EFFORTE –
'Efficient forestry by precision planning and management for sustainable environment and cost-competitive bio-based industry'

This project has received funding from the Bio Based Industries Joint Undertaking under the European Union's Horizon 2020 research and innovation programme under grant agreement No 720712
Project duration: 1.9.2016–30.8.2019.
Coordinator: Natural Resources Institute Finland (Luke).

Deliverable 3.7 – Harvester information for micro compartments		
Work Package 3 - Big Data bases and applications		
Task 3.3 Improving efficiency in silviculture operations		
Due date	28.02.2019	
Author(s), organisation(s)	Johan J Möller, Nazmul Bhuiyan, John Arlinger, Erik Willén (Skogforsk)	
Date of publication	28.02.2019	
Dissemination level		
PU	Public	X
PP	Restricted to other program participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	
Nature of the Deliverable		
R	Report	R
P	Prototype	
D	Demonstrator	
O	Other	

Big Data bases and applications

Harvester information for micro compartments

28 February 2019

Table of contents

1. EFFORTE project objectives

EFFORTE is a research and innovation project providing the European forestry sector with new knowledge and knowhow that will significantly improve the possibilities of forest enterprises to assemble and adopt novel technologies and procedures.

The project aims at enhancing the efficiency of silviculture and harvesting operations; increasing wood mobilization and annual forest growth; increasing forest operations' output while minimizing environmental impacts; and reducing fuel consumption in the forest harvesting process by at least 15%.

The project is based on three key elements of technology and knowhow:

1) Basic understanding of fundamentals of **soil mechanics and terrain trafficability** is a crucial starting point to avoid soil disturbances, accelerate machine mobility and assess persistence of soil compaction and rutting. The key findings and recommendations of trafficability related to EFFORTE can immediately be adapted in all European countries.

2) Due to decreasing Cost-competitiveness of manual work and maturity of technology it is now perfect time to realize the potential of **mechanization in silvicultural operations**. EFFORTE pursues for higher productivity and efficiency in silvicultural operations such as tree planting and young stand cleaning operations.

3) 'Big Data' (geospatial as well as data from forestry processes and common information e.g. weather data) provides a huge opportunity to increase the efficiency of forest operations. In addition it adds new possibilities to connect knowledge of basic conditions (e.g. trafficabillity), efficient silviculture and harvesting actions with demand and expectations from forest industries and the society. Accurate spatial information makes it possible for forestry to move from classic stand-wise management to precision forestry, i.e. micro stand level, grid cell level or tree-by-tree management. EFFORTE aims at achieving substantial influence to the **implementation and improved use of Big Data within Forestry** and through this increase Cost-efficiency and boost new business opportunities to small and medium size enterprises (SME) in the bioeconomy.

EFFORTE researchers will develop and pilot precision forestry applications that, according to the industrial project partners, show the greatest potential for getting implemented immediately after the project.

2. Introduction

This document describes how to use harvester data to derive microcompartments for forestry applications. Microcompartments are parts of the forest with similar forest specifications, e.g same tree height or tree species composition. Traditionally forest stand descriptions have been summarised on 2-20 hectares in Nordic forestry, but with more high resolution digital data being available it is possible to much more detailed site-specific measures.

In a parallel application area, such as agriculture, detailed data about growth within agricultural fields is used to control how much and where to apply fertilizers which can improve growth as well as lower costs for fertilizers and give less impact on the environment when fertilization is optimized based on more detailed growth conditions.

In forestry this could also be a reality although fertilization is quite limited in Sweden and Finland. There are however number of applications that currently reach implementation based on more details about the forest stands. Airborne laser (LIDAR) maps the height of the forest which makes it possible to make detailed delineation of forest stands. Forest estimates of height and basal area from LIDAR data are used in order to calculate thinning index that provides details on where, within a stand, there are needs for thinning (Willén et al 2017).

Harvester data provide details about the forest that has been cut but may also provide information about the remaining forest in a thinning as well as general growth conditions. This data may be used for effective update of forest stand databases and can also be used for detailed forest management regimes. More details in the silviculture planning could propose adjustments of the soil scarification machine and different number and types of seedlings in different parts of the stand.Harvester data may be used after thinning for describing the current forest and after final felling for planning regeneration of a new forest stand. The forest industry (sawmills and pulpmills) and the new biorefineries requires more detailed information about deliveries. Harvester data is crucial for extracting relevant information about the raw material that is about to be delivered.

Additional use of microcompartments could be as reference data for remote sensing estimates. The large amount of reference data available might support improved accuracy in tree species mapping or yield estimates for planned forest cuttings (Söderberg et al).

Microcompartments lead to more precise forest management, better description of the forest and improved descriptions of wood products to be delivered to the industry.

The aim of this report was to describe a method for using harvester data to derive microcompartments is described in this report. A scientific publication is also prepared to describe the method in detail for thinning operations.

