

Article



# Automated Method for Delineating Harvested Stands Based on Harvester Location Data

# Timo Melkas \*🕩, Kirsi Riekki and Juha-Antti Sorsa

Metsäteho Ltd., Vernissakatu 1, FI-01300 Vantaa, Finland; kirsi.riekki@metsateho.fi (K.R.); juha-antti.sorsa@metsateho.fi (J.-A.S.)

\* Correspondence: Timo.Melkas@metsateho.fi

Received: 3 July 2020; Accepted: 20 August 2020; Published: 25 August 2020



**Abstract:** The data produced by cut-to-length harvesters provide new large-scale data source for event-based update of national forest stand inventory by Finnish Forest Centre. This study aimed to automate geoprocessing, which generates delineations of operated areas from harvester location data. Automated algorithms were developed and tested with a dataset of 455 harvested objects, recorded during harvestings. In automated stand delineation, the location points are clustered, the stand points are identified and external strip roads are separated. Then, stand polygons are produced. To validate the results, automatic delineations were compared to 57 observed delineations from field measurements and aerial images. A detailed comparison method was developed to study the correspondence. Stand polygonization parameter was adjusted and areal correspondence with 1% error on average was obtained for stands over 0.75 ha. Good stand shape agreement was observed. Overall, the automated method worked well, and the operative stands towards forest inventory stands, a balancing algorithm is introduced to create a solid, unique stand boundary between overlapping stands. This algorithm is beneficial for upkeep of stand networks. In addition, the Global Navigation Satellite System (GNSS) accuracy of the harvesters was examined and estimated numerically.

**Keywords:** automated geoprocessing; forest stand; forest inventory; harvester data; GNSS positioning; clustering

## 1. Introduction

Forest and wood procurement planning requires detailed and up-to-date information on forest stands' structure and forest management operations [1,2]. Up-to-date forest resource information provides a basis for decision making and the utilization of forest resources, as well as ensures that harvesting and forest management operations are carried out right on time. This is beneficial for both the forest owner in getting a better return on their forest assets and the forest industry in getting the timber assortment that they want at the right time in the right place [3].

Forest management planning in Central and Northern Europe is nowadays based mainly on stand-wise forest inventory [4]. The stands are formed from compartments which can be defined in many ways depending on the point of view and purpose for which the data are used [5]. The patterns of the compartments may originate from forest inventory units, economy-based units, operative units or monitoring units with fixed stand boundaries.

The aim of stand delineation is to form an optimal forest inventory and treatment unit, which is delineated in terms of forest type, composition of tree species, growing stock properties (age, mean diameter and height, basal area and volume) and treatment methods and their intensity (final cuttings, thinnings or silvicultural treatments). The stand delineations also depend on topographical properties, biodiversity-related features and goals of the owners. From the economical point of view, it is very

common to combine compartments which are close to each other into larger treatment units or cutting areas. The stand delineation is a combination or compromise of the different compartment patterns which delimit a relatively homogeneous part of the forest with a uniform need for forest management operations [5].

In forest plans, variables of the stand are expressed as quantity per hectare. To calculate the total growing stock values to the total area of the stand, both hectare-wise values and the area of the stand are needed. If the location of the stand boundary cannot be precisely defined, then the area of the stand can also not be reliably defined. Consequently, volume and growth estimates of the stands can also be biased. This is the most important reason the border of the stand should be delineated accurately.

In the traditional stand-wise forest inventory, the compartments were delineated manually. The stands were digitized from aerial images by the forest planner and inspected at field. As this is time consuming, laborious and partially subjective, several automated or semi-automated delineation methods has been proposed [6,7].

Nowadays, the data of forest inventory are collected in Northern Europe by airborne laser scanning (ALS) combined with aerial photographing and manual measurements of field sample plots [8–12]. An area-based approach is used to develop prediction models for forest inventory variables between field measurements and the features of the ALS data, using 16 m by 16 m grid. The growing stock properties of the grids at entire ALS inventory area are then obtained by applying the prediction model to scanned data. The grids are automatically segmented into homogeneous compartments, called micro-stands which are then further merged to form forest inventory stand delineations [13].

The Finnish Forest Center (FFC) has been utilizing this kind of remote sensing-based inventory method from 2010 onwards, and the whole country was covered by 2019 [13]. During this inventory cycle, the traditional stand-wise forest inventory has been used as well, but as a complementary means to ensure the quality of inventory results especially in young forests. Field inspections by FFC are carried out for approximately one-tenth of the forest inventory stands [13]. Additionally, the ALS-based stand boundaries are manually checked by planner when compiling forest plans for individual forest estates. The forest inventory data are distributed by FFC as open data if the forest estate owner has not forbidden it.

Stand delineations are in constant change. Harvesting operations, forest regeneration and silvicultural treatments of young forests all cause changes to the forests and the stand delineations representing them. Therefore, it does not make sense to talk about permanent delineations of stands, but instead about delineations of stands which are changing continuously and which can be updated after different treatments or cuttings [14]. In practice, the update of continuous stand delineation is done at certain points of time. However, individual locations in the forest always belong to some stand in changing stand delineations.

The continuous updating of forest information demands faster and more efficient collecting of data about the operation events in the forests. For this purpose, the stand boundaries should already be located during forestry operations and silvicultural treatments, for instance, by utilizing Global Navigation Satellite System (GNSS). More accurate GNSS and sophisticated geographical information systems (GIS) enable faster and multifaceted collecting, processing, analyzing and visualizing of data, even in automated manner. The forest machines that work right at the harvesting sites, provide an interesting means to record GNSS data of where operations take place.

Cut-to-length harvesters record automatically the location of the machine for each cut tree in stem (STM) or harvester production (HPR) files. The location of the harvester is determined based on the global satellite navigation system (GNSS). The position accuracy of single-tree harvester location under the forest canopy varies from few to several meters when using GNSS receivers typically integrated into harvesters [15].

Some modern harvesters are also capable of recording the position of the harvester head during the cut. The exact position of the cut tree could in such case be derived from the harvester bearing and boom parameters [16].

The stem diameters and log lengths are also measured and recorded by the harvester for the part of the stem which goes through the harvester head [17]. Additionally, the driver records tree species, timber assortments and the quality of the logs. All these data are stored into Standard for Forest machine Data and Communication (StanForD) files. It is worth noting that the harvester measurements represent the removal from the growing stock of the operated stand.

The stand formation from harvester locations has been conducted previously in Sweden. In [18], a geoprocessing method based on overlay of data points upon a grid in GIS is presented. Such grid-based approach is fast to perform but has a deficit in the shape of the formed stands, namely the rectangular characteristics of the stand boundaries originating from the usage of grid.

Computational geometry relates closely to automated geoprocessing of stand delineations from point-wise harvester location data. Harvested objects often consist of several individual stands and include also other trees than those cut purely from within the marked stands, thus being not directly comparable to typical forest inventory stands. The location point datasets need to be converted into stands without additional pre-information being available. In such setup, clustering approach benefits the formulation of the individual stands based only on the point coordinates.

An interesting clustering algorithm called Density-Based Spatial Clustering of Applications with Noise (DBSCAN) is presented in [19]. This algorithm is developed to efficiently recognize clusters of arbitrary shape from large datasets using density of points. Algorithm takes only one input parameter which is the distance between points within clusters. The description of DBSCAN provides as such applicable notions to clustering of harvester location data. However, some differences exist, too, relating to the inherent spatial distribution of the point data. In [19], the DBSCAN algorithm is applied on clusters where the points are rather uniformly and isotropically distributed within the cluster. In contrast, the harvester data are not isotropic since the location is obtained by GNSS receiver on top of the harvester cabin which typically moves only at limited routes at the marked stand. The removal (point) density also varies as the growing stock can originate from natural renewing processes or can be subject to thinning operations or, e.g., damages during growth decades. In this paper, an approach, in some degree similar to DBSCAN, is presented to cluster the harvester location point datasets into individual groups of stand points.

To delineate the geometries for the stand clusters, the point data are converted into a polygon form. To obtain a representative polygon with accurate area, an approach beyond common convex hull-type solution is needed to reach the adequate level of details at the delineation. The work in [20] provides a method for generating non-convex, characteristic shapes for sets of points by  $\chi$  (chi) algorithm. In addition, Duckham et al. [20] presented discussion of related work of the topic. The  $\chi$  algorithm is closely related to Delaunay triangulation of points. It produces topologically regular polygons, which is an essential requirement in the case of applications. The importance of adjusting the delineation to desired level of details is brought up in [20], especially at boundaries of the forming polygon. This relates directly to the current problem. Harvester data also demands applying further restrictions in polygonization due to the spatial anisotropy.

The objective of this study was to develop an automated geoprocessing method to delineate the operative harvested stands based on harvester location data. The utility of the algorithms was tested, and the effects of harvesting type and GNSS positioning accuracy of harvester were studied from the results. The accuracy and representativeness of the harvested stands was demonstrated by examining areal correspondence with reference delineations in detail, including locations of the stand boundaries. To convert the harvested stands towards forest inventory stands, an additional method is introduced to create balanced boundaries for overlapping stands when compiled into a larger network of stands. The automated stand delineation is planned to be used in event-based update of Finnish forest stand inventory in FFC as a complementary source of information within the main update cycle by remote sensing methods.

