trendline), but many stands vary noticeably.

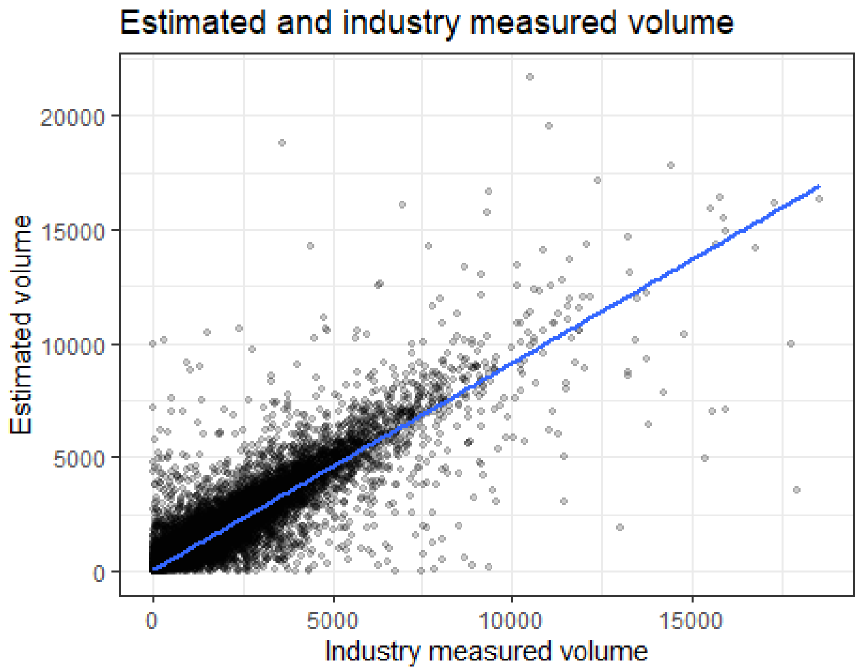


Figure 7. The scatterplot shows on the y-axis the estimated volume (m^3fub) for a specific stand, and the same stand is shown on the x-axis with the volume measured in the industry (m^3fub). Many stands are consistent between planned, and industry measured volume, which can be seen in the trendline. However, there is also many stands that vary greatly.

Not all errors result from wrong estimates, as some is due to administrative procedures. Sometimes two stands are handled as one, resulting in some stands having zero documented harvested volume and others having more. This type of error was tackled here by summarising the entire difference (m^3fub) by source, month, year, and region ($\mathrm{D_{rsym}}$) between estimated and measured volumes and then using that difference for the analysis. The average coverage time difference ($\mathrm{\mu CTD_{rsm_{rsm}}}$) is in each month, grouped by source and region. $\mathrm{\mu CTD_{rsm}}$ is defined as the mean CTD in the observed years from 2013 to 2020. Where Y is the number of years. The standard deviation of the coverage time difference $\left(\sigma_{\mathrm{CTD_{rsm}}}\right)$ is calculated similarly in each month, source and region.

$$\mathrm{\mu CTD_{rsm}} = \frac{\sum_{y} \mathrm{CTD_{rsym}}}{\mathrm{Y}} \tag{11}$$

$$\sigma_{\mathrm{CTD_{rsm}}} = \sqrt{\frac{\sum\left(\mathrm{CTD_{ryms}} - \mathrm{\mu CTD_{rsm}}\right)^2}{\mathrm{Y} - 1}} \tag{12}$$

The coverage time difference between a region year and source $(\mathrm{CTD_{rsym}})$ is defined as the unitless ratio between $\mathrm{D_{ryms}}$ and the month's harvest in that region $(\mathrm{H_{rym}})$.

$$\mathrm{CTD_{rsym}} = \frac{\mathrm{D_{rsym}}}{\mathrm{H_{rym}}} \tag{13}$$

$$\mathrm{D_{rsym}} = \sum_{\mathrm{i}} \mathrm{P_{rsymi}} - \mathrm{H_{rsymi}} \qquad [\mathrm{m^3fub}] \tag{14}$$

$\mathrm{P_{rsymi}}$ is the planned volume (m^3fub) of the ith stand, $\mathrm{H_{rsymi}}$ the harvested volume (m^3fub) of the i stand measured at the industry. Note that $\mathrm{H_{rym}}$ (m^3fub) and $\mathrm{H_{rsymi}}$ are different, as $\mathrm{H_{rym}}$ is the harvest in a certain year, month and region while $\mathrm{H_{rsymi}}$ is harvest on stand level.

2.7. Defining seasonal class utilisation and variation

Different seasonality classifications are utilised differently throughout the year, and seasonal indexes for the harvest and planning levels were calculated to explore this utilisation. The index is unitless and is calculated as the ratio between that year's average monthly harvest and each month's harvest (12). Indexes >1 indicate greater than average harvest, indexes <1 indicate less than average, and an Index =1 indicates equal to that year's average.

$$\mathrm{Si_{H_{rscym}}} = \frac{\mathrm{H_{rscym}}}{\left(\frac{1}{12}\sum_{\mathrm{y}} \mathrm{H_{rscym}}\right)} \tag{15}$$

Similarly, the seasonal index for how planned volume was added to the TB was calculated (13).

$$\mathrm{Si_{P_{rscym}}} = \frac{\mathrm{P_{rscym}}}{\left(\frac{1}{12}\sum_{\mathrm{y}} \mathrm{H_{rscym}}\right)} \tag{16}$$

$\mathrm{Si_{H_{rscym}}}$ is a seasonal index for the harvest of a seasonality classification in a particular source and region in a specific year in a unitless index and similarly $\mathrm{Si_{P_{rscym}}}$ is a seasonal index for the planned volume of a seasonality classification in a certain source and region in a specific year in a unitless index.

A mean seasonal index and its standard deviation were calculated from the seasonal indexes for each group by dividing the sum of seasonal indexes of a certain month in all the observed years and dividing it by the number of years observed.

$$\mu_{\mathrm{Si}_{\mathrm{H}_{\mathrm{rscm}}}} = \frac{\sum_{\mathrm{y}} \mathrm{Si}_{\mathrm{H}_{\mathrm{rscym}}}}{\mathrm{Y}} \tag{17}$$

$$\sigma_{\mathrm{Si}_{\mathrm{H}_{\mathrm{rscm}}}} = \sqrt{\frac{\sum \left(\mathrm{Si}_{\mathrm{H}_{\mathrm{rscym}}} - \overline{\mathrm{Si}}_{\mathrm{H}_{\mathrm{rscm}}}\right)^2}{\mathrm{Y}-1}} \tag{18}$$

$$\mu_{\mathrm{Si}_{\mathrm{P}_{\mathrm{rscm}}}} = \frac{\sum_{\mathrm{y}} \mathrm{Si}_{\mathrm{P}_{\mathrm{rscym}}}}{\mathrm{Y}} \tag{19}$$

$$\sigma_{\mathrm{Si}_{\mathrm{P}_{\mathrm{rscm}}}} = \sqrt{\frac{\sum \left(\mathrm{Si}_{\mathrm{P}_{\mathrm{rscym}}} - \overline{\mathrm{Si}}_{\mathrm{P}_{\mathrm{rscm}}}\right)^2}{\mathrm{Y}-1}} \tag{20}$$

$\mu_{\mathrm{Si}_{\mathrm{H}_{\mathrm{rscm}}}}$ is the mean seasonal index of harvest in a specific region, month, source, and seasonality classification, $\sigma_{\mathrm{Si}_{\mathrm{H}_{\mathrm{rscm}}}}$ is its standard deviation. $\mu_{\mathrm{Si}_{\mathrm{P}_{\mathrm{rscm}}}}$ is the mean seasonal index of planned volume in a certain region, month, source, and seasonality classification, $\sigma_{\mathrm{Si}_{\mathrm{P}_{\mathrm{rscm}}}}$ is its standard deviation.

2.8. Estimating a suitable tract bank

The estimated suitable TB (CT_{rs}) was calculated as the least needed coverage time to cover for the expected seasonal variation and the uncertainty in that expected variation. The expected seasonal variation is defined as the varying monthly difference between rates of planning and harvest throughout the year. The uncertainty is defined as the observed variation in that difference between years in the limiting period the year.

$$\mathrm{CT}_{rs} = \mathrm{CT}_{\mathrm{r,s}_{\mathrm{cyclestock}}} + \mathrm{SL}_{rs} \tag{21}$$

The cycle stock ($\mathrm{CT}_{\mathrm{r,s}_{\mathrm{cyclestock}}}$) covers the expected change in CT_{rs} and the security stock (SL_{rs}) covers the uncertainty defined by the limiting seasonality classification. Moreover, a total suitable coverage time per region (CT_r) was calculated for more easier comparison with the past tract, $\mathrm{CT}_r = \mu\mathrm{CTRatio}_{r,estate} \cdot \mathrm{CT}_{\mathrm{r,estate}} + \mu\mathrm{CTRatio}_{r,\mathrm{contracted}} \cdot \mathrm{CT}_{\mathrm{r,contracted}}$.