Microcompartments are to be validated in several Efforte reports. Report D2.6 describes different economic unit sizes based on harvester data. Report D3.8 describes a tool using microcompartments from harvester data for silviculture planning and the results are validated in report D4.6, *Validation of cost-efficient and productive silviculture.*

3. Material and methods

Harvester data

Modern forest machines for logging according to the cut-to-length method are equipped with control systems (machine computers). The control systems assists the operator when processing stems into forest products that industrial customers then process into sawn products, pulp, paper and energy. The forest machines are controlled by digital instructions, and production data is stored in the control system. Most of the data flows from and between forest machines are managed according to StanForD, the standard for forest machine data and communication (Arlinger et al. 2012). Today, StanForD is a global standard for cut-to-length machines and is used by all major manufacturers. StanForD2010 is the updated version of the StanForD standard that has been available since the end of the 1980s.

Figure 1. Single grip harvester in action. © Skogforsk.

Harvester data collected with StanForD provides some specific features of interest for making microcompartments:

XML-format

StanForD 2010 uses XML-format for storing and reporting information. XML is an open, general format that is used in many applications where data needs to be stored and communicated.

For software developers, XML has the advantage that there are already many available solutions for parsing, writing and managing XML-files, which saves time and development resources. In addition, files can easily be checked against an XM-schema to ensure that they comply with the standard. This simplifies the possibility to produce microcompartments using the structure and content of the XML-files.

Flexible production control

In principle there are different methods for managing digital harvesting instructions in forest machines. The simplest way is to manually send out product and object instructions to the machine before logging and then not do any changes to them during harvesting at the harvesting object (work area/logging site). .

Another more flexible method is to allow the instructions to be updated at any time during logging. When a modified version of an instruction enters the harvester database, the operator is asked whether he/she wants to update an existing instruction or reject the update. The flexible method allows the logging organization to quickly redirect production by, for example, changing the length breakdown for a product or activate/deactivate certain products in line with changes in demand. The introduction of Keys (generate automatically in the harvester) and UserIds (set by the user), together with the default of production reporting per produced log, makes this flexible method possible. It is e.g. a ProductKey that signals the conditions under which a stem is bucked into products. Production data from several loggings can now be aggregated, even if different product instructions are used.

Examples of Keys are MachineKey (a unique identification to each forest machine), StemKey (consecutive number for each stem processed) and ObjectKey (generated when a new logging object is created in the machine computer) (Arlinger et. Al 2012). They are also a prerequisite for making microcompartments as described in the coming sections.

Detailed production reporting

In StanForD 2010 production reporting is done at an individual log level (hpr message), allowing more detailed analyses of the forest products. Each log is described with dimensions, product etc (illustrated in figure 2) so reporting of the produced volumes can be customized. What is of interest to Sawmill A is perhaps of less interest to Sawmill B and not relevant at all for Pulpmill C, even if the logs originate from the same logging site. Continues reporting of production in combination with a flexible method for controlling logging allows customized production where lead times are short between a change in demand and an update product instructions/specifications.

Based on detailed production reporting it is possible to calculate forecasts of forest fuel availability as well as giving feedback about the harvesting object that the landowner can use in forest management planning. Statistical models can be used to calculate properties such as density, heartwood content and knot structure in a delivery. For planning of forwarding, reporting by log with possible associated time stamp and GNSS (Global Navigation Satellite Systems) position provides faster information about what has been logged during a certain time period and where timber can be fetched.

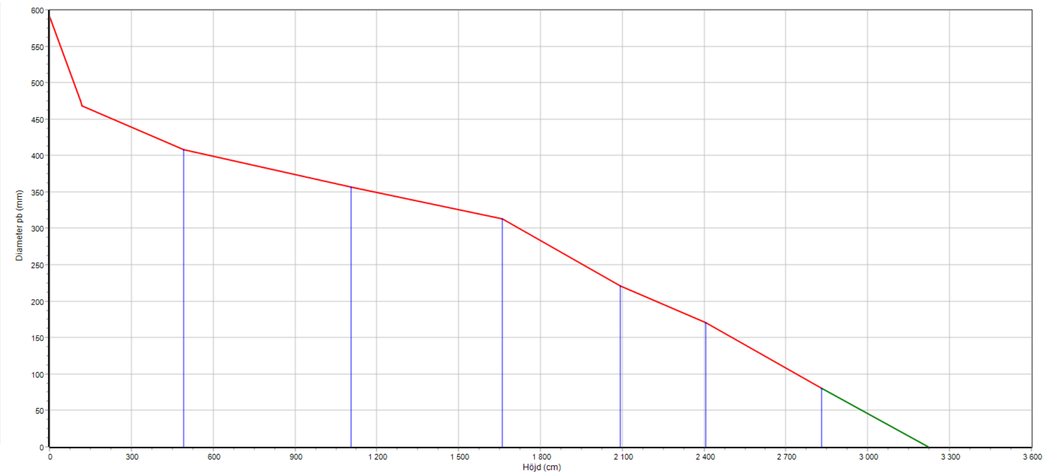

Figure 2. Detailed reporting of every stem harvested. Tree height on x-axis and tree diameter on y-axis.

Each stem is reported in detail. Tree species is set by the operator for each stem while harvesting. Figure 2 presents an example with tree height (in centimetres) on x-axis and tree diameter (in millimetres) on the y-axis All measurements are done by the harvester head. Blue lines indicate each log. The first cut is at roughly five meters with a diameter of about 400 mm. This stem is about 32-33 meters long (height while standing) and is divided into 6 logs. In this example the first 4 logs were saw log and the remaining two logs were pulpwood. The green line indicates the modelled tree top and could either be collected as biofuel or left in the forest (Möller et al 2011).