The paper proceeds by introducing the data and the detailed stand delineation algorithms in Section 2. The results of the accuracy of the method with respect to the reference data are presented in Section 3. The discussion is presented in Section 4. Section 5 concludes the applicability of the method.

#### 2. Materials and Methods

## 2.1. Study Area and Harvester Data

The study area of this work was the Uusimaa region which is located in Southern Finland (Figure 1). The landscape of the region is a mosaic of agricultural, forest and urban land, with the altitude ranging from 0 to 174 m. Finnish Forest Center has an operational inventory area at the study area. The Uusimaa study area covers 909,700 ha, of which 521,000 ha is forest land. The proportion of mineral soils is 88.5% and peatlands 11.5%. The dominant tree species of the forest land are Norway spruce (*Picea abies*) with 40% of the total volume, Scots pine (*Pinus sylvestris*) with 32% of the total volume and deciduous trees (*Betula* sp., *Populus* sp., etc.) with 28% of the total volume. The area of interest contains 14% mature (regeneration-ready) stands, 43% advanced thinning stands, 24% young thinning stands, 18% young and advanced seedling stands and 1% seed-tree and shelter tree stands. Mean growing stock volume of the forest land is 159 m<sup>3</sup>/ha, varying from 13 to 279 m<sup>3</sup>/ha depending on development class. In the Uusimaa region, the forest have slightly more structural variation than other regions in Finland [9,21,22].



**Figure 1.** The study area Uusimaa region in Finland. The laser scanning area is shown by grey color. Green points mark the harvester locations. The WGS84 (EPSG:4326) coordinates are shown at the boundary of the map. Base map © National Land Survey of Finland.

The harvester data were collected from STM files of six harvesters, from three different harvester manufacturers: Ponsse (Vieremä, Finland, four harvesters in this study), Komatsu Forest (Umeå, Sweden, one harvester) and John Deere (Moline, IL, USA, one harvester). Machines of different age were included. The data were collected during August 2015–September 2016 and contained 455 operative harvested objects [17] and 634,656 stems. Harvested objects represent units in which the operations are conducted. The STM files contain identifiers, tree-wise harvester locations and stem measurements.

To describe the marked stands at study area, average variables of the total removal were calculated from harvester data. The calculations were based on diameter profiles of the stems which were measured by harvester in 10-cm intervals for the commercial parts of the trees. For each stem, the diameters were transformed into ideal stem curves by Metsäteho's in-house tool, which utilizes least-squares fitting and is based on model by Laasasenaho [23]. The total length of the tree was directly obtained from the fitting. The diameter at breast height (DBH) from ground and volume of the entire stem, including stump and top of the tree, were calculated from the fitted stem curve. The areas of the stands were taken from the results of automated stand delineation to convert the removal variables into hectare-wise values. The average total removal variables characterizing the harvested stands at Uusimaa region are given in Table 1 for stands over 0.5 ha.

**Table 1.** The average total removal variables for harvested stands of this study, including only the stands over 0.5 ha. Other harvesting types that are not mentioned in table have been excluded. Number of stands in each harvesting type is denoted as  $N_{stand}$ , and the number of harvested trees  $N_{tree}$  is given per hectare. Abbreviations  $D_{gM}$  and  $H_{gM}$  stand for the mean basal area-weighted breast diameter and tree height, respectively. Variables *G* and *V* note the basal area and total volume. Abbreviation SD notes standard deviation.

Harvesting Type (N <sub>stand</sub> )	Proportion of Tree Species: Pine/Spruce/Birch/Other, %		A, ha	N <sub>tree</sub> per ha	$D_{gM}$ , cm	$H_{gM}$ , m	<i>G</i> , m²/ha	V, m <sup>3</sup> /ha
		min	0.5	380	11.8	11.6	3.6	26.1
First thinning (33)	28/26/41/5	max	7.9	1020	27.1	20.0	13.8	123.8
First unitility (33)	28/28/41/5	mean	1.9	680	15.3	14.9	8.9	67.4
		SD	1.6	180	3.0	1.8	2.8	25.6
		min	0.5	160	13.7	13.2	5.1	49.0
Later thinning (108)	25/52/19/4	max	11.8	950	33.9	23.5	26.6	286.5
Eater timining (100)	23/32/19/4	mean	2.8	530	20.5	18.2	11.2	103.5
		SD	2.3	200	4.2	2.4	4.1	46.3
		min	0.5	220	14.9	15.8	8.2	65.8
Clear cutting (216)	11/7//12/3	max	12.4	1130	46.8	29.6	41.9	562.8
Cical cutting (210)	11/74/12/3	mean	2.4	600	28.6	22.7	25.4	281.5
		SD	1.9	180	4.4	2.4	5.4	75.0
		min	0.5	90	16.7	14.6	2.9	30.1
Seed tree or shelter tree	22/62/11/4	max	7.2	1200	39.3	27.7	34.5	423.7
cutting (43)	25/05/11/4	mean	2.0	460	28.7	22.4	19.5	214.6
		SD	1.5	190	4.9	3.0	6.7	87.0
		min	0.8	40	23.0	20.3	2.6	29.0
Hold over suffing $\frac{1}{10}$	34/27/34/5	max	7.9	540	37.1	28.0	19.9	198.3
(19)	34/2//34/3	mean	2.3	190	31.0	23.8	9.0	96.8
		SD	1.8	150	4.3	2.1	5.0	51.5

<sup>1</sup> The delineations of the hold-over cutting stands were separately produced for this calculation, to cover the entire area from which the trees were harvested. The area was needed for hectare-wise stand variables.

## Preprocessing of Harvester Data

The purpose of the preprocessing is to extract the relevant information for stand delineation algorithms from harvester data files (STM/HPR). The contents of interest are harvested object information, comprising the identifiers, starting time of harvest, harvesting type and the stem-wise coordinates of harvester while cutting the tree. This information is read from the harvester files and inserted into a relation database. However, the harvesting type was not recorded by harvester into STM files used in this work; instead, it was obtained from forest companies and manually added into the data. However, HPR files of StanForD 2010 standard contain harvesting type, although in some cases a harvesting type group information is given in the HPR file instead of detailed harvesting type (e.g., for regeneration fellings). The stems are identified in the data by a unique, continuous StemNumber, which is recorded by the harvester.

The division of harvester data into harvested objects was fully automated. This means that a unique bookkeeping label called "ObjectID" was produced based on the identifiers, starting time and harvesting type of the data. The label was given for each harvested object and their stems. It is worth

noting that the automated preprocessing results separate objects if, e.g., only starting times of two objects differ.

The harvesters of the study were divided into two GNSS accuracy groups: "more accurate" and "less accurate" GNSS receivers. The classification was done based on machine properties that were manually interpreted from StanForD file extractions. The machine properties appear in miscellaneous optional fields in StanForD files, which prevented automated extraction of machine data. Some additional information was also inquired from the machine manufacturers, including the age, model and software of the harvester. The availability of machine data allowed such accuracy division and enabled studying how receiver accuracy affects the stand delineation. Such set of information is not currently available outside this study, and therefore a numerical estimation method of GNSS accuracy was developed.

The GNSS accuracy estimation method averages the distance of consequent trees in cutting order based on StemNumber-variable in harvester data. Long distances were excluded, as they can be even hundreds of meters, and that would skew the result. The details of the method are presented in Appendix A.

#### 2.2. Reference Stand Data

The reference stand delineations were obtained from two difference sources. Firstly, stand boundaries were recorded as field observations at harvesting sites with a high-accuracy GNSS receiver, resulting  $N_{GNSS}$  = 27 stands. The GNSS receiver was Trimble GeoXH6000 (Trimble, Sunnyvale, CA, USA), and VRS H-Star correction was used. Most of the stands were recorded during the harvester data collection period at autumn 2015 when the boundaries of the harvested area have been clearly visible. A smaller, complementary dataset was obtained at summer 2019 from harvested sites which allowed recording of the stand boundary despite the time elapsed from the harvesting. At this time, the receiver was Trimble R2 GNSS (Trimble, Sunnyvale, CA, USA) with VRS H-Star correction.

The other source of reference stands was the aerial images by National Land Survey of Finland which are available as open data. Reference stands ( $N_{AI}$  = 42) were digitized from aerial images in GIS program. The reference was recorded only, if the harvesting operation was visible in the aerial image. From this reason, the aerial image references were mainly clear cuttings or seed tree cuttings. The aerial images of areas used in this work were obtained at springs 2017 and 2018. Based on information provided by National Land Survey of Finland, the spatial accuracy of the aerial images is 0.5–2 m and pixel size of the aerial images is 0.5 m [24].

In total, 57 separate stands were included into references. Both references were obtained for 12 stands. The harvesting types of the stands were distributed as follows: 48 clear cuttings, 4 thinnings and 5 seed tree cuttings.

Delineating reference stands for operative harvested stands differs from recording a typical forest inventory stand. The forest inventory stands form a topologically covering network of forested area where each stand is neatly bounded by other stands, roads, fields, water areas, non-forested areas or other terrain objects. In contrast, the operative stands appear merely where the operation has occurred, independently of the surrounding properties of forest or terrain. Therefore, attention had to be paid when recording the operative stand references to not record the reference by terrain feature geometries around the site. The harvester location data were used to indicate the overall region where the operative stand reference lied.