2.8.1. Calculating the cycle stock

The cycle stock $\left(CT_{r,s_{cyclestock}}\right)$ was expressed as the minimal offset needed to keep the expected CT in all seasonality classifications above 0 throughout the year. With the assumption that the fluctuations in the TB, is the difference between the mean rate of planning and harvesting of the different seasonality classifications in the observed years for each month. The solutions for how small $CT_{r,s_{cyclestock}}$ can be at the start of the year ($CT_{rs_{offset_{min}}}$) and not reach 0 in any seasonality classification or month were found using linear programming to find the theoretical minimum (fig. 8).

$$\begin{aligned} CT_{rs_{cyclestock}}(m) &= CT_{rs_{offset_{min}}} + CTD_{rs} + \triangle CT_{rs}(m) = \\ &= \min\left\{CT_{r,s_{offset}}\right\} + CTD_{rsym} + \sum_{c} \triangle CT_{r,s,c,m} \end{aligned} \tag{22}$$

Subject to

$$CT_{r,s,c} + \triangle CT_{r,s,c,m} \geq 0 \tag{23}$$

where

$$CT_{r,s,c} = CT_{r,s_{offset}} \cdot \mu CTRatio_{r,s,c} \tag{24}$$

and

$$\triangle CT_{r,s,c,m} = \triangle ECT_{r,s,c,m} + \triangle CT_{r,s,c,m-1} \tag{25}$$

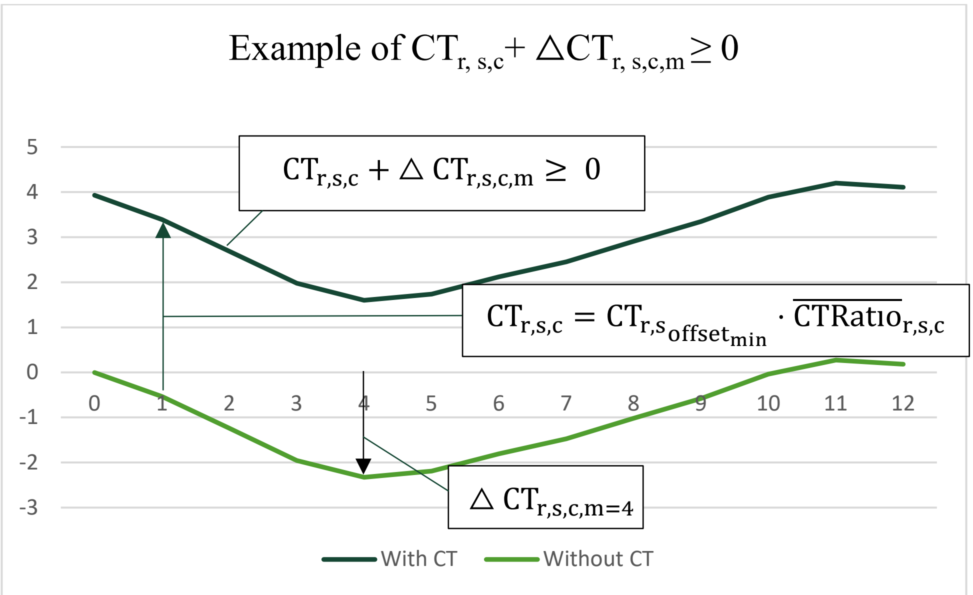


Figure 8. The graph shows a yearly fluctuation of a Stand bank for a region, source and seasonality classification and shows how the offset increases the stand bank such that it is never below 0. Notice that it is higher than 0 in all months and that is because another seasonality classification is limiting.

As the composition the stand bank is assumed as constant an increase in one classification forces an increase in all.

The expected monthly change of coverage time of a seasonality classification by source ($\triangle\ ECT_{rsmc}$). Was defined as the difference between the expected planned seasonality classification in that source ($E_{P_{msc}}$) and the expected harvest of that month, seasonality, and source ($E_{H_{msc}}$).

$$\triangle\ ECT_{rscm} = E_{p_{rscm}} - E_{H_{rscm}} \tag{26}$$

The expected harvest ($E_{H_{msc}}$) is defined as the mean seasonal index ($\mu Si_{H_{rmsc}}$) multiplied by that seasonality classifications ratio of the yearly harvest ($\mu HRatio_{rsc}$) (20).

$$E_{H_{rmsc}} = \mu Si_{H_{rmsc}} \cdot \ \mu HRatio_{rsc} \tag{27}$$

The expected planned volume ($E_{P_{msc}}$) is defined similarly as the mean seasonal index ($\mu Si_{P_{rmsc}}$) multiplied by that seasonality classifications ratio of the yearly planning ($\mu PRatio_{rsc}$) (22).

$$E_{P_{rmsc}} = \mu Si_{P_{rmsc}} \cdot \ \mu PRatio_{rsc} \tag{28}$$

It was done as $\mu HRatio_{rsc}$, $\mu PRatio_{rsc}$ and the mean composition of the TB ($\mu CTRatio_{rsc}$) are different.

$$\mu HRatio_{rsc} = \frac{H_{rsc}}{H_r} \tag{29}$$

$$\mu PRatio_{rsc} = \frac{P_{rsc}}{P_r}$$

$$\mu CTRatio_{rsc} = \frac{\sum_m^y \frac{CT_{rymsc}}{CT_{ryms}}}{Y \cdot M} \tag{30}$$

2.8.2. Calculating the security stock

The TB should also be large enough to handle a greater-than-expected demand. In other words, the TB should also have a safety stock (SL) (Lumsden et al. 2019). Only the uncertainty in utilising seasonality classifications and the variance of CTD is considered when estimating the security stock, as variation in the lead time is seen as constant. Variance in the difference between harvesting and planning is considered since both the variation in the creation and harvest of stands in the TB make up its associated uncertainty. It is further motivated by the fact that both have a degree of control over what is in and is used in the TB. The forest planning must fulfil the harvest estimate and the wood supply must meet the delivery plan and can demand stands not in the TB if needed (fig. 1.).

$$\mathrm{SL}_{rs} = \mathrm{Z} \cdot \sqrt{\sigma\mathrm{CTD}_{rs}^2 + \sigma\mathrm{Si}_{\mathrm{diff}_{\mathrm{r,s,c=limiting}}}^2} \quad (31)$$

σ_{Reg}^2 is the variation between what is estimated to be in the stock and the actual stock: it is set to be equal to $\sigma\mathrm{CTD}_{rs}^2$. The general model for calculating the SL (2) is adapted by setting $(\sigma_{\mathrm{LT}} \cdot \mathrm{D})^2$ as 0 as the standard deviation $\sigma_{\mathrm{LT}} = 0$. Z is assumed to be 1.28 in this study, which is the corresponding factor of a service level of $\pm 10\%$ or 90%. Delivery precision at 90 % is the norm in many parts of the Swedish wood supply chain, for example, mill-specific deliveries. The expression, $\mathrm{LT} \cdot \sigma_{\mathrm{D}}^2$, is written as $\sigma\mathrm{Si}_{\mathrm{diff}_{\mathrm{r,s,c=limiting}}}^2$ since the variance of the expected total difference in the period of uncertainty for the limiting seasonality classification is considered (fig. 9). The standard deviation of the total difference in the period ($\sigma\mathrm{Si}_{\mathrm{diff}_{\mathrm{r,s,c,}}}$) is calculated from the observed differences in the years 2013 to 2020.

The period of uncertainty is the period at the end of which the limiting seasonality classification is expected to reach 0, if there is no SL.

$$\mathrm{Si}_{\mathrm{diff}_{\mathrm{r,s,c,y}}} = \sum_m \mathrm{I}_m \cdot \left(Si_{Plan_{\mathrm{r,s,c=limiting,y,m}}} - Si_{Harv_{\mathrm{r,s,c=limiting,y,m}}} \right) \quad (31)$$

I_m are the months of uncertainty. Functionally I_m is a filter of months that define that period of uncertainty. For example, the contract bank of region middle has the limiting seasonality classification winter. The winter stands have an expected negative change in january, february and marsh (1,2,3), as such these are the months of uncertainty, and the security stock is estimated to cover these months (fig 9).

$$\mathrm{I}_m \begin{cases} = 1 & \mathit{For\ januari, february\ and\ march} \\ = 0 & \mathit{Otherwise} \end{cases}.$$

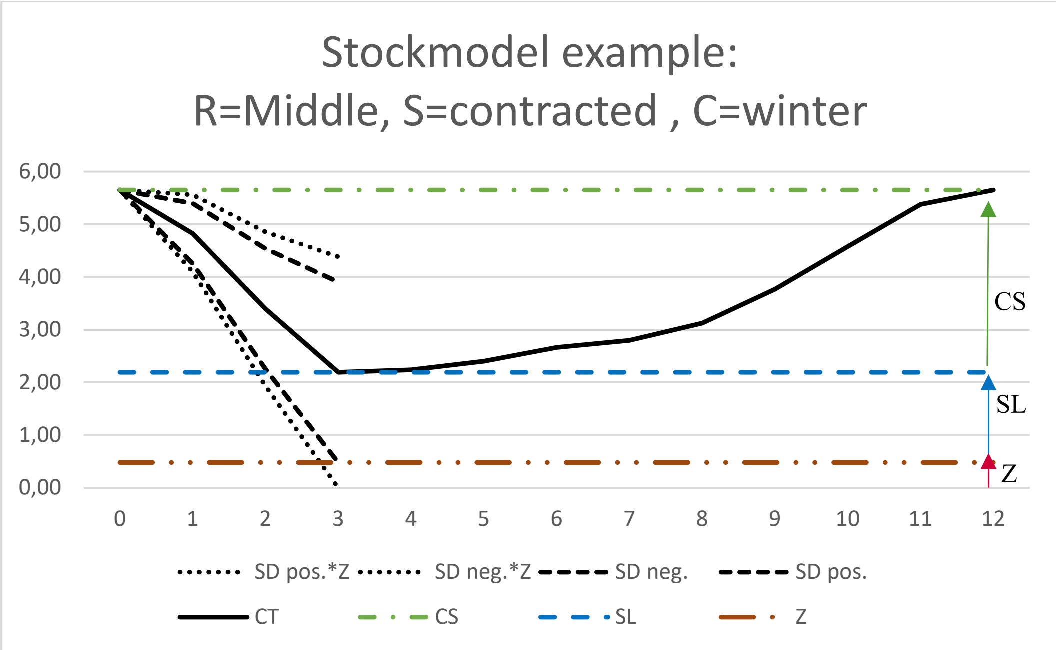


Figure 9. The figure visualizes an estimated stock model and visualizes its individual parts. The spread of uncertainty is visualized as accumulated uncertainty, where visually a third of the SL is added each month. However, it is calculated as one value based on the different differences observed between harvest and planning in the period 2013 to 2020. The continuous line is the coverage time in both the SL and the cycle stock in a given month as calculated. The top line that is green and dashed with one dot, CS, is the cycle stock. The blue dashed line is the SL not adjusted for service level. The bottom line, red dashed with two dots, is the added stock Z to cover for desired service level.

2.9. Modelling and predicting the seasonal index

To model how stands are utilised seasonally a random forest model was made for each group of an administrative region, source and seasonality classification (not by month and year). The model target variable was the seasonal index each month $\mathrm{Si}_{\mathrm{H}_{\mathrm{rymsc}}}$ or the $\mathrm{Si}_{\mathrm{P}_{\mathrm{rymsc}}}$ for the years 2013 to 2020. The explaining variables (fig. 10) included the cover time of the stand bank and contract bank (CT_{rsm}), and how much coverage time of each seasonality classification that were in them (CT_{rscm}). The explaining variables also included weather data (SMHI 2022a) on the administrative region's arithmetic mean and maximum monthly precipitation and temperature. The means and maximums were calculated from all the weather stations measured monthly mean precipitation and temperature (as described in chapter 2.3.1). The information also included data on that month’s spruce and pine round wood prices from the Food and Agricultural Organization of the United Nations (FAO).