Root rot occurrence

Root rot in spruce is caused by a fungus in the forest soil that disintegrates the wood quality and primarily affects Norway spruce (*picea abies*) in Nordic countries. At harvest each tree is first cut at it's base in order to optimize the outcome of timber volume. But the harvester operator continually keeps an eye on root rot occurrence at the base cut. If a tree is found to be rotting at its base, the operator will discard as much of the lowest part of the log as necessary in order to only collect fresh timber. This manually directed cut is detected and logged in the harvester file, and during the file processing it is interpreted as occurrence of a rotting stem base. When calculating root rot occurrence in spruce forests it is assumed to be root rot when the first log is cut manually into a low value product such as pulp or energy

(Möller et al 2011). It might be manually cut for other quality reasons, but the overall dominating cause for manual cuts in the first log in spruce is root rot.

The interpretation of tree height and decay occurrence is visualized in figure 3, which also clarifies the estimation of volume for each cut product.

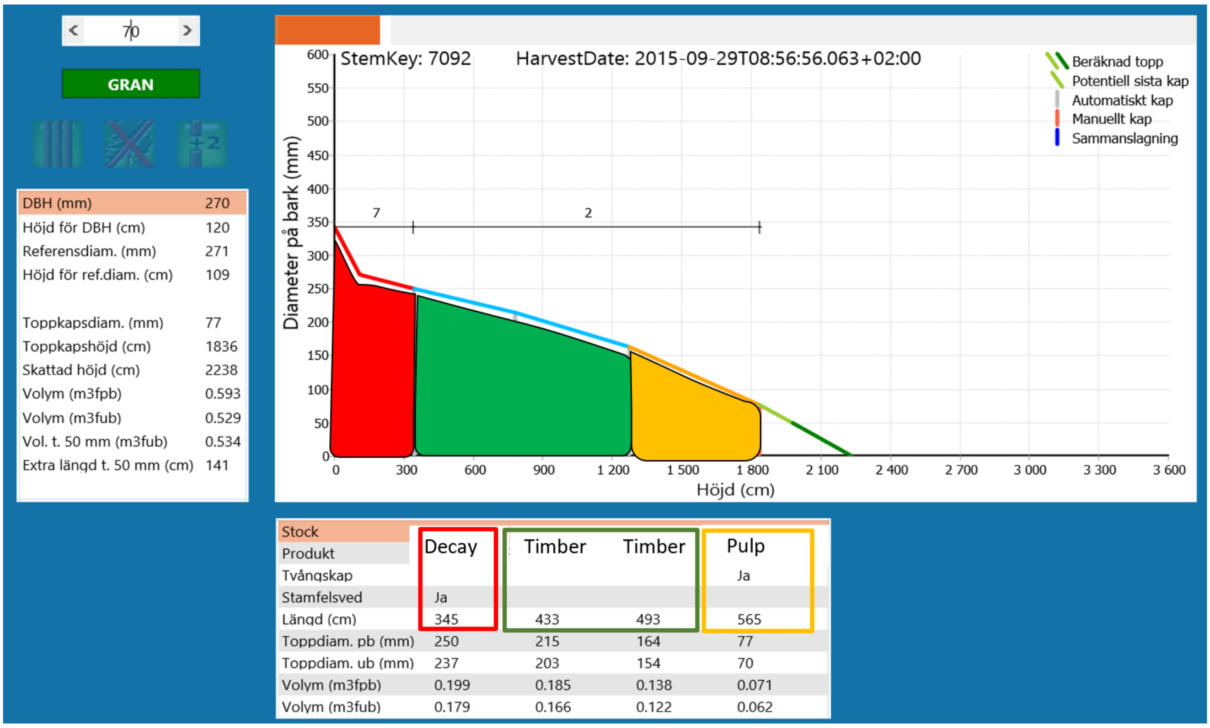

Figure 3. Visualization of the harvester data, hpr. Stem diameter on y-axis [mm] and stem length on x-axis [cm]. In this particular case a cut was inserted at 345 cm which is interpolated as root rot. After that two cuts were made for timber lengths of 433 cm and 493, and the final cut gave rise to a 565 cm piece classified to be used for pulp. Volume of each product is calculated by integrating the curve at each section.

Quality assurance and calibration

In order to ensure that the harvester and forwarder systems for measuring length, diameter and weight are accurate, there are procedures for assuring measurement quality. For harvester measurement, quality assurance includes making random control measurements of a number of stems and comparing the results with the machine data. An independent auditor making regular follow-ups can also be linked into the system. For the forwarder, quality assurance comprises checking the weight scale.

Positioning

The GNSS antenna is placed on the top of the harvester cab to be up in the air to get better conditions in retrieving satellite signals, but also to be out of the way while harvesting. Each stem harvested gets a position based on the position of the harvester. A harvester crane reaches 7-9 metres, but normally the distance is shorter than 7 metres. It means that you often get several trees with similar location based on the harvester position. Figure 4 shows an example. Current R&D at manufacturers of forest machines includes methods for calculating a more precise location for each stem based on GNSS and machine data of crane angle and crane length during harvesting.