At GNSS field recording, the reference boundary was located halfway between the stumps of cut trees and the remaining trees standing next to the harvested area. The same principle was guiding also the recording of aerial image references, but the resolution of the image did not quite allow such accuracy. Expert opinion was therefore used in estimating the best possible delineation for the reference stand, considering the shadows and the direction of shooting of the aerial image. However, some degree of subjectivity possibly remains in the delineations. Determining the exact boundary of harvested area was difficult if, for example, young tree bushes remained after harvesting at stand

boundaries near fields, ditches and roads. In some cases, an open area was located next to the clear cut harvester data area, and the reference had to exclude since stand boundary between two open areas could not be observed from the aerial image.

#### 2.3. The Automated Stand Delineation Method

The automated method was developed based on examination of harvester data and its properties. The method was implemented as three individual, consecutive geoprocessing algorithms, which are called Algorithms A1–A3, respectively.

Before performing the automated algorithms in GIS, the harvester data were put through the preprocessing described in Section 2.1. The automated stand delineation method relies on the two unique identifiers, ObjectID and StemNumber, which identify both the objects, and the stems within the objects. Then, the input data are transferred into geographical information system (GIS) program. The harvesters record their coordinates in WGS84 (EPSG:4326) coordinate system which are then reprojected into Finnish metric coordinate system ETRS89/ETRS-TM35FIN (EPSG:3067).

Algorithm A1 classifies the data points into clusters of stand points by the locations. It excludes strip road points which are leading to the stands rather than belonging to the stands. Algorithm A2 then converts the clusters of stand points into stand delineations which represent the operative stands realized at the harvesting site. Algorithm A3 balances the boundaries of overlapping operative stands, and thereby converts the operative stands towards forest inventory stands. After Algorithm A3, the information of the removal trees from the stand is compiled, in addition to the stand geometry. Both are needed in further analyses when calculating the stand variables, e.g., cutting removal per hectare. Geoprocessing workflow of algorithms is shown in Figure 2.



Figure 2. The workflow of geoprocessing in automated stand delineation.

The algorithms contain various numerical parameters that are used in geoprocessing. The values of these parameters correspond to average values appearing in practical harvestings at field and are selected based on best expert knowledge. However, currently, the values are inferred from the limited number of harvester data that were available for this work.

Developing and testing of the algorithms was done in QGIS 2.14 long-term release (Open Source Geospatial Foundation), writing scripts by Python language. Using Python enabled the use of spatial processing libraries SAGA (SAGA User Group Association, Germany) and GRASS GIS 7 (Open Source

Geospatial Foundation) that are available via QGIS. The algorithms were also implemented in PostGIS (Open Source Geospatial Foundation) to provide information of the computation times at cloud environment. Other GIS programming environments may provide additional or different geoprocessing tools, and their suitability in implementation of the algorithms has to be considered case by case.

Algorithms A1–A3 are described in detail in the following subsections.

# 2.3.1. Algorithm A1: Separation of External Strip Roads

Algorithm A1 (Appendix B) handles tree-wise point data of one harvested object at a time. Each point is the location of the harvester while felling the tree. The algorithm separates the trees that belong to stands from trees that belong to external strip roads leading to stands (see Figure 3). The separation is done based on the spatially averaged locations of the points. The special benefit of spatial averaging is that it is independent of the felling order of the trees. This is needed since the harvesters drive back and forth at the marked stand when changing work shift, refueling, etc. They often fell some trees during such driving, from locations that situate somewhere at the strip road network, if, e.g., a damaged tree is observed at thinning. Therefore, the felling order of the trees is "discontinuous" in that sense, thus some consequential trees of the data have not been felled proximately from each other at the marked stand. In order-based averaging, all different driving occasions would be separately handled, but, in spatial averaging, all recorded locations around the point of study are included and only one, unique averaging result follows for all locations. The positioning accuracy of the harvester's GNSS receiver is used to adjust parameter values in Algorithm A1.



**Figure 3.** Example of results of Algorithm A1 at clear cutting object. The trees of the object have been separated into two geographically separate stands and external strip road trees. The ETRS-TM35FIN (EPSG:3067) coordinate system is used in the figure. Base map © National Land Survey of Finland.

The tree-wise classification of the points into stand/non-stand trees is done with help of finding if there exist adjacent strip roads for the point under study. This is done by determining directions within the data which resemble the shape of the strip road network—the actual bearings of harvester movements at the strip road are irrelevant here. The locations within stand do have adjacent strip

roads, but external strip roads leading to stands do not have adjacent strip roads. As the external strip roads can be either shorter spur trails or longer external strip roads, the algorithm separates them based on maximum allowed length of spur trails. The spur trail trees are then added into the stand trees.

The algorithm returns stand-wise group of tree points, i.e., clusters, as a result. The trees belonging to external strip roads are labeled differently for further use, but they do not belong to any stand. One object can turn into several stands, if it contains geographically separated harvested stands, as in Figure 3.

The details of the Algorithm A1 are presented in Appendix B.

Algorithm A1 has some similar characteristics as the DBSCAN clustering algorithm [19]. The density of the points is in major role in recognizing the stand points in both DBSCAN and Algorithm A1. In DBSCAN, the points have uniform spatial distribution and the neighboring points are searched by circular area, which has distance parameter as radius.

The harvester data recorded by GNSS receiver at top of the cabin have, however, pronounced property in the spatial distribution of the points: they are restricted mainly into the strip road network at the stand, especially at thinnings. Small perpendicular movements of the harvester with respect to the strip road network routes do appear during the harvesting work, but they are most often limited to length scale of the harvester itself, a few meters. Recommended strip road interval at thinnings in Finland is 20 m based on typical lengths of harvester cranes [25]. In this study, the maximum lengths of harvester cranes were 10–11.7 m, and on average somewhat less.

Due to this inherent spatial restriction of harvester data, the density-based clustering distance is relevant to determine perpendicularly from the strip road network positions, i.e., the neighborhood is rather adjacency to the point of study. The direction of the strip road's "tangent" is needed for directing the search of neighbors. Observing the curves of the strip roads at stand, the angle of perpendicularity was enlarged from the strict 90°, thus resulting the sector-shaped intermediate geometry in finding the neighboring points (see also Figure A2 in Appendix B). The amount of found neighbors is one in Algorithm A1, and a constant value is suggested for use in DBSCAN [19].

In combining the stand points into a cluster, rather similar approach is used in DBSCAN and Algorithm A1. In DBSCAN, the cluster is collected starting from a seed points by putting together all such points that are "density-reachable" from each other, as described in [19]. In Algorithm A1, the stand trees are collected in very similar manner, starting from the first cut tree points and combining more points into same cluster if the buffered point groups intersect—this is another formulation of "density-reachability" using two distance parameters due to spatial restriction of harvester data.

In stand recognition of Algorithm A1, the parameters were selected to include tree point into stand if possible. This is in concordance with the work of Ester et al. [19], which suggests that the density of the "thinnest" cluster that is still accounted as a cluster instead of noise would be an appropriate density for the clustering algorithm. In Algorithm A1, this relates to the value of perpendicular distance that is used to find the neighboring points at adjacent strip roads. The numeric value selected reflects the combination of maximum length of harvester boom, GNSS positioning error and the preferable strip road interval in perpendicular direction—these all are further dependent on, e.g., harvesting type.

Similarly, the limit length of spur trails was selected to rather include the spur trail into stand than to leave it as an external strip road. When the resulting stand-wise tree groups are merged together, the algorithm rather combines two close but separate stands into one. For this reason, it is possible that Algorithm A2 later results multipolygon-type stands which are then converted to single polygon type (see Section 2.3.2).

The GNSS accuracy was taken explicitly into account in parameter values of Algorithm A1, which adjust the strength of spatial averaging. In practice, the range of spatial averaging was enlarged for objects harvested with less accurate GNSS positioning at harvester. That way, the algorithm was able to recognize a single strip road, longer than the maximum spur trail, from between two separate stands.

As Algorithm A1 is strongly based on the locations of the cut trees, two types of cases were found from the dataset in which the automatic separation of the stands and external strip roads is difficult.

Firstly, if external strip road curves along the edge of stand, it will be merged into stand (Figure 4a). Secondly, if stand is elongated and consists only of one strip road, the algorithm will not recognize that as a stand (Figure 4b; see also Figure 3 strip road in the middle). To tackle the later problem, obvious stand cases have to be recognized otherwise after Algorithm A1. Appendix C introduces a short estimation method based on areal density of cut tree points at the strip roads, where an effective area for strip road parts is temporarily created. In that way, the dense strip roads are reverted to stand-wise tree groups into stand formation along the previously obtained stand point clusters. For example, most of the strip road tree points in Figure 4b turn into stand points.



**Figure 4.** Averaged tree point data of harvested objects. The ETRS-TM35FIN (EPSG:3067) coordinate system is used in the figure. Aerial images © National Land Survey of Finland. (**a**) The external strip road is recognized as part of stand. GNSS positioning of harvester was more accurate. (**b**) The parts of stands which have only one strip road are recognized as external strip roads at final felling objects. GNSS positioning of harvester was less accurate based on available information.