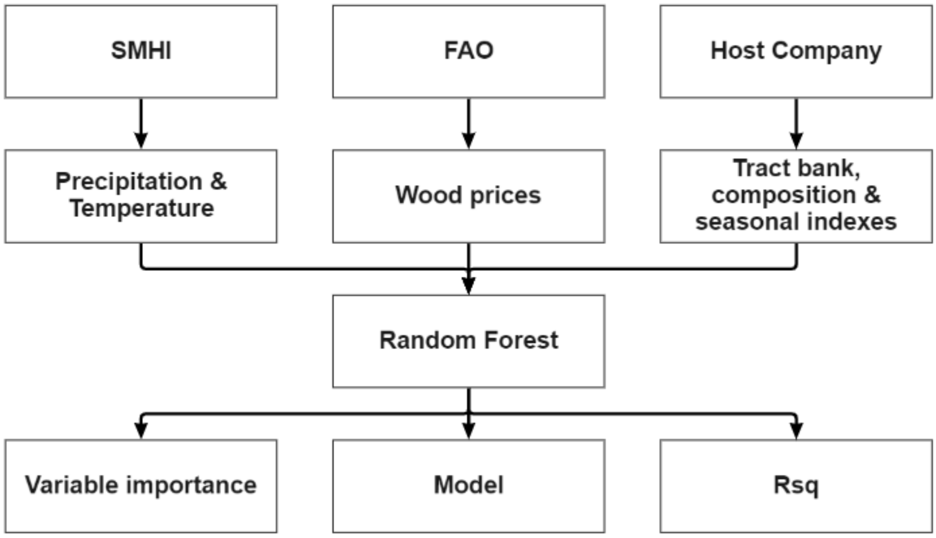


Figure 10. The flowchart visualizes the inputs and outputs of the models. SMHI, the Swedish meteorological and hydrological institute have weather stations throughout the country with the measurements available to the public, from these measurements' monthly levels of precipitation and temperature were estimated. FAO have numbers on wood prices for each month, that is submitted per country. The tact bank size, composition and seasonal indexes were calculated by historical data from the host company.

The randomforest model is an ensemble of decision trees that are generated by considering a random set of features at each node of each tree (fig. 11). The number of variables considered at each node is by default set to the square root of features but can often benefit from finetuning. The tree decides on the best performing feature at each node either by finding the data split with the least residuals or least impurity. The trees of the randomforest then independently come to a solution and vote for the best solution. In the case of regression that means the average of the forest's predictions and for classification actual votes for the possible classifications (James, G. et.al. 2021).

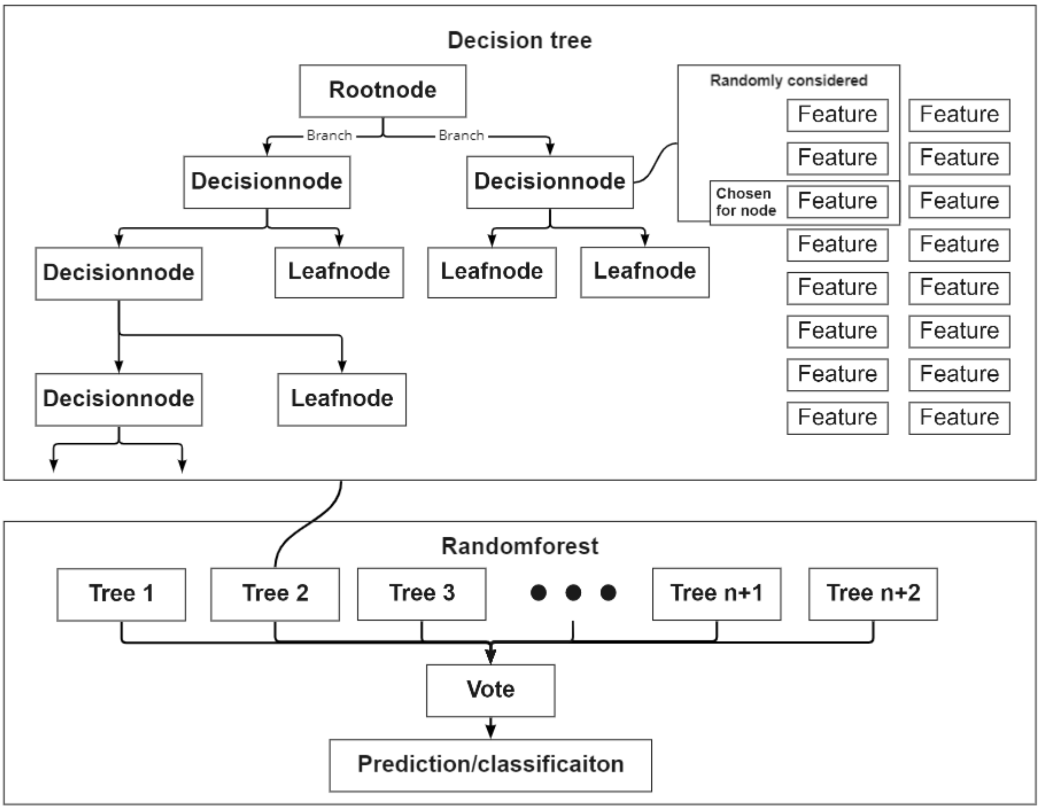


Figure 11. The diagram visualises part of an example of a decision tree in the randomforest, and a random forest. At each tree node, random features are considered and the best chosen for that decision node or root node. After it's been chosen it won't be considered in the other nodes of that tree again. The segments between nodes are referred to as branches, the root node is the first split in the tree and the leafnodes also called terminal nodes are at the end of a branch. The trees in the random forest vote for the best solution.

The features were grouped by administrative region, year, month, source, and seasonality classification (figure 5). Each induvial model was made with the settings: ntrees =1000, proximity = True, other settings as default.

Additional output from these models other than the predicted seasonal indexes was the variable importance for the predictions and a rsq for each model. The variable importance is a measurement of how much a particular variable improved the model in all decision trees created for the model. This gives insight into what improved the model's predictive power in the 1000 trees and how well the models performed. The rsq is a pseudo R^2 (R^2 = 1 - mse / Var(y)), a measurement of how well the model explains the variation of the variable it predicts. A low rsq would mean that the model explains little of the variation, and if it is very low, such as a negative value, it then imposes more variation than it explains. For a more thorough explanation of random forest modelling read James, G. et.al. (2021), L, Breiman (2001) and L, Breiman (2002).

3. Results

3.1. Historical tract banks

Table 2. The table show the mean coverage time by region for both sources and for all observed periods.

Region	North	Middle	South
Tract bank	13.1	9.5	5.8

In the observed years, 2013 to 2020, the host company's TB is smaller than the 24 months coverage time in the stand bank and 8 months coverage in the contract bank that is often referred to in Swedish forestry (Table 2;3). The two regions, Middle and South, which are more heavily dependent on contracted harvests, are closer to 6 months of harvest in their contract bank rather than the 8-month upperbound (Table 2; fig. 12). While region North that is less dependent on contracted harvests and only has around 4 months coverage time in the contract bank.

Table 3. The table shows the mean and standard deviation (SD) of coverage time (months) in the tract bank (TB) in the host company's three observed regions, North, Middle, and South of Sweden. The tract bank is divided by its two sources: Estate and contracted stands. The number of months of harvests in the tract bank in each source and month was calculated by dividing the sum of the wood volume in the tract banks by the mean monthly harvest of its source that particular year and region. The mean value and its SD that is presented here is the mean TB of all observed months in the period 2013-2020 by its region and source.

	Estate TB			Contracted TB		
	North	**Middle**	**South**	**North**	**Middle**	**South**
Mean	15.4	13.3	7.1	3.9	6.4	5.4
SD	3.3	3.5	2.7	1.0	1.2	0.8

The TB in the three regions fluctuate differently throughout the year (fig 12; table 4). The North region varies most with 3.2 months of harvest less in its stand bank in June compared with January and its contract bank having a difference of 1.6 between January and May despite being the smallest TB. The Middle region has a difference of 1 month's harvest between its peak in the stand bank in January and May with a difference of 0.6 in its contract bank between two peaks in January and

July, the lowest point in April. The south region has the least yearly variation in the size of its tract banks, with a peak size in April with 7.5 months of harvest and November with a difference of 0.4 in its stand bank, whilst the contract bank has 5.7 in July and 5.0 in November with a difference of 0.7 (Table 4).

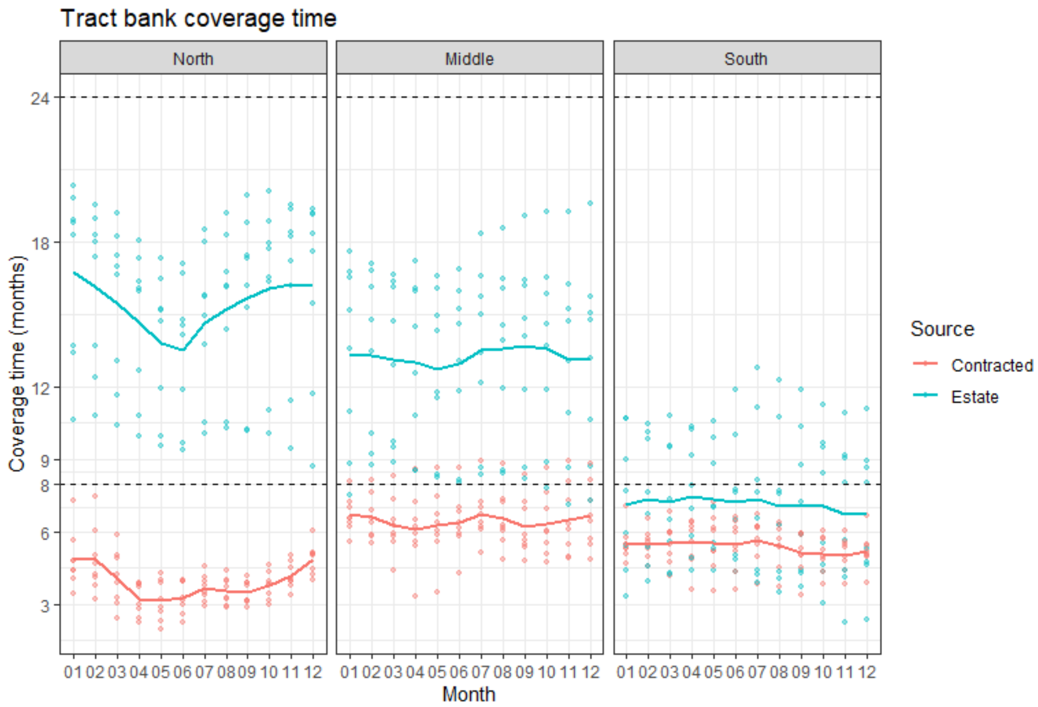


Figure 12. *The figure shows the monthly mean and all monthly values of coverage time (months) in the tract bank (TB) in the host company's three observed regions, North, Middle, and South of Sweden. The tract bank is divided by its two sources: Estate and contracted stands. The number of months of harvests in the tract bank in each source and month was calculated by dividing the sum of the wood volume in the tract banks by the mean monthly harvest of its source that particular year and region. Each dot is a region's TB in a certain month, the continuous coloured lines are the mean TB for that month, source and region. The dashed lines are TB levels where the 24 months is the stand bank level often refered to and the 8 months is a upper limit often refered to in dimensioning the contract bank.*