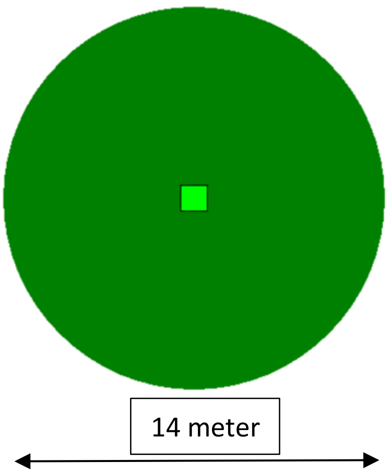

Figure 4. Position of the harvester (light green) that is the registered position in the harvester data and the area where the crane reaches in green. Radius is 7 metres in this example.

GNSS positioning errors in the forest varies depending on availability of satellite signals and possibilities for correction of satellite signals. Empirical experiences of forest machine positioning often show positioning errors of 5-10 meters in the forest (Kaartinen et al 2015, Cedervind 1997).

Processing of harvester data

The processing of harvester data into microcompartments is performed in two steps:

The pre-processing includes filtering and rebuilding of stems by each trees logs and there is also a top added (Möller et al 2009). The calculation of tree height from the last cut to the tree top as the green section in figure 2. It is based on tree species specific functions that utilise harvester measurements (Kiljunen, 2002):

d_1 = tree diameter 1 meter above first cut (mm)

d_2 = *tree diameter 2 meters above first cut (mm)*

h = *tree height without the stump (dm)*

$h_{pulpwood}$ = *height from stump to last cut (dm)*

$d_{pulpwood}$ = *tree diameter at last cut (mm)*

$d_{pulpwood-1}$ = *tree diameter 1 metre below last cut (mm)*

$h_{pine, spruce, birch}$ = *tree species specific top tree height functions*

$h_{pine} = 93.780 - 24{,}579 \times Ln(d_2) + 1.093 \times h_{pulpwood} + 25{,}374 \times (d_{pulpwood}/d_{pulpwood-1}) + 0{,}507 \times d_{pulpwood}$

$h_{spruce} = 69{,}244 - 20{,}755 \times Ln(d_1) + 0{,}686 \times d_{pulpwood} + 1{,}086 \times h_{pulpwood} + 21{,}651 \times d_{pulpwood}/dpulpwood\text{-}1)$

$h_{birch} = 160{,}388 – 37{,}517 \times Ln(d_1) + 1{,}151 \times h_{pulpwood} + 0{,}832 \times d_{pulpwood} + 10{,}533 \times (d_{pulpwood}/d_{pulpwood-1})$

The function provides top tree height and might also be used for tree volume calculations when a cone is assumed with the diameter at the last cut and the calculated tree length (Möller et al 2009).

In addition, the pre-processing also includes merging of harvester data as several files might cover the same object as well as several quality checks on harvester data such as removing of possible duplicates by using the Keys in harvester data.

The next step, *area calculation,* aims at providing reliable area statistics prior to the segmentation and is a key step also for further processing of the harvester data as all forest statistics (basal area, number of stems and more) are area-based (Bhuiyan et. Al.2016).

For area calculation two alternative grid methods have been tested:

1. Grid cells based on position of harvester data
2. Grid cells based on harvester data, including crane angle

The first method for area calculation (Möller et al. 2015) is based on harvester positions by putting a grid on the harvester position coordinates (normally 13 meters spacing, but possible to adjust). Grid cells are used to first summarise the harvested area. trees in an appropriate minimum area (crane length) and reduce positioning errors in individual records. In every grid where trees been harvested an area is allocated. However, the grid area might be reduced to compensate along borders of the cutting or sparse parts of the forest. Figure 5 illustrate how this area calculation is performed with the grid and harvester positions (yellow). Dark green areas indicate grid cells with harvester positions where at least one stem was harvested. Light green areas, marked 1, indicate cells where there is no harvester position but has at least six neighbouring cells where trees were harvested. Red areas and light green areas with other number than "1" is split, they only contribute partly to the overall area. In the red grid cells there are trees cut. If they have less than 6 neighbouring cells the size is reduced according to the number of neighbouring down to 1/9 part if it is no neighbouring grid cells with trees in. The light green, empty grid cells, is reduced if they have four or five neighbouring grid cells.