## 2.3.2. Algorithm A2: Stand Formation

The starting points of Algorithm A2 are the stand-wise tree groups of one object from Algorithm A1. Each stand is handled separately. At first, this tree-wise point data of raw locations are checked for whether the geoprocessing can progress further. The minimum requirement is three separate points adequately near to each other. If found appropriate, Delaunay triangulation of the data points takes place. Triangulation is somewhat demanding in computation resources, and therefore the duplicate locations are excluded. The base of the stand is formed by selecting certain triangles depending on their geometric properties. The selection corresponds to the region that is inside the strip road network of the stand, as the points are the locations of the harvester (see Figure 5a). The stand area is then obtained by buffering the area outwards with constant value  $B_2$ , to represent the area from which the trees were harvested from the location of the harvester during felling (see Figure 5b). The buffering width roughly corresponds to a net distance of both average boom length and GNSS error of the harvester location data. The suitable buffer width was determined by adjusting the stand areas with respect to the reference stands. This process is detailed in Section 3.3.3 and resulted in the buffer width  $B_2 = 5.0$  m for more accurate and  $B_2 = 3.2$  m for less accurate GNSS positioning.

The outcome of Algorithm A2 is a delineated polygon of operative harvested stand which indicates the realized area and location of the harvesting. Note that the harvested stand is different from the corresponding forest inventory stand, as explained in the Introduction. Algorithm A2 details are presented in Appendix D. The parameters of the algorithm were adjusted so that the selected area follows the edges of the stands closely, but then some empty slots tend to appear within the stand area more frequently (see Figure 6a). Buffering reduces the size of the slots. The slots smaller than  $A_{max} = 1000 \text{ m}^2$  can be filled afterwards, as they are judged small enough to belong actually into the stand (Figure 6b). Interestingly, the appearance of empty slots allows recognition of non-harvested areas at the scale depending on the parameters of Algorithms A1 and A2. Some of them are retention tree groups to maintain the biodiversity at final fellings, but some are merely locations where growing stock is not mature for harvesting. Whether a certain empty slot is a retention tree group cannot be directly deduced from harvester data only.



**Figure 5.** Examples of Algorithm A2 at clear cut objects where harvester was known to have more accurate GNSS. The ETRS-TM35FIN (EPSG:3067) coordinate system is used in the figure. Aerial images © National Land Survey of Finland. (a) The selected triangles from Delaunay triangulation form the preliminary stand area. (b) The triangles are dissolved and buffered to obtain the stand polygon.



**Figure 6.** Empty slots appear into thinning stand during polygon formation. The ETRS-TM35FIN (EPSG:3067) coordinate system is used in the figure. Base map and aerial image © National Land Survey of Finland. (a) The selected triangles for stand delineation. (b) After polygon has been produced, the empty slots are smaller as the buffering of the stand polygon reduces their area. The empty slots which have area under  $A_{max}$  are indicated in the figure.

Despite the parameter adjustments done during the development of the algorithm, isthmuses (necks of land) sometimes appeared between different parts of the stands (see Figure 7). That is caused by the restricted spatial distribution of the harvester location data points into the strip road network. The dimensions of those Delaunay triangles that cause the isthmuses are such that they have to be

accepted into stand area elsewhere in the data. Thus, there appears to be a delicate balance in adjusting the parameters in Algorithm A2 formulated as described.



**Figure 7.** An example of isthmus which has appeared during triangle selection in Algorithm A2 at clear cut object. Harvester was known to have a less accurate GNSS receiver. The ETRS-TM35FIN (EPSG:3067) coordinate system is used in the figure. (a) Three triangles have been selected into stand since the tree points at opposite edges of the concave non-harvested area are within the allowed maximum distance of points. (b) Aerial image confirms that the area of isthmus has not been clear cut. Aerial image © National Land Survey of Finland.

The comparison of Algorithm A2 to  $\chi$  algorithm generating characteristic shapes [20] shows interesting findings. The  $\chi$  algorithm examines the length of one edge of triangle at a time to be selected into the final polygon. It is stated to provide rather stable results at wide range of input length parameter. At the stable region, the  $\chi$  algorithm tolerates changes in point density and homogeneity of point distribution. Instead, outliers are not handled as well, but they are advised to take care of by preprocessing [20]. These observations also apply for Algorithm A2 by visual inspection of the dataset—small changes in parameter values do not change significantly the outcoming stand delineation. The outliers are a problem for Algorithm A2 as they expand the area of the polygons relatively too much. Attempts are made in this work to clean them, both in entire Algorithm A1 and in stage D.1. of Algorithm A2 (see Appendix D). One major difference is that  $\chi$  algorithm cannot create empty slots into the polygon geometry, in clear contrast to Algorithm A2 of this work. This capability is seen as a benefit for stand delineation.

#### 2.3.3. Algorithm A3: Intersecting Stand Polygons

Whereas Algorithms A1 and A2 produced the object-wise delineations for realized operative harvested stands, Algorithm A3 balances the overlapping stand delineations originating from different objects. It turns the stands towards forest inventory stands which need a unique stand boundary between them to form a topologically clean and covering network of stands, e.g., for determining the growing stock related properties of the forest inventory stands. The starting point of Algorithm A3 is the stand delineations from Algorithm A2.

In this study, the complete network of balanced operative harvested stands could not be achieved due to the small number of harvester data. An attempt was, however, made to balance those harvested stands that were available. To simplify the test, the stand geometries (results of Stage D.2. of Algorithm A2 in Appendix D) without dense strip roads were used. In that dataset, roughly 20% of the stands intersect with adjacent stand polygons.

The reasons for stand intersections are that the positioning inaccuracy may move the trees outwards at the stand boundaries, or the harvester may have been situated and/or felled some trees from the other marked stand. In Algorithm A3, the intersection zone of stands is split and the parts are merged into the stands. The new, balanced stand boundary is created into the middle of the intersection zone.

If two stands have several intersection zones, empty slots appear between the intersection zones. Small empty slots (area less than  $A_{max}$  introduced earlier, having value 1000 m<sup>2</sup>) are divided for the stands as well, as in Figure 8. Instead, slots over  $A_{max}$  are left intact—the limit size of the area was selected here to represent the desired outcome, and the value is to be adjusted for other purposes. A practical minimum size limit for splittable intersection and slot polygons was set as  $A_{min} = 10 \text{ m}^2$ . Smaller polygons are simply merged as such into either of the intersecting stand polygons. With this value of  $A_{min}$ , the area error caused into resulting updated stands due to not splitting small polygons is judged negligible. The details of Algorithm A3 are presented in Appendix E.



**Figure 8.** A new, solid stand boundary is created between the intersecting stands. The intersection zone and the small slot between the intersections are split and merged into stands. Note that the other small slot is smaller than  $A_{min}$  and therefore not split, but only merged into one of the stands (here to the blue stand at left). Harvester was known to have a less accurate GNSS receiver. The ETRS-TM35FIN (EPSG:3067) coordinate system is used in the figure.

Algorithm A3 facilitates more general stand network operations by embedding updated single stand delineations cleanly into existing stand network. When stand has several adjacent stands around it, the algorithm handles the intersections for each stand pair separately (see Figure 9). It is possible that an empty slot appears between three intersecting stands; however, it is possible to update the algorithm to fill these slots easily if necessary.

The algorithm has rather detailed stages in how the divided parts of intersection polygon are addressed to the originally intersecting stand polygons. The main solution is based on the straightforward geoprocessing logic, but sometimes a numeric approach based on certain intersection areas is needed (see Appendix E, Stage E.4. of Algorithm A3). In the current dataset, there were few cases in which this areal comparison was necessary. Despite performing the areal comparison, it is rarely possible that, with small enough (but over the limit size  $A_{min}$ ) divided parts of intersections, the divided part may end up combined with "wrong stand polygon".



**Figure 9.** Several adjacent stands are intersecting each other, and Algorithm A3 has provided the solid updated stand boundaries for all of them to the middle of the intersection zones. The stand pair-wise intersection zones are shown in different colors. All stands were harvested with machines with known more accurate GNSS receiver. The ETRS-TM35FIN (EPSG:3067) coordinate system is used in the figure. (a) A separate forest stand (area 0.75 ha) lies within the harvested stands. (b) Nine intersecting stand geometries are converted into a stand network. The intersections are handled for two stands at a time in the current algorithm. Therefore, the unfilled areas remain between some stands.

In optimal cases, the trees that lie within the intersection zone are well balanced with respect to the new stand boundary (Supplementary material, Figure S1a). When the tree-wise point data are considered, few points seem to be at the "wrong" side of the new split stand boundary, but they still belong into their original stand. In some cases, the tree densities of the intersecting stands differ at the intersection zone (Figure S1b). Taking tree density and tree dimensions into account when creating the new stand boundary may be beneficial in some cases.

There are certain restrictions for the polygon intersections to be splittable in current Algorithm A3. The problematic cases are identified and excluded from the algorithm by the constraining conditions. In those cases, the amount of vertex points, i.e., points between which the division line is created, differs from two (see Stage E.2. of Algorithm A3). A typical situation is an overlap where one polygon reaches over, outside the other polygon when intersecting (for examples, see Figure S2). Considering the inclusion of dense strip road areas into stands, such non-splittable intersections of the stand polygons would appear more often.