Table 4. *The table shows the monthly mean values of how many months of harvest there were in the tract bank (TB) in the host company's three observed regions, North, Middle, and South of Sweden in the months of the year. The tract bank is divided by its two sources: Estate and contracted stands. The number of months of harvests in the tract bank in each source and month was calculated by dividing the sum of the wood volume in the tract banks by the mean monthly harvest of its source that particular year and region. Each value is a region's TB in a certain month, source and region.*

	Mean Estate TB			**Mean Contracted TB**		
Month	**North**	**Middle**	**South**	**North**	**Middle**	**South**
Jan	16.7	13.4	7.1	4.8	6.7	5.5
Feb	16.1	13.3	7.3	4.8	6.6	5.5
Mar	15.5	13.1	7.2	4.1	6.3	5.5
Apr	14.7	13.0	7.5	3.2	6.1	5.5
May	13.8	12.7	7.3	3.2	6.3	5.6
Jun	13.5	13.0	7.3	3.3	6.4	5.5
Jul	14.7	13.5	7.4	3.7	6.7	5.7
Aug	15.2	13.6	7.1	3.5	6.5	5.4
Sep	15.7	13.7	7.1	3.5	6.2	5.1
Oct	16.1	13.6	7.1	3.8	6.3	5.1
Nov	16.2	13.2	6.7	4.2	6.5	5.0
Dec	16.2	13.1	6.7	4.9	6.7	5.2

3.2. Storage time in the tract bank

On average, for all regions, sources and stand types, the storage time of volume in the TB was 10.2 months. But for the volume in the stand banks the mean storage time was 13.8 months with a larger than the mean standard deviation of 17.0 months where 8.2 % were harvested in its first month and 61% within the first year. The contract banks have a mean volume storage time of 5.9 months and a larger than mean standard deviation of 8.2 where 18.8 % was harvest in its first month and 87% within the first year. The trend of the standard deviation being larger than the mean storage time can be seen in all regions and sources (table 5).

The share of planned wood volume that spend more than 3 or 5 years in the stand bank are greater the further north its located, and the largest difference can be seen between the South and middle regions. The same trend cannot be seen for contracted stands.

Table 5. The table shows how much time stands spent stored in the TB on average and its standard deviation, in different regions and by source.

Storage time in tract bank						
Region	**Source**	**[months] Mean**	**[months] SD**	**% ≥3 years**	**% ≥5 years**	**% ≤8 months**
North	Contracted	4.1	6.4	0.3	0.1	83.9
North	Estate forest	16.4	18.2	9.5	2.3	41.1
Middle	Contracted	7.0	9.0	1.0	0.1	70.4
Middle	Estate forest	13.0	15.1	8.7	1.0	50.9
South	Contracted	5.4	7.7	0.7	0.1	78.1
South	Estate forest	6.4	10.7	4.0	0.4	70.0

3.3. Monthly outcome deviation

The deviation between estimated and industry measured outcomes are a few per cent when seen as a monthly summation (fig. 13; table 6). However, it does indicate that the average stand tends to be underestimated, as the mean value for all regions is negative. A trend can be seen from north to South that contracted stands are underestimated more with greater variation in the South. The same trend cannot be seen in the stand banks stands, as they all are similarly underestimating stand volume with no clear geographic trend to that variation, but the middle region has the largest variation.

Table 6. The table shows by source and region the mean and the standard deviation of the difference between the estimated wood volume and the wood volume measured at industry (m^3fub) for all harvested stands as a summation per month in all months in the period 2013 to 2020. The differences are grouped by their two sources: Estate and contracted stands and region North, Middle and South of Sweden. The unit is the region and sources monthly harvest.

	Estate TB			**Contracted TB**		
	North	**Middle**	**South**	**North**	**Middle**	**South**
Mean	-0.02	-0.01	-0.01	-0.01	-0.04	-0.07
SD	0.01	0.01	0.00	0.01	0.02	0.02

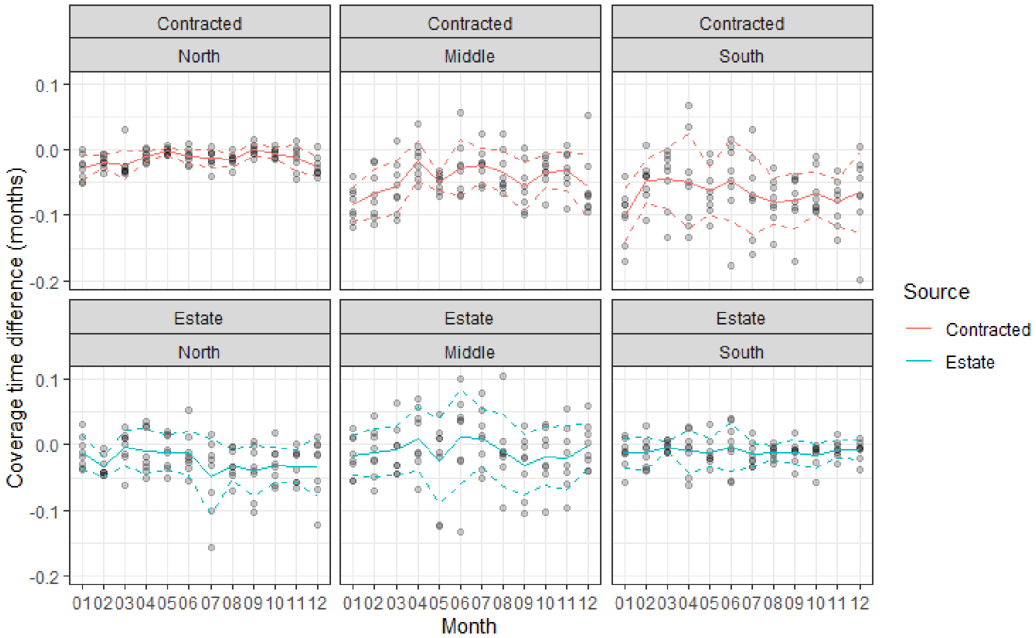


Figure 13. *The figure shows the standard deviation (dotted line) and the mean difference (continuous line) between the estimated wood volume and the wood volume measured at industry (m^3fub) for all harvested stands as a summation per month in all months in the period 2013 to 2020. Each dot is one month's difference divided by the months total harvest which turns it into a relative difference compared with the other regions and sources. The relative differences are grouped by their two sources: Estate and contracted stands as well as by region North, Middle and South of Sweden.*

3.4. Seasonal indexes for the planning and harvest

In exploring the seasonal difference between wood volume harvested and planned throughout the year, the seasonal indexes for planning and harvesting of the seasonality classifications were plotted against each other, fig. 14-17. A seasonal asymmetry between the work processes planning and harvesting was identified in how they create and utilise stands with different seasonality classifications throughout the year. Consequently, different trends could be observed for different sources, regions, and seasonality classifications.

For the seasonal indexes of spring and winter (fig 14), there are distinct peaks in the estate source of when the harvesting occurs. During the winter months the winter harvesting index have a distinct peak to then during spring and summer decrease to an index close to 0. Similarly, the spring harvesting index have a distinct peak in when it occurs and then for the rest of the year staying close to 0. However, the contracted spring stands does not follow the same seasonal pattern. These stands instead are utilised similarly throughout the year.

There is a large asymmetry in the estate source of when a stand of these two seasonality classifications are planned and harvested. This can especially be observed in the north, and middle regions estate, here more distinct peaks in the seasonality can be seen gradually diminishing on a southward gradient. Harvesting is mainly at the beginning of the year, while the planning of the same stands are spread out throughout the year. The southern region has a similar but less distinct seasonal pattern. The asymmetry is more similar between regions for contracted stands than in the estate stands.

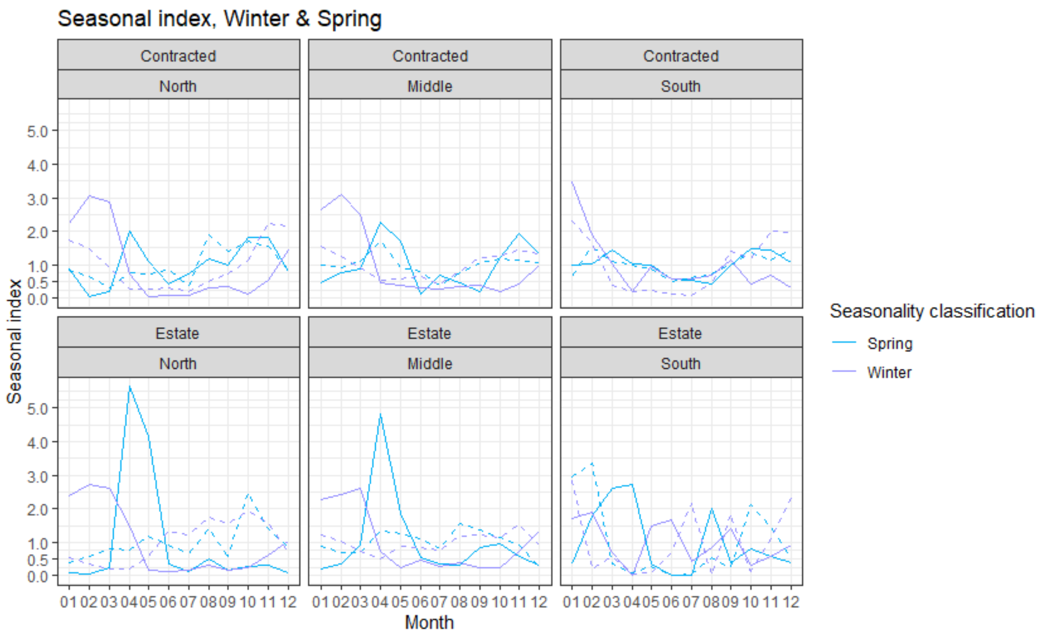


Figure 14. The figure shows the mean seasonal indexes for planning and harvesting of winter and spring stands, as they were classified in this study (Table 1). The continuous lines are harvesting indexes, and the dotted lines are planning indexes. The sum of the index values in each line equals 12, similarly, the mean index in all months (the entire year) is 1. If the harvest index is 4 in one month, then the harvesting in that month equals 4 average monthly harvests for that index. The mean seasonal indexes seen in the figure are calculated by dividing the monthly outcome with the mean monthly outcome, in all years 2013 to 2020, and then calculating the monthly mean of the indexes.