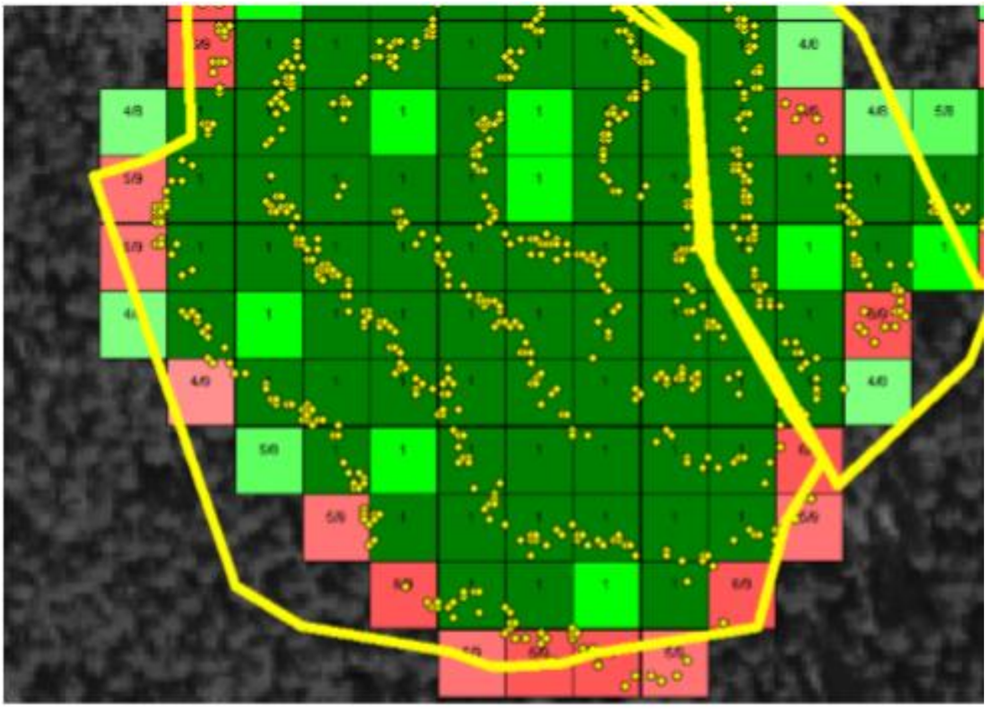

Figure 5. Area calculation principles.

If a crane angle is registered (second method) in harvester data then the stems can be distributed accordingly, figure 6. This option has recently been implemented in harvesters (Bhuiyan et al 2016). If the crane angle is 0 the position is the same as the harvester position, but when the crane position is +/- 90 degrees the position will be spread in the cranes actual direction as the average crane length, 7-8 m. It will then be possible to use a more detailed grid than 13 meter spacing and about 7 meter is proposed.

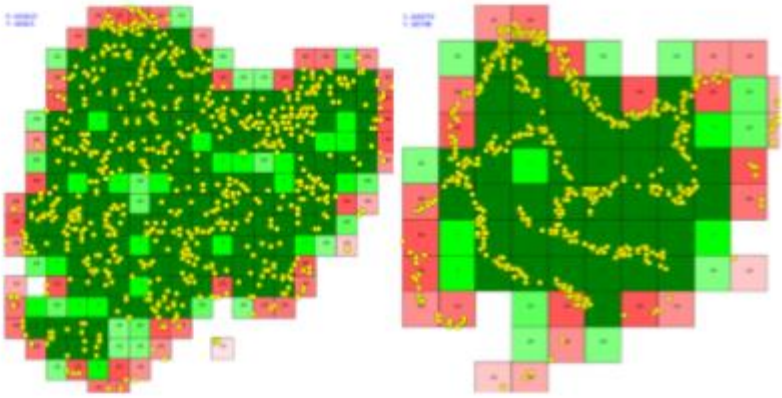

Figure 6. Same object with area calculation based on crane length (left) and only harvester position (right).

In order to evaluate grid spacing Skogforsk conducted a study in 2016 (Bhuiyan et al) where 59 sample plots were put in forests with thinning as the upcoming management regime. The sample plots were distributed in different parts of Sweden and different harvesters used. 26 of the sample plots include crane angle positions. The quadratic sample plots sizes varied from 7-10 m when crane angle was included and 12-16 meters when only harvester position was available. A relative area was calculated showing the estimated area from the grid cells compared to manual GNSS measurements of the sample plots. Also a standard deviation was calculated based on the area deviations. The results showed an increased relative area with larger grids regardless crane angle or not. This was explained by the fact that larger grids will include more area also outside the sample plots. Regarding the standard deviation it was higher for the smallest and largest sample plots and lowest for about 8-9 meters using the crane angle and 13-14 meter without (Bhuiyan et al 2016). These results actually also reflect the possible GNSS position error in the forests. As described, experiences reveal that it has been around 5-10 meter positioning accuracy for harvesters in the forests, below canopy, and is therefore also suitable as minimum grid spacing for area calculations. Recent development with improved GNSS systems such as Galileo and Bei-Dou may increase the accuracy and possibly improve area calculations shortly.

The final step, segmentation, is based on the results from the area calculation and dominant height calculation (Möller et al 2011). The segmentation may actually be applied to most of the collected parameters in harvester data. One useful segmentation is based on upper height also used for site index. The accurate of different area calculation model show that to use the coordinate of the machine gave similar accuracy of area calculation as using the crane angel (Bhuiyan et al 2016). That result mean that we choice the machine position which all forest machine model can produce in StanForD files.