However, problematic intersections can be handled in upkeep of a more complex stand network by formulating logical embedding rules based on, e.g., harvesting types and starting times of harvesting. Presenting such case-specific rules in a comprehensive manner is left out from this work as upkeeping of complete stand networks requires good areal coverage of the harvester data.

In current work, the forthcoming research aspects induced to form as many forest inventory-style stands from the current harvester data as possible. Inspection of data gave insight of few simple combination rules for non-splittable overlap cases, after which the stand polygons could be split. These rules included merging of stand tree groups if only the starting time of harvesting was different at two overlapping objects. Such rules were deemed applicable for further purposes and generalizing them has to be considered case by case. Obtaining the tree-wise removal results for the balanced stands may be beneficial for further studies based on automated harvested stands (for a more detailed description, see Appendix F).

#### 2.4. Method for Validating the Stand Delineations

An automated pairwise comparison of stand boundaries was needed to estimate how well the harvested stands correspond with the reference stands. Previously [26–29], area overlap approach of determining the correspondence of the stand boundaries was used, and it was taken as a starting point in this work. However, more information of stand boundary locations was strived for. Harvested stands turned out to have sometimes rather complex shapes, and the method was therefore aimed to compare various polygon shapes.

To respond to the abovementioned needs, a detailed, comprehensive, automated comparison method of stand properties was developed. It covers the entire stand boundary and can be adjusted to take adequate amount of samples. It is rather independent of the shape of the stand, however requiring some correspondence in the shape of the compared stands for giving reasonable results. The method can be used also more generally to compare any stand delineations, regardless how the stand delineations are obtained. The pre-requisite for the method is that the corresponding stand and reference stand are addressed, e.g., in the attributes of the GIS data.

Firstly, the area and perimeter, i.e., length of the boundary of the individual stands, are obtained directly from GIS program. Relative area with respect to the reference stand is obtained. The method determines areal match of the compared stands by union and intersection of the stands. The areas of regions outside the intersection are determined for both stands. Thus far, the method is similar to earlier works [26–29]. The relative perimeter describes the correspondence of the boundary lines, although this variable may need to be used with some caution if reference delineations have different level of details in geometry (e.g., GNSS-receiver observations at short intervals).

The displacements of the harvested stands are studied by sampling the perpendicular distance between stand boundaries around the whole reference stand (see Figure 10). Determination of distances and systematic shifts is presented in Appendix G.



**Figure 10.** The principle of method to compare different stand delineations comprehensively and accurately. The entire boundary of the reference stand is sampled into lines, and the perpendicular distance to the boundary of the studied stand is determined. The distance distribution is obtained from all samples and provides information of the match of the stand boundaries. The ETRS-TM35FIN (EPSG:3067) coordinate system is used in the figure.

The obtained average and standard deviation of the perpendicular distances describe how well the position of the boundary of studied stand corresponds with the reference stand's boundary. Additionally, systematic shifts can be found with help of averaged perpendicular distances between stand boundaries at the main directions. The magnitude of subtracted values at four directions N–S, NE–SW, E–W and SE–NW will reveal if there are significant shifts towards the first-mentioned directions. An example of interpreting the shifts is that if N–S = -10.5 m, the stand has shifted significantly towards south with respect to the reference stand.

Sometimes the shift values can also relate to partial expansion or reduction of the stand. In such case, for example half of the stand boundary corresponds the reference boundary, but the other half is either inside or outside the reference stand.

Giving exact numeric limit value for shifts/partial size changes was difficult with current amount of data, but shift values of roughly 4–5 m and above correspond with a clear shift in that direction in visual inspection. The magnitude of the "limit" shift value is related to overall GNSS positioning accuracy.

## 3. Results

In this section, the results of putting all input data through geoprocessing Algorithms A1–A3 are presented. The numerical estimates of GNSS positioning accuracy are examined. The overall accuracy of stand delineations is studied by comparing the references mutually. Then, a subset of automated harvested stand delineations is compared to the reference stand delineations in detailed stand-wise manner.

## 3.1. Robustness of Automated Method and Overall Results of the Dataset

The automatic method was run on all input harvester data that were available for the study. The results were examined visually.

Algorithms A1 and A2: The main finding was that all the input harvested objects (455) were managed to put through from both Algorithms A1 and A2 in a technical sense. The stand delineations of operative harvested stands for these objects were obtained, when the natural limiting conditions related to input data requirements were fulfilled. One cannot perform this kind of stand delineation on harvested object which contains, e.g., 10 identical point locations as a whole or only two separate points in total since the triangulation cannot be created for less than three points.

When examining the results visually, it was found that harvesting type strongly affected the sensibility of the results. For cuttings of hold-over trees (22 objects), the resulting stand delineations were of varying quality and found not further useful. Instead, for other harvesting types (thinning, final felling or other similar), the obtained stand delineations were reasonable (433 harvested objects, 623,176 trees). For these objects, 581 successfully delineated operative stands resulted, without the dense strip road polygons. Thus, roughly four stands appeared from three harvested objects at the study area. The areas of the stands are summarized in Table 2.

Dataset	N		Area, ha			
Dataset	1	Min	Max	Mean	SD	
Operative stands	581	0.02	12.01	1.73	1.88	
Dense strip road polygons	206	0.01	1.00	0.13	0.11	
Oper. stand $\cup$ dense strip roads	653	0.01	12.01	1.58	1.89	
Objects without dense strip roads Objects with dense strip roads	399 413	0.03 0.02	19.47 19.56	2.52 2.49	2.46 2.48	

**Table 2.** The numbers (*N*) and areas of harvested tree stems for harvested stands, strip roads leading to stands and harvested objects. Note that the hold-over cuttings were excluded. The numerically estimated GNSS accuracies were used. Abbreviation: SD, standard deviation.

The reason the stand delineation of hold-over cutting is not reasonable lies in the density and spatial distribution of cut trees (the harvester data points). The hold-over cutting marked stands can have only few trees harvested sparsely around the marked stand, e.g., in cutting seedling pines, or sometimes the density of harvested stands can be roughly the same as in typical thinnings, e.g., in

shelter tree cuttings for spruce (see also Table 1). In the case of sparse spatial distribution, Algorithm A1 cannot detect uniform harvested area from the location points.

The amount of external strip roads was calculated for the abovementioned 433 objects. For 70% (304) of objects, external strip roads leading to stands were found by Algorithm A1. The number of trees cut from external strip roads was roughly 4.6% of the total number of trees. Further, 3.3% of trees belonged to dense strip roads. When the dense strip road tree points were polygonised in Algorithm A2, 206 stand polygons resulted (for details, see Table 2). When taking these polygons into account in total amount of stands produced, roughly every second harvested object contained a dense strip road polygon in addition to stand(s) from Algorithm A1. To complement the operative harvested stands with the dense strip road polygons, the latter were added into the stand geometries if the polygons of same harvested object intersected (results of Stage D.3. of Algorithm A2). The individual strip road polygons were also added. Table 2 provides more details of the areas of the merged stands.

Table 2 also summarizes the realized areas for objects with and without the dense strip road polygons. The areas were compiled from the stand results. The difference in areas of stands and objects is clear. This, together with the amount of resulting stands, reflects how the harvested objects consist of more than one stand.

Erroneous location recorded by GNSS-receiver prevented the reasonable stand delineation for three objects from total 433 objects. The GNSS-receiver had recorded same single coordinate value for all stems in two objects. In addition, for one object the GNSS receiver had recorded same coordinate value for significant amount of stems, which in practice did not allow proper stand delineation from the harvester location data, although a dense strip road polygon was produced. All these objects were harvested by machine which was known to have less accurate GNSS receiver.

Regarding Algorithm A2, in visual inspection of operative harvested stands, the problem of remaining isthmuses was observed (see Figure 7). The area of isthmus is erroneously interpreted into automated stand, implying that harvesting took place there, although indication of that is not seen in aerial images. The appearance of isthmuses also increases areas, decreases perimeters of the stands and generates empty slots to areas that are not even part of the marked stand under harvesting.

Algorithm A3: The dataset of this study was very incomplete in the sense of compiling areally covering stand network from the harvested stands. However, 130 splittable stand overlappings (22% of total amount of stands, not including dense strip road polygons) were found after delineating operative stands, so the local coverage of stands was adequate to try and study the performance of Algorithm A3 in balancing the stand boundaries. Note that one stand can be part of several overlappings.

From this dataset, seven non-splittable stand overlapping intersections could not be handled with Algorithm A3 due to the geometries of the intersecting stands (see Section 2.3.3). The dense strip road stands were left out from stand network in current study, as they often can have more complex geometries. The non-splittable intersection cases can be, however, solved during upkeep of a more complex stand network by adding more deduction rules regarding the harvesting types and starting times of the intersecting stands. Then, the geometry problems of intersections disappear. For example, final felling geometries can override other harvesting types.

Practical geoprocessing caused three additional problem intersections in running Algorithm A3 in QGIS 2.14. In two intersections, the geoprocessing did not work as expected, although the geometries were programmatically valid. In one intersection, geometry error prevented the correct stand delineation, although automated cleaning was used in Algorithm A3 (incorrect stand delineation did appear in the results). In total, when all the stands of this dataset were compiled into (low-coverage) stand network, 98.3% (571) of the stands were successfully put through Algorithm A3, based purely on stand polygon geometries.