For stands classified as dry summer (fig 15) there seems to be a large variation between how the different regions and sources utilise the stands. In the south region, they seem to be utilised evenly throughout the year with a slight peak at the end of summer. Also, in the north and middle estate harvests a similar pattern of being utilised evenly throughout the years can be seen. For the contracted middle region on the other hand a tendency to be used during winter and less so during spring and fall can be seen. However, the contracted stands in the north seem to be used similarly to winter stands, but with a slightly larger tendency to be used at the end of summer. The main asymmetry between planning and harvesting is mainly in the north estate at the beginning of the year. There the planning index is low similarly to the spring and winter indexes.

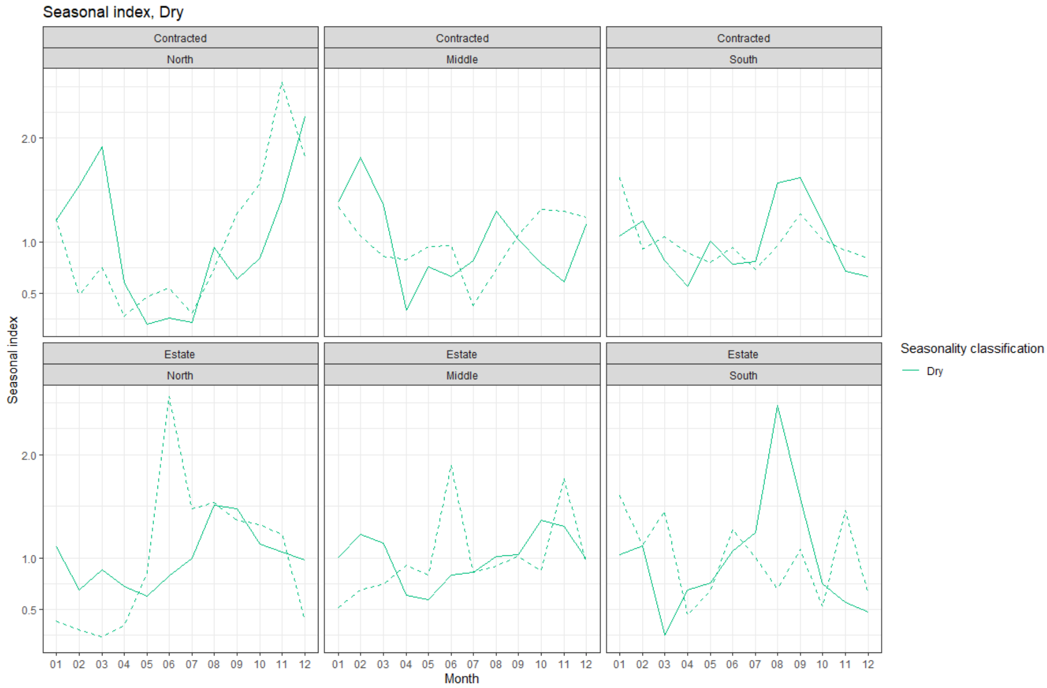


Figure 15. *The figure shows the mean seasonal indexes for planning and harvesting dry summer stands, as they were classified in this study (Table 1). The continuous lines are harvesting indexes, and the dotted lines are planning indexes. The sum of the index values in each line equals 12, similarly, the mean index in all months (the entire year) is 1. For example, if the harvest index is 4 in one month, then the harvesting in that month equals 4 average monthly harvests for that index. The mean seasonal indexes seen in the figure are calculated by dividing the monthly outcome with the mean monthly outcome, in all years 2013 to 2020, and then calculating the monthly mean of the indexes.*

Looking at the summer indexes (fig. 16), the harvesting and planning of the stands seem to occur with similar seasonality. The stands in the south region seem to be utilised evenly without seasonality throughout the year. In the middle region there seems to be a tendency to use them during summer slightly skewed towards the end of summer. For the north estate stands, a clear seasonality in utilisation during summer can be seen and for the contracted stands in the north a pattern of less use in the first half of the year and more in the second half can be seen.

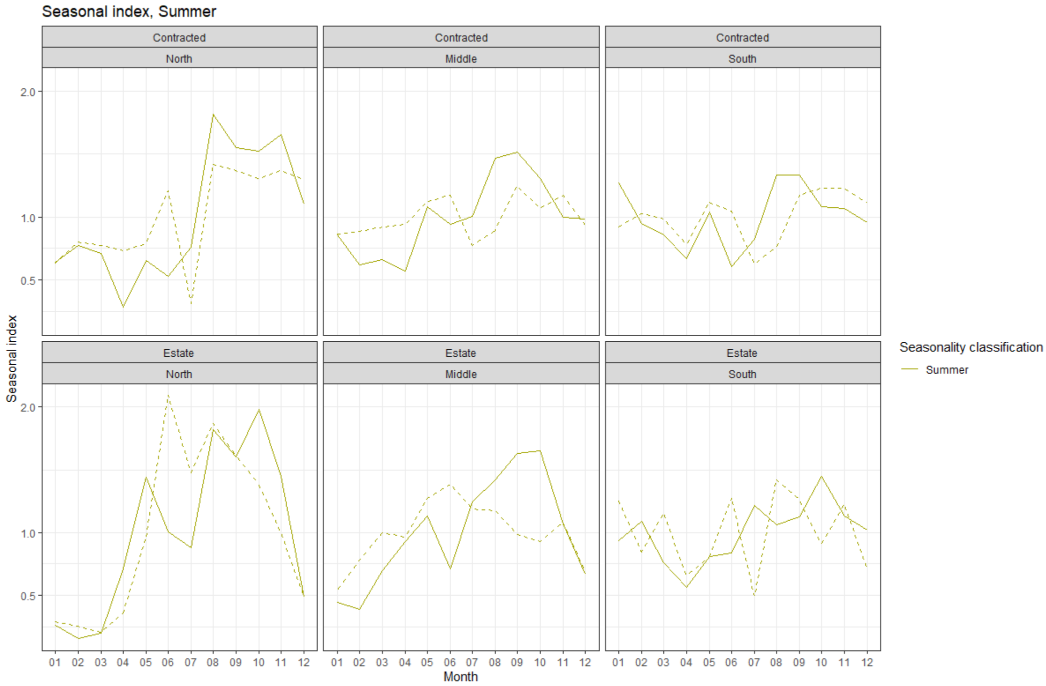


Figure 16. The figure shows the mean seasonal indexes for planning and harvesting of summer stands, as they were classified in this study (Table 1). The continuous lines are harvesting indexes, and the dotted lines are planning indexes. The sum of the index values in each line equals 12, similarly the mean index in all months (the entire year) is 1. As such if the harvest index is 4 in one month, then the harvesting in that month equal 4 average monthly harvests for that index. The mean seasonal indexes seen in the figure are calculated by dividing the monthly outcome with the mean monthly outcome, in all years 2013 to 2020, and then calculating the monthly mean of the indexes.

The fall indexes (fig. 17) are utilised differently for different regions and sources. In the south an even utilisation can be seen with a tendency towards use in winter for the estate source. The contracted stands in the north and middle regions are utilised with an expected aggregation towards the fall months. On the other hand, the estate sourced stands in the north and middle region are utilised as both spring and fall stands. The harvesting and planning indexes follow each other to a large extent, except in the north where the estate begin to utilise fall harvests before their planning season starts and stop to utilise more abruptly towards the end of the year compared with the planning.

Figure 17. The figure shows the mean seasonal indexes for planning and harvesting of fall stands, as they were classified in this study (Table 1). The continuous lines are harvesting indexes, and the dotted lines are planning indexes. The sum of the index values in each line equals 12, similarly, the mean index in all months (the entire year) is 1. If the harvest index is 4 in one month, then the harvesting in that month equals 4 average monthly harvests for that index. The mean seasonal indexes seen in the figure are calculated by dividing the monthly outcome with the mean monthly outcome, in all years 2013 to 2020, and then calculating the monthly mean of the indexes.

3.5. Suitable tract bank, cycle stock and security stock

Table 7. The table show the suitable coverage time estimated per region for both sources and for the observed periods.

Region	**North**	**Middle**	**South**
Tract bank	9,5	5,9	7,1

The north and middle regions are smaller than mean past tract banks while the south region is larger than the past tract banks of the region (table 2;7). Most of the cycle stocks show a similar level of around 3,3 the estate in the North region stands out with 8,5 months of harvest too cover the expected development of the TB (table 8). Most of the security stocks also show similar levels of around 2-3 months of harvest to cover the uncertainty of the limiting period and seasonality classification. The estate in region south stands out here with 5,2 months of harvest to cover for the uncertainty. Both of the cases that stand out are limited by the spring seasonality classification and have the estate as source.

Table 8. The table shows suitable tract bank sizes (CT_{rs}) in months of harvesting in stock as calculated in this study. As well as its two main components the cycle stock ($CT_{r,s_{cyclestock}}$) and the safety stock (SL_{rs}). It also shows which is the limiting seasonality classification and the months of uncertainty for the limiting seasonality classification.

Source	Region	$CT_{rs_{cyclestock}}$	SL_{rs}	CT_{rs}	I_m	Limiting C
Estate	North	8,5	1,9	10,5	4,5	Spring
Contracted	North	3,6	2,2	5,8	1,2,3,4	Winter
Estate	Middle	3,3	2,9	6,2	1,2,3,4	Winter
Contracted	Middle	3,3	2,2	5,7	1,2,3	Winter
Estate	South	3,2	5,2	8,4	4,5	Spring
Contracted	South	3,2	3,5	6,7	1,2,3	Winter

3.6. Predicting regional seasonality utilisation

The seasonal indexes for the different seasonality classifications can too varying degrees be predicted by randomforest regression models (table 8; 9). For harvesting the North region has the greatest degree of predictability, closely followed by the Middle region. There the rsq is around 0.5 in many cases indicating that the models can predict a large part of the variation. The rsq for predicting the winter harvesting indexes in north and middle are considerable as well as the summer harvesting in the north estate.

However, the models in this study fail to predict region South's seasonal indexes (seasonal utilisation) by any source or seasonality classification. Also, the seasonal indexes of planning could not be predicted either except for some cases in the north regions planning (table 9).