The principles of the algorithm are (Möller et al 2015):

- Assign harvester data to each grid cell
- Split the object in parts along grid cells, but there must be at least 100 trees in each part as upper height site index should include 100 trees or more. This step is iterative.
- Compile tree lists for each segment and by summing up volumes for spruce and pine the dominant tree species is decided. The dominant height is set to 90^{th} percentile of the trees to avoid unrealistic upper height. The prerequisite for this last step is that it must be at least 30 stems of the dominating tree species with the highest volume otherwise the dominating tree species is set to the tree species that dominate the number of stems.
- In order to then derive microcompartments based on dominant height grid cells with similar dominant height is merged into sub- or microcompartments with a minimum area that is possible to specify, normally 0,5-2 hectares for practical use, se figure 7-9 below.

4. Results and discussion

Figure 7 show the resulting from a harvester production file covering a harvesting object. The colour of the dots represent different tree species. Within the same object there are substantial differences showing the need to divide the object into microcompartments for more site specific forest management.

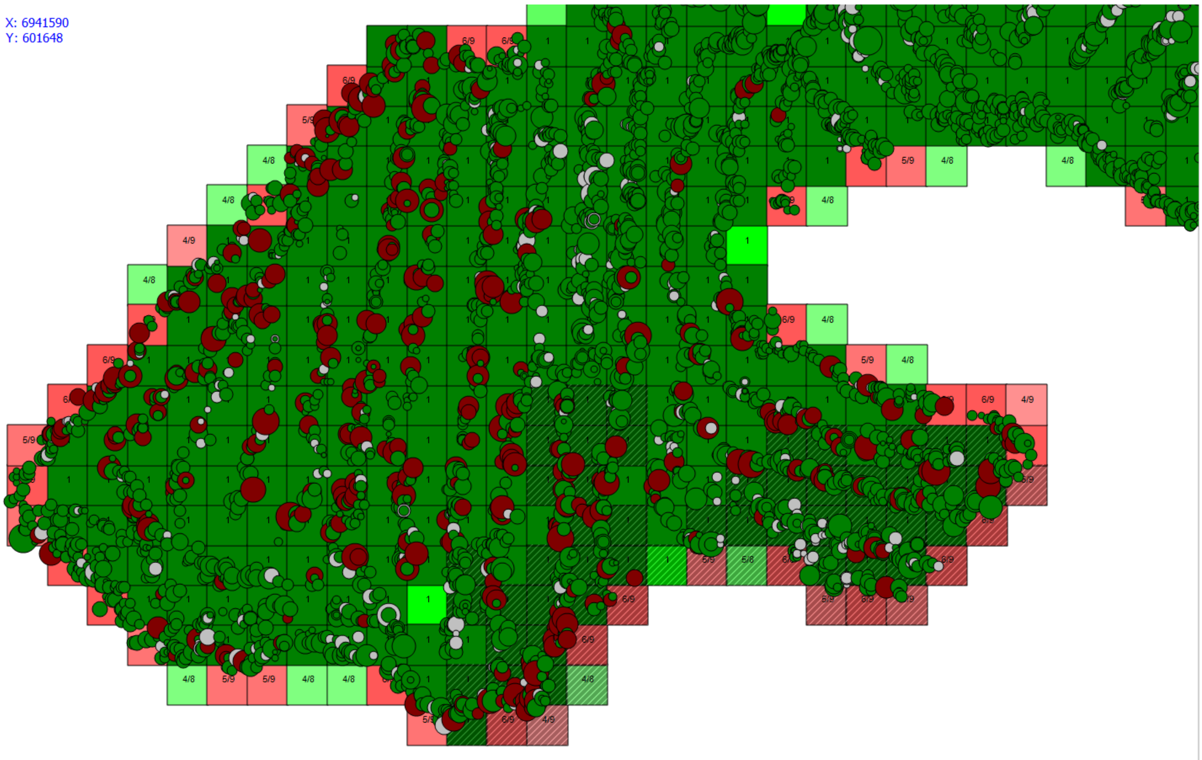

Figure 7. Harvester data showing stems in an object. *Red dots indicate pine trees, green dots spruce trees and white dots broadleaved trees, the position is the coordinates for the harvester. The size of the dot indicates stem size. In the figure is also the area grid for the object shown.*

Figure 8 show the results after calculations of dominant height in grid cells and then the large variation within the object, the forest stand, is evident. In the west part tree height are about 18-19 meters and in the east part about 22-24 meters. This often reflect various site conditions such as soil moisture or soil depth.