The computing time of automated stand delineation is typically few seconds to several minutes at PostGIS cloud environment, and few seconds to hours in desktop QGIS. The amount of trees in object affects strongly the computing time, especially at Algorithm A2 where Delaunay triangulation has computational efficiency  $O(n \log n)$  with n data points. In this dataset, the largest object contained

approximately 15,000 harvested trees, corresponding to total harvested area of roughly 19.6 ha (see also Table 2).

#### 3.2. Results of Estimated GNSS Accuracy

In this study, the GNSS positioning accuracy of harvesters was examined at harvested object level both visually and analytically. A two-fold classification was used: more accurate or less accurate positioning. In this dataset, some information of the harvesters was available and the GNSS accuracy was managed to estimate from those. The harvesters of the study are referred to as Harvester 1–6 and each harvester was given a label "more accurate" or "less accurate" by the known properties. This label was further given for the objects harvested by that machine. Typical visually observed scales for spread of location points are roughly under 4 m for more accurate GNSS receivers in this dataset. For less accurate receivers the scale of spread is often roughly 4–6 m, and can even exceed 10 m locally.

The two-valued GNSS accuracy was also numerically estimated from the point data for the harvested objects by the method presented in Appendix A. Table 3 shows how the numerically estimated accuracy corresponded with the accuracy obtained from known machine properties. The amounts of thinning objects, including here first and later thinnings, are also given in Table 3. The hold-over cuttings were not included in these results due to observations presented in Section 3.1.

**Table 3.** The correspondence of numerically estimated GNSS accuracy with GNSS accuracy obtained from known properties of harvesters. The numbers indicate how many objects belong into each category. Thinnings include in this table both first and later thinnings. The hold-over cutting objects have been excluded. Additionally, two objects were left out as all the points had same one GNSS location and the numerical estimate could therefore not be determined.

Harvester ID	Accuracy Label from Known Properties	Numerically Estimated More Accurate GNSS, All Harvesting Types (%)	Numerically Estimated More Accurate GNSS, Thinnings	Numerically Estimated Less Accurate GNSS, All Harvesting Types (%)	Numerically Estimated Less Accurate GNSS, Thinnings	Total Number of Objects
1	more accurate	53 (96)	10	2 (4)	2	55
2	more accurate	105 (92)	42	9 (8)	3	114
3	more accurate	103 (98)	48	2 (2)	0	105
4	less accurate	13 (34)	0	25 (66)	0	38
5	less accurate	19 (45)	12	23 (55)	14	42
6	less accurate	47 (61)	8	30 (39)	16	77
sum		340	120	91	35	431

As Table 3 indicates, the objects where harvesters were known to have a rather accurate receiver, turned out mostly as "more accurate" in numerical estimation of GNSS accuracy. Only 4.7% of objects by Harvesters 1–3 were estimated as "less accurate". Instead, from the objects where harvesters had a somewhat less accurate receiver based on known properties, 50% turned out to be classified into category "more accurate" GNSS receiver. Visual inspection of the data supported these results, and indeed showed that notable fraction of objects of Harvesters 4–6 had larger variation in the positioning accuracy than objects of Harvesters 1–3 which had less variation and rather accurate positioning. Especially the variation within single objects of Harvesters 4–6 took place at topologically varying landscapes: typically on top of the hills or otherwise open landscapes the accuracy was seemingly similar to Harvester 1–3, but if the harvested object reached below steep hills or cliffs, the local accuracy was clearly lower (for example, see Figure A1 in Appendix A).

Regarding harvesting type, the amounts of thinning objects in Table 3 imply that, if the harvester is known to have less accurate GNSS positioning, the growing stock remaining at the stand and its canopy cover may lower the numerically estimated accuracy. Overall, the results indicate that actually the less accurate receivers based on machine information are in themselves capable of recording rather accurate positioning data if the conditions are favorable.

#### 3.3. Comparison of Stands

The validation of the automatically delineated harvested stands was done by comparing the operative harvested stands from Algorithm A2 into reference stand delineations. The operative harvested stands were used instead of the balanced, inventory-style stands from the reason that balancing could not be done equally for all harvested stands due to the sparseness of the harvester data in this study. Algorithm A3 affects to the area of the delineated stands. In addition, the reference stand data were collected from operated areas which correspond with the operative harvested stands. From these reasons, the balanced stands would not give fully comparable results.

The stand pair-wise comparison was made in detailed manner as presented in Section 2.4 and Appendix G. The variables included relative areas, relative perimeters, average perpendicular displacements of the stand boundaries and estimated net shifts of the stands with respect to the reference stands.

Different pairs of stands were compared. The automated stand delineation was compared separately to field references and aerial image references. In addition, the digitized reference stands from aerial images were compared to field references, keeping the latter as "true stand delineation" in this comparison. To set the overall level of stand correspondence, these results are presented before the harvester data results.

The stands were divided into four categories by their area. The area of the harvested stand was used, to ensure that the stand belongs to same category in different comparisons.

#### 3.3.1. Ensuring the Similarity of Compared Stands

Due to the fact that the field references were collected mostly at the time of the harvesting when the results of the automated stand delineation were not yet available, it was necessary to ensure that the automatically delineated stands and reference stands can indeed be compared. There were a few considerations to take care of before making the comparison. If the field reference consisted of two separate stands and only one delineated stand appeared for the respective site from the automated method, the data points were split artificially to correspond with the field reference. An example of such a case was when the marked stand spans over road. If instead the marked stand had split into several objects in harvester data, and only one reference stand of unique harvesting type covered the marked stand, the harvester data points were artificially merged for automated stand delineation for comparison.

Unfortunately, some reference stand delineations from both sources, field and aerial images, had to be excluded from the comparison. For certain field references, later inspection with respective harvester data and aerial images showed that reference was erroneously recorded and had to leave out. One reason was that the harvester data available were not complete with respect to the observed outcome at marked stand, e.g., the clear cut at field was notably larger than what the harvester data implied. This is rather expected finding, taking into account how small part of the total amount of harvested data in time and areal coverage was included into this study. It is also possible that the data collection of this study has started in the middle of harvesting certain marked stands and therefore part of the data may be missing. Another reason for excluding reference data related to differences in harvesting types. For example, harvester data indicated a clear cut object, but at marked stand thinning area was found.

In total, 57 harvested stands were found appropriate for comparison and had either or both reference stands available. The stands consisted of 27 field observations recorded with a high-accuracy GNSS receiver and 42 digitized stands from aerial images.

There was one stand which was excluded from the harvested stand comparison, but both references could be recorded. Therefore, this stand is included in the comparison of the references, in addition to 12 stands for which the both references and harvested stands are compared.

#### 3.3.2. Comparison Results of Aerial Images and GNSS References

At first, comparison of GNSS references into aerial image references was made. The results are shown in Table 4 for 13 stands in total (one of which could not be included into harvested stand comparison). The relative areas vary on average from few percent to roughly 10% at smaller stands (for individual data points, see Figure 11). The stands digitized from aerial images were systematically slightly larger than the GNSS references. The area obtained from union of polygon geometries indicated that the stands are almost overlapping with each other.

**Table 4.** The averaged comparison results of aerial image (AI) references to GNSS reference stands from field recordings (abbreviated as GNSS) which were kept as the reference stands. Areas are given in hectares. Abbreviations: Relative area, Area(AI)/Area(GNSS); *N*, number of stands; PD, perpendicular displacement of stand boundaries in meters. The standard deviations are given in parentheses.

Area Category, ha	N	Area (AI)	Area (GNSS)	Area (AI ∪ GNSS)	Relative Area	Relative Perimeter	Average PD	RMSE PD
<0.75	3	0.23 (0.03)	0.21 (0.05)	0.25 (0.04)	1.13 (0.22)	1.06 (0.14)	1.51 (2.57)	4.80 (3.44)
0.75 - 1.5	5	1.10 (0.13)	1.02 (0.16)	1.14 (0.14)	1.09 (0.07)	0.99 (0.06)	1.68 (1.62)	4.75 (2.01)
1.5-3.0	4	1.87 (0.31)	1.77 (0.29)	1.92 (0.31)	1.06 (0.03)	0.97 (0.03)	1.44 (0.49)	4.73 (1.45)
>3.0	1	5.28 <sup>1</sup>	5.19 <sup>1</sup>	$5.44^{1}$	$1.02^{\ 1}$	0.96 <sup>1</sup>	$0.68^{1}$	5.81 <sup>1</sup>



<sup>1</sup> Since this area category had only one stand, the standard deviation could not be determined.

**Figure 11.** The area correspondence of harvested stands with respect to the references: (**a**) the relative areas; and (**b**) the differences of areas. In addition, the mutual comparison of GNSS and aerial image references is shown.

The relative perimeters mostly under value 1.0 mean that the aerial image references have shorter stand boundaries than the GNSS references. This originates from the amount of details in boundaries of GNSS references as the recording interval between positions has been rather short. Instead, in digitizing the aerial image references, the interval of points at stand boundary is related to the accuracy of the image and therefore cannot be as short as in GNSS references.

The average perpendicular displacement distances of the stand boundaries were roughly around 1 m. Positive values mean that the boundaries of the aerial image stands were outside the reference stands. The consistency of stand boundaries is deduced from the root mean square errors (RMSE) of the perpendicular displacements. Between the two reference datasets, the stand boundary agreement appears rather constant over all area categories. This finding is interpreted that both reference stands have in overall similar boundaries, and no major differences occur. Visual inspection of the data supports that.