Table 7. The table shows the rsq output from a random forest model. The models aim to predict the seasonal index for harvest of a seasonality classification in a certain month and there is one model per seasonality classification, region, and source.

Harvest rsq		Seasonality classification				
Region	Source	Winter	Spring	Summer	Dry summer	Fall
North	Estate forest	0.83	0.41	0.67	0.08	0.53
North	Contracted	0.79	-0.01	0.34	0.49	0.16
Middle	Estate forest	0.62	0.10	0.40	-0.06	0.21
Middle	Contracted	0.79	0.10	0.36	0.28	0.24
South	Estate forest	-0.21	-0.22	0.01	-0.07	0.08
South	Contracted	0.18	0.13	0.09	0.12	0.21

Table 8. The table shows the rsq output from a randomforest model. The models aim to predict the seasonal index for planning of a seasonality classification in a certain month and there is one model per seasonality classification, region and source.

Planning rsq		Seasonality classification				
Region	Source	Winter	Spring	Summer	Dry summer	Fall
North	Estate forest	0.55	-0.13	0.61	0.51	0.49
North	Contracted	0.43	-0.16	0.02	0.27	-0.05
Middle	Estate forest	-0.21	-0.07	0.03	-0.22	-0.05
Middle	Contracted	0.26	-0.21	-0.03	0.14	0.10
South	Estate forest	0.01	-0.13	-0.11	-0.14	-0.12
South	Contracted	0.11	-0.12	-0.01	-0.08	0.18

The data that has the highest variable importance in making the predictions were data on monthly temperature and the month of the year (fig 18). The lack of predictability in region south indicates the forestry is more seasonally constrained in the north than in the South.

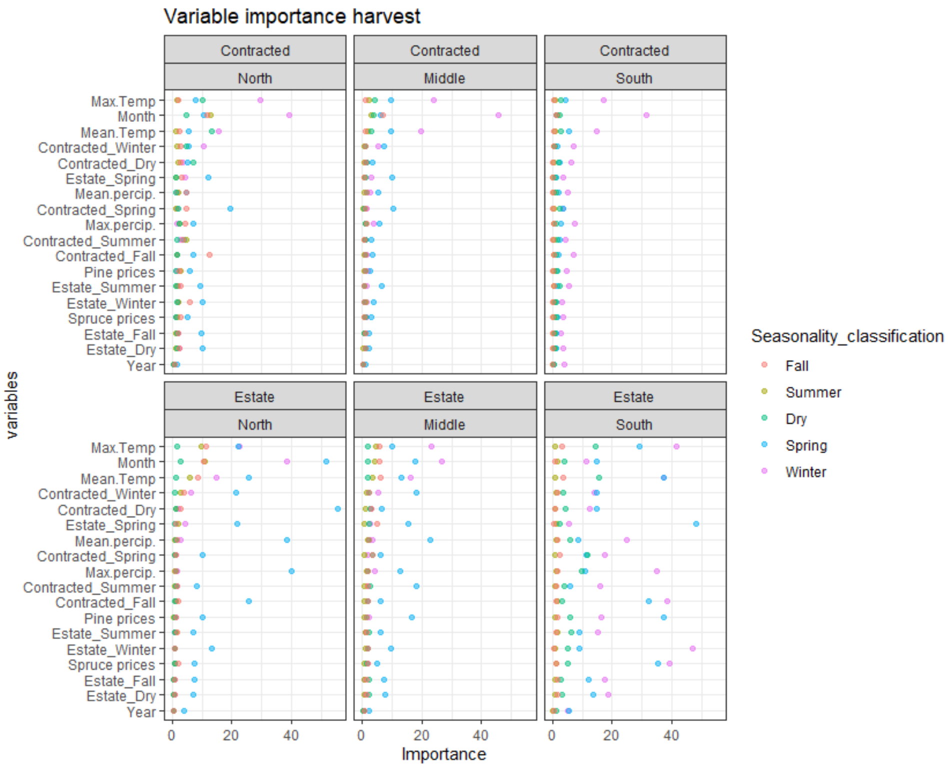


Figure 18. The figure shows variable importance in predicting certain seasonality classifications harvest seasonal indexes. See table 12 for rsq for each model, as many of them are very low.

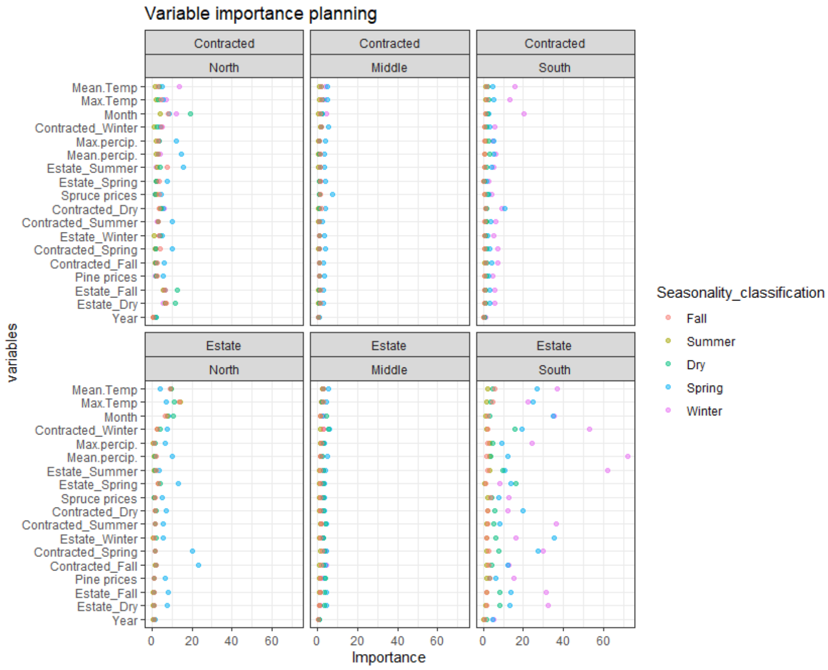


Figure 19. The figure shows variable importance in predicting a certain seasonality classification planning seasonal indexes.

4. Discussion

This study indicates that there are spatially and source dependent uncertainties that affect the TB. Spatially, the most apparent tendencies can be seen on a regional scale from North to South and vice versa. Northwards a more significant part of stands spend a longer time in the TB, a greater asymmetry between the planning and harvest indexes can be seen and models can explain more of the utilisation of different seasonality classifications. On the other hand, southwards in the contracted stands there is a larger monthly coverage time difference (CTD) such that they tend to underestimate estimated stand volume more.

The analysis was done on a strategic regional level, as regional harvesting and planning levels were more stable from 2013 to 2020. Aggregating regionally simplified the problem and enabled capturing trends over a long period. The aggregation also made it possible to analyse similarly in all observed areas and make the results more comparable. However, further analysis on a tactical level would be needed before implementing the results from this study. The tactical analysis would instead look at data on a local level and consider things this study has set as constant – such as industry demand, composition, rates of harvesting and planning.

Further analysis on a local tactical level could possibly determine a suitable TB size per district and be more operationally applicable to practical forestry. The local knowledge and experience should be utilised when doing this. For example, the seasonality of road and terrain were merged similarly in all regions, and different districts could want to merge differently than done in this study. However, as the weakest part of a stand limits its accessibility, it was considered reasonable. They could also want to have more or fewer seasonality classification. Furthermore, as the analysis of this study was performed over a long period, spatial variables such as administrative area were standardised. When analysing on a local level, the local knowledge of how and why areas have been clustered differently over time could give insights into what data to use.

Examples of local analysis would be to simulate how well demand can be met with current and potential TB sizes and composition (Staland 2001). Another analysis would be to optimise old routes to see how potential routing is influenced by TB size and composition (Jacobsson 2005). Consequently, combining the methods and considering things such as current and potential entrepreneur geography. A significant difference between previous and future studies is the current and future access to large datasets with detailed history. Simulating how different TB sizes and compositions could be optimised in future scenarios regarding industry demand, seasonality and cost minimisation could give insight into suitable TB sizes. Say you find that a TB with three years of harvesting in the TB is optimal; you would also have to consider not keeping stands in the TB for too long. A large TB could potentially have a larger portion of stands that need replanning, adding extra costs. Long storage time potentially also ages the field planning, physical marking bands in field loose colour and knowledge of road conditions decrease as a single significant rain downpour can greatly change road conditions.

Another aspect of having a larger TB would be an increased incentive to use optimisation models for route optimisation. There are algorithms to optimise scheduling, an example being Frisk et. al. 2016, but they are not utilised to a large degree in practical forestry. Possibly as there is too little room for improvement with the constraints that are put on the decision making. Such as already discussed – long term forest plan, entrepreneur satisfaction and delivery plans etc. However, as there would be more available nodes for optimisation with a larger TB, the benefits of optimizing could increase sufficiently for implementation.

4.1. Tract bank size

In all regions and sources, the mean TB of the host company is lower than the levels often referred to in the Swedish industry (table 2; 3). This indicate that those TB levels even if known to many in the Swedish forest industry aren't generally used in practical forestry. Also, the calculated CT (The coverage time of the suitable tract bank that was estimated) (table 7; 8) suggests an even lower TB. However, it should be considered that the uncertainty in a supply chain often is lower in higher planning levels than lower, as the outcomes on lower levels often even each other out. As such a TB calculated on a lower level for local planning will probably indicate a larger uncertainty than is done on the regional. Moreover, as the TB is operationally used locally, the general suitable TB size if there is any should be determined at that level.

Further analysis of the company could also be to look at how the supply structure has changed in the years following the analysis (2021-) as the company has made a large acquisition of wood industries in the northern region. This has probably changed their supply structure to require more flexibility and a larger TB since they now need to provide for their own industries to a larger degree. It would also be of interest to compare this company with one or more organisations that are more heavily clustered in one region, as compared to the host company that is spread over three.

4.2. Storage time and the relative difference in the outcome

The mean CTD (1-7% of the monthly harvest) and stands with storage time that requires replanning (0.1-2.3% of the total TB) both only make up a small part of the monthly uncertainty compared to the calculated CT (5,7-10,5 months of harvest). This validates the tolerance for data deficiencies in the TB that is expressed in Nilsson et al. (2013).