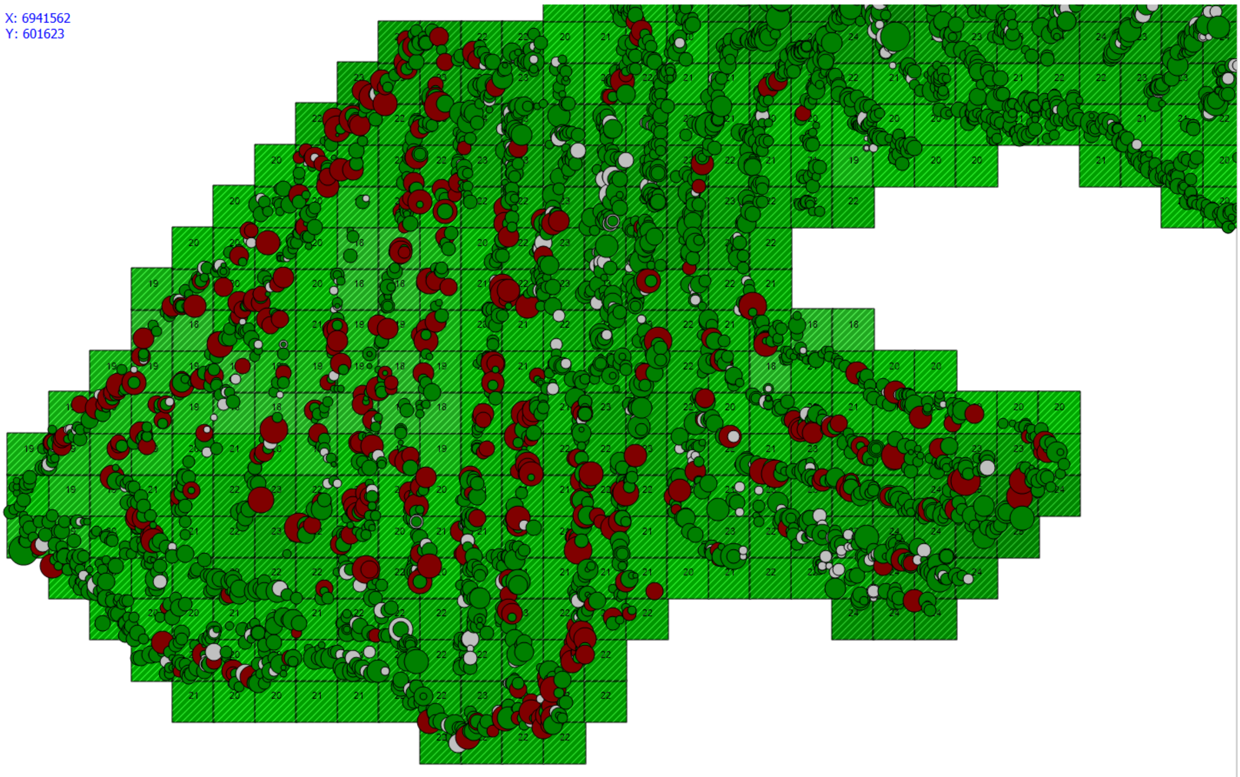

Figure 8. Dominant tree height from harvester data in grid cells. Red dots indicate pine trees, green dots spruce trees and with dots broad leave trees, the position is the coordinates for the harvester.

In order to derive microcompartments it is possible to segment grid cells with similar dominant height to 0.5-2 hectares, as shown in figure 9. Similar dominant height reveals approximately equal growing conditions. After segmentation forest variables may be summarised within the microcompartment. The microcompartment in this example (figure 9) includes 29 % pine, 65 % spruce and 6 % broadleaved trees. The basal area 26.8 m^2/ha, basal area weighted height is 19.7 meters and the mean DBH is 260 mm.