The estimated shifts in four main directions, N–S, NE–SW, E–W and SE–NW, range from sub-meter level to several meters (Figure S3). In comparison of references, the shifts may occur due to

orthorectification or radial displacement of ortho images. It is also possible that some other errors have happened in recording of the references.

The results benchmark the overall accuracy level of recording references as well as stand-wise variation, and how area category affects the values of variables.

## 3.3.3. Adjusting the Area Parameter of the Automated Stand Delineation Method

One target of the work was to obtain as correct area as possible for automated harvested stands. The stand delineation results were initially calculated with value  $B_2 = 6.5$  m based on typical crane length of harvesters and visual inspection of data with aerial images. The stand comparison results with reference delineations were calculated using the method described in Section 2.4 and Appendix G, and the results revealed excess of area at automated stands. Additionally, a dependence was found on GNSS accuracy of the harvester. For more accurate GNSS positioning, the excess area stabilized to approximately 4%, and, in the case of less accurate GNSS positioning, roughly 10% excess of area was observed. Thus, it was obvious that some adjustment of the method is needed. The area of the stand is most directly affected by the buffering width parameter  $B_2$  in stand formation Algorithm A2. Therefore, adjusting value of  $B_2$  is the natural first choice to begin with.

To tackle the dependency of excess area on GNSS accuracy of harvester, the numerical classification of GNSS accuracy was examined in more detail for the stands in comparison. It turned out that four stands got label "less accurate" from numerical method, whereas 16 stands had label "less accurate" based on the known properties of harvester. All 41 previous more accurate GNSS stands retained the same label in numerical estimation of accuracy. Thus, 53 of total 57 stands were classified to have a more accurate GNSS positioning based on the harvester data themselves.

The readjustment of the stand area continued by selecting such stands where the match between harvested stand and reference stand was very good in visual inspection—at the best possible level taking into account the error sources in recording reference stands. A set of 9 + 3 stands was found, from more and less accurate GNSS classes, respectively. The fourth stand with "less accurate" label differed significantly from other three stands and was therefore excluded. All the selected stands had area over 0.75 ha.

The perpendicular boundary displacements of the selected stands were used to determine the average amount by which the harvested stand's boundary outreaches the reference boundary (positive value of average). To balance the stand boundaries over the whole stand, the harvested stand should be shrunk "inwards" by this average distance. Few individual distance values were discarded from average since they were clearly deviating. The reason for such local deviations were various details, e.g., few single harvester data points that would not appear in ideal cases. In averaging, the amount of distance observations per stand was used as a weighting factor (for details of the selected stands, see Table S1).

After the average perpendicular displacements for both GNSS accuracies were available, the updated values of buffer width  $B_2$  were obtained by subtracting the distance values from initial buffer width. The resulting values are  $B_2 = 5.0$  m for stands with more accurate numerical GNSS label and  $B_2 = 3.2$  m for stands with less accurate GNSS in numerical estimation.

#### 3.3.4. Comparison Results of Harvested Stands and References

The automated stand delineations were produced using the adjusted buffer widths  $B_2$  depending on the numerically estimated GNSS accuracy. The total 57 stands were compared to the both reference datasets separately, 27 stands to GNSS-recorded field references and 42 stands to aerial image references. Table 5 shows amounts of stands' numerical GNSS labels in stand area categories.

**Table 5.** The stand area categories and amounts of stands in comparison. The numbers in parentheses indicate amount of stands where the numerical GNSS estimate of harvester is more or less accurate, respectively. The abbreviation AHS denotes automated harvested stands and other abbreviations are as in Table 4.

Stand Pair	# of Stands	<0.75 ha	0.75–1.5 ha	1.5–3.0 ha	>3.0 ha
AHS vs. GNSS	27 (26/1)	7 (7/-)	11 (11/-)	8 (7/1)	1 (1/-)
AHS vs. AI	42 (39/3)	7 (7/-)	10 (10/-)	13 (10/3)	12 (12/-)

The comparison results for relative areas and perpendicular displacements of stand boundaries (PD) are in Table 6. Additionally, relative perimeters and the root mean square errors (RMSE) of PD's are detailed in Table S2. Figure 11a shows all stand pair-wise values of relative area with respect to the references. In Figure 11b, the differences of the areas ( $A_{stand}-A_{reference}$ ) are plotted to study the absolute area differences. Example delineations are shown in Figure 12.

**Table 6.** The relative areas and perpendicular displacements for compared stands. The abbreviations are as in Tables 4 and 5. The standard deviations are given in parentheses, and missing values indicate that there was only one stand in that area category.

Variable	Stand Pair	<0.75 ha	0.75–1.5 ha	1.5–3.0 ha	>3.0 ha
	AI vs. GNSS	1.13 (0.22)	1.09 (0.07)	1.06 (0.03)	1.02 (-)
	AHS vs. GNSS	1.00 (0.19)	1.00 (0.08)	1.07 (0.12)	1.03 (-)
D.L.C.	more accurate	1.00 (0.19)	1.00 (0.08)	1.03 (0.06)	1.03 (-)
Kelative area	less accurate	-	-	1.36 (-)	-
	AHS vs. AI	1.16 (0.24)	0.98 (0.08)	1.01 (0.05)	1.00 (0.02)
	more accurate	1.16 (0.24)	0.98 (0.08)	1.01 (0.06)	1.00 (0.02)
	less accurate	-	-	1.00 <sup>1</sup> (0.03)	-
	AI vs. GNSS	1.5 (2.6)	1.7 (1.6)	1.4 (0.5)	0.7 (-)
	AHS vs. GNSS	-0.6 (2.5)	-0.3 (2.2)	2.7 (3.6)	2.8 (-)
Perpendicular	more accurate	-0.6(2.5)	-0.3 (2.2)	1.7 (2.8)	2.8 (-)
distance, m	less accurate	-	-	9.2 (-)	-
	AHS vs. AI	1.8 (2.9)	-0.7 (2.0)	0.35 (2.4)	0.4 (1.5)
	more accurate	1.8 (2.9)	-0.7 (2.0)	0.6 (2.6)	0.4 (1.5)
	less accurate	-	-	-0.6 <sup>1</sup> (1.2)	-

<sup>1</sup> All (and only) these stands were used in updating the buffer width value for numerically estimated less accurate GNSS receivers.

The relative areas show that good agreement of harvested stands to references is obtained in automated method. Table 6 and Figure 11a indicate that the relative areas level close to aimed value 1.0. The average relative area is 1.01 (standard deviation 0.08) for stands over 0.75 ha, i.e., the areas of the harvested stands exceed on average only 1% the areas of reference stands. When all stands were included, the average relative area was 1.025 with standard deviation of 0.13. The reason is that small harvested stands, under 0.75 ha, have slightly larger relative area (see Figure 11a). This originates in some degree from the small values of area, as dividing by area increases the relative errors originating from small deviations in stand boundaries. When only >2-ha stands were considered, the average relative area was 1.007 (standard deviation 0.03). An additional observation from the results in Tables 4 and 6 is that the aerial image reference stands with GNSS references have on average slightly higher excess of area than automated harvested stands.



**Figure 12.** Example of correspondence of stand boundaries. The overall agreement is good, and agreement with GNSS reference is very good (stand is #2 of Table S1). The ETRS-TM35FIN (EPSG:3067) coordinate system is used in the figure. (**a**) The automated stand is on both reference stands. References are mutually shifted. (**b**) The aerial image is shown at the background. The harvested area is clearly visible, and the reference was digitized from the image. Aerial image © National Land Survey of Finland.

The area differences (see Figure 11b) keep rather constant over stand area. Most of the area differences lie within range of approximately -0.2 to 0.2 ha, independently of the stand area. The average difference is 0.02 ha and standard deviation is 0.11 ha, from which the latter can be interpreted as the rough overall error of the area of automated harvested stands. The error is assumed to originate from varying, case-specific properties somewhat randomly present in the harvester data and details of the automated method, as well as the cooperative action of these two. The average difference being close to zero, together with the observation from the relative area values, supports the inference that no major systematic bias is caused by the method.

The perpendicular displacement (PD) values measure how far the stand boundaries are from each other on average throughout the stand. The results show that the distance values are often in sub-meter scale. In wide scope, taking into account the general positioning accuracy of the GNSS receivers, this means that the stand boundaries are in balance with the reference stands. However, the stand-wise variation (i.e., standard deviation of PD in Table 6) appears to be larger than in mutual comparison of references.

The relative perimeters of automated harvested stands are given in Table S2. The values are lower than value 1.0 which indicates that the delineated stands have systematically few percent to 10% shorter boundary line than the references. This relates to the interval of recording the boundary points of the reference stands, as described already in Section 3.3.2. Based on visual observations, the boundaries of the automated stands have on average rather long intervals of the points with respect to the reference stands which shortens the perimeters.

The RMSE of perpendicular displacements in Table S2 describes the agreement of compared stands via amount of deviation between the boundaries of compared stands. Harvested stands have clearly larger values of RMSEs of PD than the references compared mutually (Table 4). The estimated accuracy of the GNSS receiver of harvester increases slightly the RMSE values if the stand has less accurate GNSS positioning.