Nevertheless, the stands that must be planned more than once should be considered when setting goal sizes for the TB, especially in region North where they have the largest share (Table 5). Partly as these stands to some extent, were not part of the TB, and if they are stands with crucial characteristics, be it seasonality classification or wood assortment, it could pose a problem. An increase in the TB could counteract the uncertainty in the TB's actual size and composition, but it is no solution as it could risk increasing the absolute size of that unharvested share. Instead, the share could be used as a key performance indicator for the planning process. The planning process is responsible for the long-term prioritisation and clustering of the stands in the TB spatially and temporally. A large share could indicate less good clustering and a small share the opposite. It could also be used in the discussion between the planning and harvesting to evaluate work and the current situation. It can also be seen as a key performance indicator of harvest management, as they should aim to minimise this share.

The mean CTD should also be considered and should decrease the TB size somewhat as stand volume tend to be underestimated. The regional trend that stand volume is underestimated more southwards would suggest that this could decrease the TB somewhat more southwards. Also, it is worth discussing the relationship between the mean CTD and the standard deviation of the CTD. Region north, for example, seems to have a 1:1 relationship (fig. 20, centre plot) between the monthly underestimation and its standard deviation. The relationship indicates that planners here could actively be underestimating the volume to cover the uncertainty of the outcome. If the relationship is genuine, planners actively but skilfully underestimate the estimated stand volume to the same degree that it is wrong. However, if true it might be worth considering changing the operational practice of underestimating the outcome and tackling the same problem of variation of the outcome on a higher hierarchical level instead.

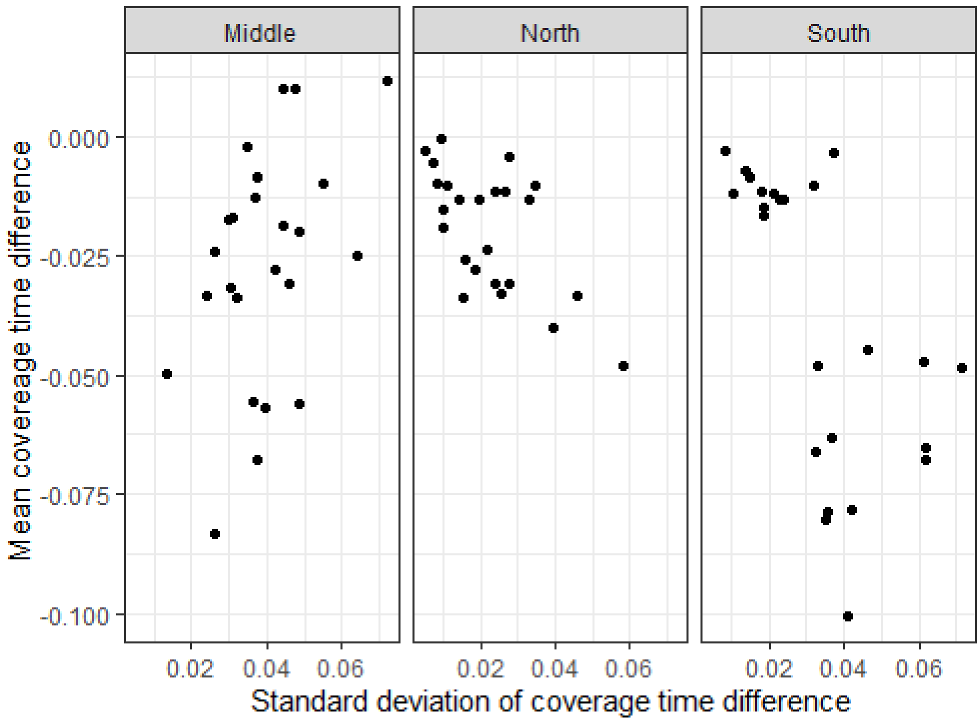


Figure 20. The figure shows three scatterplots of the Mean relative difference of outcome per month (CTD) and region (y-axis) against the standard deviation of the relative difference of outcome per month.

However, as the reverse relationship is indicated in the middle region (fig 20) and no relationship can be seen in the region south (fig 20), it might also be a coincidence of outcome.

4.3. Seasonal indexes of harvest and planning

It is interesting to see the clear differences in how the different sources are utilised within regions but also across all of them. That harvest of contracted spring stands barely shows a seasonal spike in the spring months (fig. 14) or that the dry summer contracted stands are harvested so similarly to winter stands in region north (fig. 15). Moreover, it’s interesting that the estate fall stands so clearly have a seasonal spike in spring covering for that lack of spring stands.

That lower trafficability classes to some degree cover for the lack of higher classes, is not entirely surprising but it does imply that the quality of those stands are not on par with the stands used in the right season. To some extent it is probably not as dramatic as it sounds, that the harvest managers with their experience know: which harvester teams can handle more difficult conditions, which harvest planners tend to classify stands with lower trafficability or what areas tend to be more or less accessible than their seasonality classification implies. Some harvest managers also make their own field assessment and seasonality classification of the stands when time allows. A study looking into the implications of using stands out of season would be of interest. Implications such as an increase in soil damage or increased costs from a slower or more difficult harvest.

That the estate stands are not planned in winter in region north while the contracted stands are (fig 14) and can probably be explained by their differing workflows. The contracted stand has its ecological values assessed before the contract is made while the estate planning is done all at once. This can probably explain how contracted stands continue to be planned during winter despite the snow cover. On the other hand, the estate stands are probably constrained by the continuous snow cover, as the ecological assessment can't be made at a satisfactory level. However, some estate winter stands in the north are still planned in this period. These stands could either be thinning's, other harvests because of abiotic or biotic stand damages or that the plan was finalised in this period and the field planning was done earlier.

The South region has a considerable variation in the seasonal indexes of the south estate harvest compared to middle, north, and contracted harvests (SL in table 7). The variation could result from it being small, smaller than what is suitable or from having an unfavourable composition forcing the utilisation of stands with certain seasonality classifications to be used out of season to a larger degree. By being small, it would take a larger share of its size to cover for deficiencies in the contract bank. It could also be that the difference in climate over time leads to greater variation in the utilisation of the TB. Region south has a warm humid continental climate while the north has a subarctic climate, and the middle region is in the transition between the two (climatedata.org 2022). The longer summer in the south could explain the larger variation to some degree, as winter has more stable weather, frozen ground, and snow. Furthermore, summer has large downpours up to 198 mm in a day (SMHI 2021), which impacts ground and road conditions.

Another way to visualise the seasonal variation is to look at boxplots of the seasonal indexes of the different regions and sources (fig. 21). Region souths estate harvests have noticeably more outliers (fig. 21) than other regions and sources throughout the year. However, perhaps with the exception region Norths contracted harvests, which also has many outliers. These two have in common that they generally have a smaller TB than their counterparts in the other regions (fig 12). These outliers are not isolated to one year either but seem to be reoccurring in most years (fig. 22). The outliers of the smaller TB:s are probably a result of them having to cover a variation in the other source of that region.

Figure 21. The figure shows the seasonal harvest indexes as boxplots coloured by seasonality classification in all regions and seasonality classifications for all the observed years.

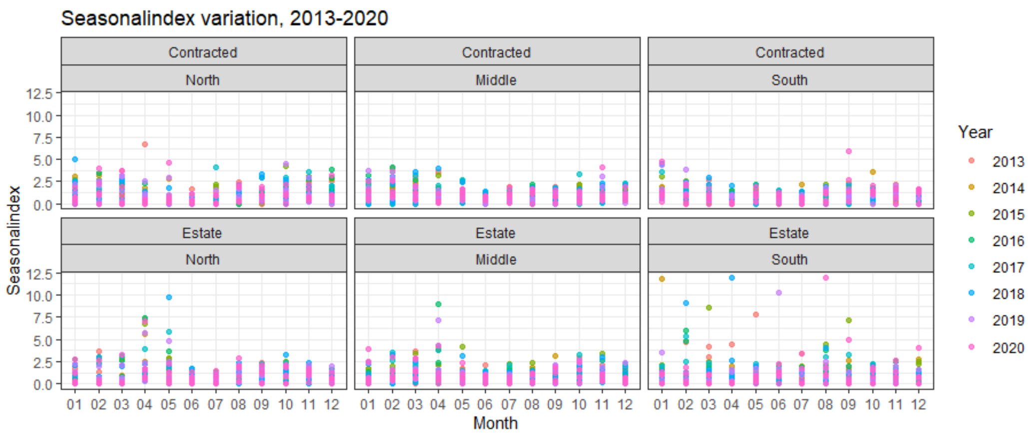


Figure 22. The figure shows the seasonal harvest indexes coloured by seasonality classification in all regions and seasonality classifications for all the observed years.

4.4. The estimated suitable tract bank

The suitable TB as calculated in this study can only estimate a TB size that is expected to be needed before a period of uncertainty starts. The TB estimate could be used as a basis for discussion as to how the size of the actual TB relates to that of the estimated suitable TB. As this study assumes a constant industry demand, yearly harvesting capacity, yearly harvest planning capacity and composition of the tract bank. The Composition is set as constant, but it naturally changes with time partly as the forest is heterogenic. Moreover, as different strategies regarding investment levels in the road network also affects the seasonality of the stands available for harvest. If the forest plan indicates a lack of higher seasonality classifications, then the long-term planning could take this into account and invest more into the road network to reach a suitable composition of the TB as well as size. As stated in previous studies the composition of a TB is just as important as the actual size of it (Staland 2001; Nilsson et al. 2013; Jacobsson 2005; Renström 2008). Also, if for some reason harvest or planning rates are expected to decrease for some reason compared with previous levels then the suitable TB could be adjusted for that potential scenario.

Since data on current and past TB sizes and their composition exists, a tool for visualising it in real-time could be helpful. For example, it would be recommendable to have a powerBI tool or similar for visualising the past and expected change in TB as well as the current TB. For example, it could show the TB has changed over that past year or months and show what is currently available and how it is expected to change with the current plan. Both change in absolute size but also in composition, as it is shown in this study the composition varies throughout the year. It could also forecast how the TB is expected to change with historical values with upper and lower bounds, for example, building on the methodology used in this study.

A decision support tool such as this could be an asset in letting the harvest estimate take a more central role in the day-to-day dialogue. It could visualise the stands in the TB and the stands that the long-term forest planning has planned for the harvest planners to plan. The visualisation could then assist the harvest manager too more easily make a conscious decision when weighing the delivery plan to the harvest estimate. It would also give the higher hierarchical levels more transparency into planning and harvesting and how the processes are currently working towards different goals. The tool could be a part in clearing up how harvests deviating from the long term forest plan to serve the delivery plan are accounted for in the long term (Nilsson et al. 2013).

4.5. Predicting seasonal indexes

Creating models of the seasonal indexes had varying results, but there is potential here to decrease the uncertainty in the utilisation of seasonality classifications and thus also decrease the TB. For example, the utilisation of winter stands in the middle and north region could be predicted to a large degree. That potential predictive power could also be greater when having more data and tailored parameters for a particular region, district, or production unit. Also, exploring a more comprehensive array of ML algorithms, as they all have their pros and cons, could potentially increase the predictability of the indexes.