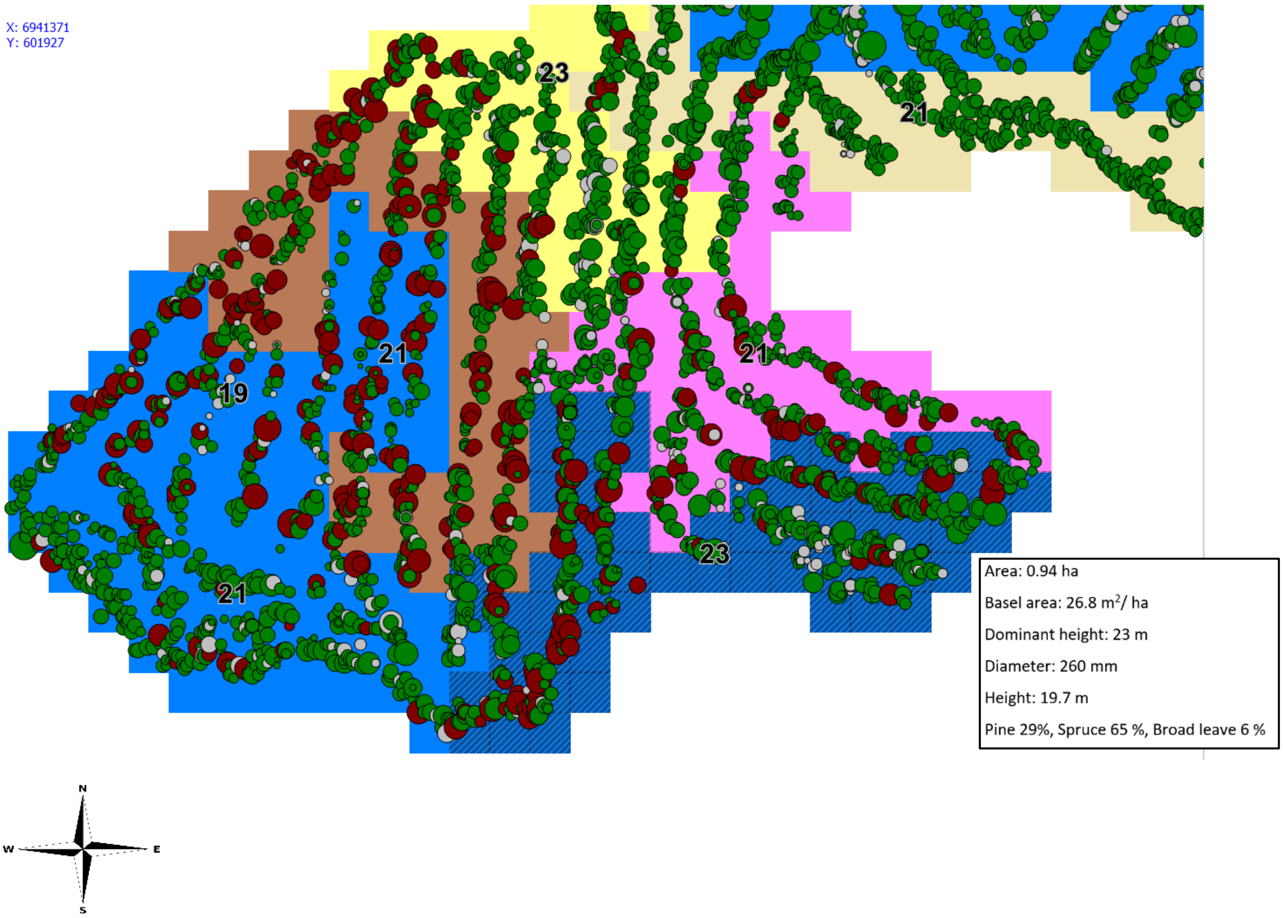

Figure 9. Microcompartments from harvester data based on dominate height. Microcompartments divided in different colours with average dominate height. Forest variables shown reflects the dark blue south- east microcompartment.

Microcompartments from harvester data is possible to use after thinnings for further forest management. Current thinning regimes include 1-3 thinnings in Sweden and Finland and microcompartments may be used for detailed feedback to forest management GIS-systems, but also to describe content and possible management regimes for coming thinnings and final fellings (Möller et al 2015, Hannrup et al 2015).

Another application would be to use the microcompartments from harvester data after final felling for more detailed silviculture planning. This tool is described in Efforte D3.8 and undergo validation to be reported in D 4.6.

Of interest for silviculture planning is the occurrence of root rot (Möller et al 2011). As described in this report it is defined as manual cuts of spruce of the fist log. Figure 10 presents an example of a forest stand with large root rot frequency. Based on this information it might be possible to alter the tree species composition when re-planting after the final felling. Information on root rot frequency may also be derived in microcompartments in order to propose alternative forest management in large forest stands.

Figure 10. Root rot frequency(white dots) in a spruce forest stand

5. References.

Arlinger, J., Nordström, M., Möller, J.J. 2012. StanForD 2010. Modern kommunikation med skogsmaskiner. – StanForD 2010. – Modern communication with forest machines. Report 784, Skogforsk.

Bhuiyan, N., Möller, J.J., Hannrup, B., Arlinger, J. 2016. Automatic follow-up of thinning. Stand area estimation and use of crane angle data to identify strip roads trees and calculate thinning quotient. Report 899. Skogforsk.

Cedervind, J. 1997. GPS under krontak (below canopy). Master thesis. Report 20. Swedish university of Agricultural Sciences.

Hannrup, B., Bhuiyan, N., Möller, J.J. 2011. Evaluation of a system för calculation and feedback of harvester based information to forestry planning system. Report 757, Skogforsk.

Hannrup, B., Bhuiyan, N., Möller, J.J. 2015. Nationwide evaluation of a system for automated follow-up of thinning. Report 85. Skogforsk.

Kaartinen, H., Hyypää, J., Vastaranta, M., Kukko, A., Jaakoa, A., Yu, X. 2015. Accuracy of kinematic positioning using Global Satellite Navigation Systems under forest canopies. Forests, 6 (9), 3218-3236.

Kiljunen, N. 2002. Estimating dry mass of logging residuals from final cuttings using a harvester data management system. International Journal of Forest Engineering 13(1): 17-25.

Möller, J.J., Hannrup, B., Larsson, W., Barth, A., Arlinger, J. 2009. A system for calculation and geographical visualization of cutting quantities of forest fuels based on harvester data. Report 677. Skogforsk.

Möller, J.J., Arlinger, J. Barth, A. Bhuiyan, N., Hannrup, B. 2011. A system for calculation and feedback of harvester based information to forestry planning system. Report 756, Skogforsk.

Möller, J. J., Bhuiyan, N., Hannrup, B. 2015. Development and test of decision support tool for automated monitoring of thinning. Report 862. Skogforsk.

Siljebo, W., Möller, J. J., Hannrup, B., Bhuiyan, N. 2017. hprCM – module for using harvester data to calculate tree properties and forest fuel quantities. Report 944. Skogforsk

Skogforsk. StanForD2010 – Modern communication with forest machines. https://www.skogforsk.se/english/projects/stanford/

Söderberg, J., Willén, E., Bohlin, J. 2017. Identification of thinning needs using remote sensing. Report 922. Skogforsk

Söderberg, J., Willén, E., Möller, J.J., Arlinger, J., Bhuiyan, N. 2017. Evaluation of yield forecasts produced by forest laser estimations and harvester data – results from three case studies. Report 937. Skogforsk