This study used the random forest algorithm with weather data from SMHI and wood prices from FAO. The weather data used was monthly data on temperature and precipitation, but weather data on wind speed, snow depth and ground conditions, vegetation period and sunshine could also have been used. Large spikes in the wind indicate a storm and would then force harvest in some stands. The snow depth enables the use of snowmobiles in the planning but limits the possibility of reaching the stands with a car for the planners; it also affects the scheduling of stands as it affects the harvesting operations. Data on the vegetation period and sunshine could give the models insight into the evapotranspiration, and in so doing, the ground conditions as the precipitation levels are also known. More detailed data on wood prices could also be used as the numbers used in this study were one general number per country, month, and species. Wood prices could probably explain more than the general numbers provided here on the assortment, district and company level. For example, models were also made on a production unit level, which indicates that the monthly mean precipitation is the most important variable in the spring harvest estate indexes of the production units of region north (fig. 23).

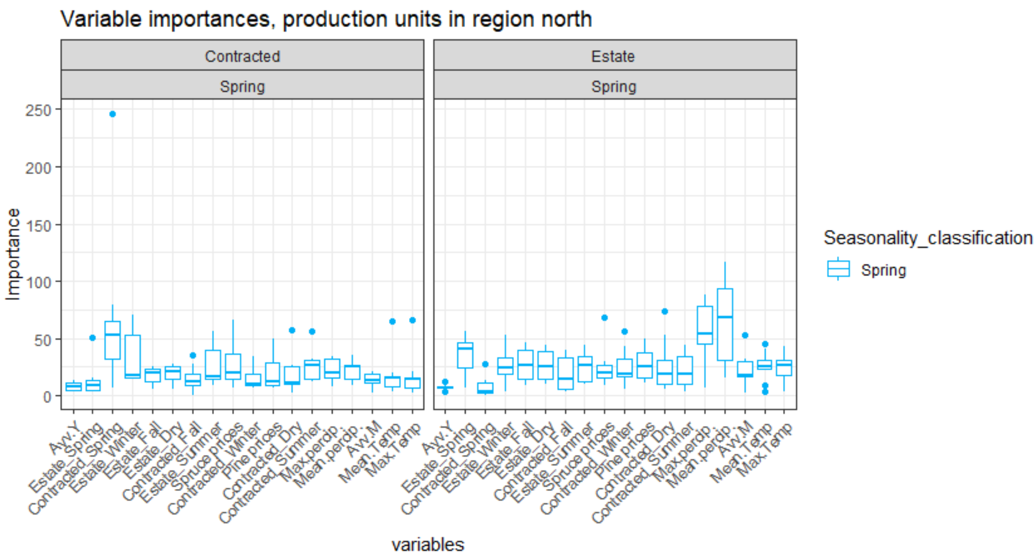


Figure 23. The figure shows boxplots for variable importance for models that predict the seasonal harvest spring indexes in the production units of region north. The Models themselves do not explain a large part of the seasonal indexes, so the variable importance should only be seen as an indication of what variables that are important for future studies.

However, since the rsq for these models are low, the variable importance should only be seen as an indication that there could be a possible relationship. On the other hand, the relationship is very reasonable, as precipitation affects decision-making to a large degree. It directly affects ground conditions in stand terrain and roads and the risk of forest fires. It is also reasonable to explain variation on the local level as rain will only affect the area with rain and not the production unit 30 km away.

5. Conclusion

To conclude, this study explored how on a strategic level the tract bank (TB) per geographic region and stand source. The study also applied concepts from inventory theory (Lumsden 2019) to regional planning of the TB.

Historically, based on the case data, the TB in months of harvest has on average been: 13,5 in the north; 9,5 in the middle and 5,8 in the south, while the estimated suitable TB in this study was 9,5 in the north; in the middle 5,9 and in the south 7,1. Where the smaller historical TB size in the north and middle region to some extent can be explained by the often-smaller uncertainty on a higher planning level, compared with the local level at which the TB is dimensioned. The estimated suitable TB in the south is larger than the past average.

In all observed areas and sources, estimated stand volume is systematically underestimated too varying degrees.

The study also found that forestry is noticeably more seasonally constrained in company-owned stands in the north and middle region compared with the south region and contracted stands. Moreover, that temperature and month of the year are the most important variables in explaining year-to-year seasonal index variations on the regional planning level. These results are relevant to those that work with the tract bank and could be used in further research of the tract bank.

References

1. Holmen (2020). Holmen Årsredovisning 2020.
2. Wildmark, L. (2014). Flyttavstånd och flyttkostnader för skogsmaskiner samt ruttplanering hos Södra Skogsägarna. Umeå: Sveriges lantbruks universitet.
3. Bugge, M, T Hansen, & A Klitkou (2016). 'What Is the Bioeconomy? A Review of the Literature. in Sustainability (Basel, Switzerland), 8, 691–.
4. Söderholm, J. (2002). *De Svenska Skogsbolagens System för Skoglig Planering*; Swedish University of Agricultural Sciences: Umeå, Sweden; p. 56.
5. Church, R (2007). Tactical-Level Forest Management Models. In Handbook of Operations Research in Natural Resources; Weintraub, A., Romero, C., bjørndal, T., Epstein, R., Eds.; Springer: New York, NY, USA; Volume 99, pp. 343–363.
6. Epstein, R.; Karlsson, J.; Rönnqvist, M.; Weintraub (2007). A. Harvest Operational Models in Forestry. In Handbook of Operations Research in Natural Resources; Weintraub, A., Romero, C., Bjørndal, T., Epstein, R., Eds.; Springer: New York, NY, USA; Volume 99, pp. 365–377.
9. Willén, E. Friberg, G. Skarpman Sundholm, J & Webber, L. (2021) ARBETSRAPPORT 1071-2021 Planeringsrobot för traktplanering Automation in forestry site planning. Skogforsk. Uppsala, Sweden.
10. Erlandsson, E, (2016). *The triad perspective on business models for wood harvesting tailoring for service satisfaction within forest owners associations*. Umeå, Department of Forest Biomaterials and Technology, Swedish University of Agricultural Sciences.
11. Lumsden, K. Stefansson, G. & Woxenius, J. (2019). *Logistikens Grunder*, Lund, Studentlitteratur AB.
12. Breiman, L (2001). Random Forests, Machine Learning 45(1), 5-32.
13. Breiman, L (2002). "Manual On Setting Up, Using, And Understanding Random Forests V3.1", https://www.stat.berkeley.edu/~breiman/Using_random_forests_V3.1.pdf.
14. SMHI (2022a). *Ladda ner metrologiska observationer.* https://www.smhi.se/data/meteorologi/ladda-ner-meteorologiska-observationer/#param=precipitationMonthlySum,stations=active,stationid=188790. [2022-01-20].

15. Food and Agriculture Organization of the United Nations (2022). *Forestry Production and Trade*. https://www.fao.org/faostat/en/#data/FO .[2022-01-24].
16. Roth, M. (2010). *Ledtidens betydelse för privata skogsägares kundnöjdhet i samband med gallring*. (48 s). Umeå: Skoglig resurshushållning.
17. SMHI (2022b). *Jordens huvudklimattyper.* https://www.smhi.se/kunskapsbanken/klimat/jordens-klimat?fbclid=IwAR07q5rVDbI010fjFP16ravrIUCmxkl761TBXL4Q9SpkZtf17Ko9nvQmmg0. [2022-05-20].
18. SMHI (2021). *Sveriges klimat*. https://www.smhi.se/kunskapsbanken/klimat/sveriges-klimat/sveriges-klimat-1.6867. [2022-05-20].
19. Climate-Data.org (2022). *Climate: Sweden.* [Table]. Available: https://en.climate-data.org/europe/sweden-105/?fbclid=IwAR0fiEF2i--BfbaUlR2cr7SVnSWo7aMExqB51BWY53UtJEdCYe1PgMKO6Oc. [2022-05-20].
20. Dag Fjeld, Bo Dahlin (2017). *Nordic logistics handbook – Forest operations in wood supply*. Swedish University of Agricultural sciences, Helsinki University.
21. Jakob Staland (2001). *Control of timber flows tailored to customer needs - Influence of size of tract register and accuracy of log tally projection.* Sveriges lantbruks universitet. Umeå: Institutionen för skoglig resurshushållning och geomatic.
22. 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: Sveriges lantbruksuniversitet.
23. Jacobsson, L (2005). *Förbättringspotential i avverkningsplanering -En fallstudie av ett års avverkningar på två distrikt inom SCA skog, Jämtlands förvaltning.* Umeå: Sveriges lantbruksuniversitet.
24. R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
25. Frisk, M. Flisberg, P. Rönnqvist, M. & Andersson, G. (2016). *Detailed scheduling of harvest teams and robust use of harvest and transportation resources.* Scandinavian Journal of Forest Research, 31:7, 681-690, DOI: 10.1080/02827581.2016.1206144
26. James, G. Witten, D. Hastie, T. & Tibshirani, R (2021). An Introduction to Statistical Learning with Applications in R. Second Edition. Pages 327-361.
27. Aggeryd, B. (2009). Optimalt råvarulager för biobränsleföretaget. Umeå: Sveriges lantbruksuniversitet.
28. Renström, J. (2008). Säsongsvis avverkning det operativa traktvalets påverkan på den säsongsmässiga uthålligheten i ett bolagsdistrikts traktbank. Umeå: Sveriges lantbruksuniversitet.

29. Larsson, E.W. (2003). Funktionsspecifikation till Holmen Skogs framtida drivningsweb = Specification of functions on a future contractor's web made for Holmen Skog AB. Umeå: Sveriges lantbruksuniv.
30. Nilsson, S. (2008). Dimensionering av industrilager för biobränsle. Umeå: Sveriges lantbruksuniversitet.
31. Erlandsson, E. & Fjeld, D. (2017). Impacts of service buyer management on contractor profitability and satisfaction - a Swedish case study. International journal of forest engineering, 28 (3), 148–156. https://doi.org/10.1080/14942119.2017.1367235
32. Liu, Z., Peng, C., Work, T., Candau, J. N., DesRochers, A., & Kneeshaw, D. (2018). Application of machine-learning methods in forest ecology: recent progress and future challenges. Environmental Reviews, 26(4), 339-350.
33. Jones, M.-F., & Arp, P. (2019). Soil Trafficability Forecasting. Open Journal of Forestry, 9, 296-322.