The values of both PDs and their RMSEs are also implying that the case-specific properties of the harvester data play notable role in positioning of the stand boundaries by causing local deviations. However, the areas correspond well and overall match of the stand boundaries is good, thus local differences in the stand delineations are regarded as less significant.

The systematic shifts of the harvested stands were also studied. The numerical shift values obtained in four main directions, N–S, NE–SW, E–W and SE–NW, are presented in Figure S4. The shifts were inspected visually as defining clear boundary values for shift was difficult with current small data. In addition, the amounts of samples in each direction varied largely from few to several tens. Therefore, the visual examination was the most reliable way here to detect the real shifts. The stands of significant shifts (see Figure 13) and size changes were counted. The results are in Table S3.



**Figure 13.** Examples of shifted harvested stands. The ETRS-TM35FIN (EPSG:3067) coordinate system is used in the figure. (a) The stand has shifted towards east with respect to the GNSS reference. (b) The stand has shifted towards northwest with respect to the aerial image reference.

One observation was that the automated method itself caused size changes for six stands. The reasons were difficult geometries, for example deep, narrow non-harvested areas inwards to the stands. Such areas were delineated outside of the references, but the automated method did interpret them as stand areas and caused expansion of stand. Another typical case where expansion occurred was at the beginning of the external strip roads leading to stands. If the external strip road continues directly from stand tree points, few trees that visually would belong to strip road, are classified as stand tree points in Algorithm A1.

More shifts originated from the differences of the harvester data and the reference stands than from the method. In counting the results during visual inspection, case where both shift and size change were observed for same stand, the shift was prioritized over size change (i.e., some stands included into shift-category also included size change). In total, roughly 1/3 of the stands had either shift or size change. Direction dependence of the shift values was not observed in this dataset.

Regarding the GNSS accuracy of the harvesters at stands, the data contained four cases of less accurate GNSS label in numerical estimation. All these stands displayed either shift or change of size. However, quite many shifts occurred also for the stands of more accurate GNSS estimate (Table S3). This implies that the less accurate GNSS probably causes significant shift values, whereas the more accurate GNSS does not guarantee low shift values for automated stands. It is also worth noting that even the mutual comparison of references included few shifts and size changes of stands (Table S3).

Certain reasons for shifts of harvested stands were identified. As clearly indicated by the results, the positioning error of GNSS can itself cause a systematic shift of several meters. The positioning error further indirectly depends on various factors, e.g., topography, properties of growing stock at stand and conditions including weather. The working preferences of the harvester driver may also lead to shifts, since in the stand delineation method the use of harvester boom is averaged. The driver can and has to make case-by-case decisions and for example limit to half angle of operation at one

side of harvester, if the harvester drives along the stand boundary. These can occur at any stand, independently of GNSS accuracy of harvester.

#### 4. Discussion

The results confirm that automated stand delineation does produce appropriate stand delineations overall. The method manages to handle various input cases that appear in operative harvester data, as the dataset of this work included real, operative data. Method recognizes the stands within the harvested objects and excludes the sparsely harvested tree points outside the marked stand. Therefore, it allows conversion of point-wise data from wood procurement-based objects into harvested stand geometries efficiently and with good accuracy, and it provides the stand classification at the level of single harvested trees. This will, in general, provide great opportunities for further analyses based on harvester data.

The delineated stands were found to correspond well with the reference stand delineations, when the buffer width of Algorithm A2 was adjusted with selected stands and the numerical estimation of GNSS accuracy was applied at harvested objects. The shapes of the stands were in good overall agreement, although some local deviations were observed, for example at locations where the external strip roads join the stands or if the GNSS positioning of the harvester was less accurate. Local geometries of the harvester data points were in some cases difficult to delineate correctly automatically. Stand shifts over approximately 4–5 m took place in roughly 1/3 of the data, and they remain currently as an intrinsic error of the harvested stands.

The amount of reference stands from two different sources was limited, being 57 in total. Acquiring reference data has rather strict conditions that need to be filled which affected the total amount of references in this work. Additionally, the data of this study were collected mainly from one laser scanning inventory area, being a very limited part of Finland. Despite the somewhat larger structural variation of the forests at the study area, a spatially covering data of the entire country is needed to confirm that the method works at geographically different sites. Increasing the amount and expanding the spatial coverage of the reference dataset to other geographic areas of Finland are seen necessary in the future research, to ensure the fitness of the buffer values. This is especially the case with those objects where the GNSS positioning of harvester is less accurate by numerical estimation.

Automated validation method of stand boundaries which was developed for this work, turned out to work well and provide relevant information of the agreement of the compared stands. It enables to repeat the stand comparisons effectively in the future with more data. Only the automation of recognizing directional shifts of compared stands was not achieved yet and therefore it remains as a further study topic. In contrast to previous comparison methods of stands [26–29], the current method is more detailed and accurate.

The references themselves are always subjective, as found also in [27]. The subjectivity relates to the fact that, in principle, the stand delineations may not necessarily be determined exactly from forests due to features of nature environment. The references may contain errors, too. The field references typically have positioning error approximately 1 m from the GNSS device only (see also [30]). The accurate determination of harvested area may also contain uncertainties when using aerial images. The identification of individual tree crowns from aerial images is a non-trivial task where forming larger entities from aerial image pixels for interpretation results more reasonable delineated areas [31]. Orthorectification of aerial images depends mostly on the digital elevation model used [24,32]. The accuracy of rectification is studied in [33] where error level of  $\pm 5$  m was observed. For the aerial image, sensors and imaging height may have caused errors. Together these sources of errors and the positioning accuracy of the harvester are roughly at the estimated lower limit of stand shift. It is noteworthy, that the area errors between harvested stands and reference stands were at similar level as the errors between aerial image references and field references recorded with high-accuracy-GNSS device. Considering the philosophical aspects of where correct stand boundary should locate, general level of GNSS positioning accuracy of harvester, and the possible errors in delineating reference stands, the overall correspondence of the operative stand boundaries was judged good. In earlier studies [26,27], the area errors were few to tens of percent, and this work obtained notably smaller error, roughly 1% error for stands above 0.75 ha. Such average error in automated stand areas is regarded as adequate accuracy level for the stand delineations for operative purposes. The error is expected to have minor economic implications, if the stand data were used in, e.g., evaluating the costs of regenerating the marked stand. Areas of individual automated stands do contain error, but, e.g., similar percentual error in basal area at final fellings is estimated to have larger economic impact. It is also worth noting that in the current situation the forest use declarations or harvesting instructions are the best available knowledge of the stand area, and the realized area at marked stand can differ from the a priori information.

The obtained accuracy of automatically delineated stand polygons enables using them in the update of the forest inventory data. The stand delineations provide up-to-date information of harvested sites (starting time and harvesting type) to complement the remote sensing-based inventory between its repetition cycle of years. The harvested stands can be available for forest inventory update in principle in few days from ending the operation, when assuming an automated processing chain of harvester data and including the latencies in sending the harvester data from the machine to stand delineation. The computing times of the automated stand delineation in cloud environment are typically few minutes per object, meaning that the computing time facilitates obtaining the operative stand delineations for inventory update at fast schedule. It is difficult to estimate the economic aspects of the up-to-date forest inventory data for the forest sector since the benefits are distributed between different actors. However, one way to illustrate the importance is the cost of laser scanning cycle, and even that cannot be done on short time intervals in which the automated stand delineations can be provided. In addition, the automated stand boundaries are adequately correct, and amount of manual work in updating the forest inventory would notably lower if harvested stands are used.

The operative stands obtained from Algorithms A1 and A2 are regarded suitable for updating the forest inventory. In adding the operative stand into existing stand network, the stand geometries can intersect at various manners, somewhat similarly as in tests of Algorithm A3 in current work. However, the update of forest inventory requires, in any case, detailed rules how the overlapping stand geometries are solved based on stand attributes. Therefore, a topic of future study is to develop the rules for solving the apparently problematic overlapping cases. This could be done utilizing the stand attributes, most importantly harvesting type and start time, in addition to mere stand geometry. By using such rules, it is expected that all the automatically delineated stands can be utilized in updating of the forest inventory. Using additional data from, e.g., topographic database [34], such as location information of roads, agricultural lands, water areas, etc., is also recommended in converting the operative stands to forest inventory stands. In a wider scope, Algorithm A3 is seen as a beneficial tool for maintaining balanced forest stand geometries in inventory, regardless the origin of the stands to be embedded to the stand network. The local errors in stand boundaries are also expected to smoothen in the embedding.

The harvesting type is essential input information for the automated method, as the results are found to depend on harvesting type in different manners. In overall, the results confirm the fact that delineating hold-over cutting stands is not very meaningful even in principle, using any method, due to the sparse, varying density of cut trees. Instead, in cases of other harvesting type, the removal density is found to be notably more uniform, even when taking into account the natural variation in density of trees within different harvested stands (as shown in Table 1). This facilitates limiting the use of density-based classification approach for the most common harvesting types in Algorithm A1. Regarding hold-over cuttings, what could be done based on harvester data is identifying the stands from some other existing stand network (e.g., Finnish forest inventory stand network) from which the hold-over trees were harvested.

The success of the automated stand delineation depends on the quality of the input data. How harvested objects are outlined affects directly the major-level classification within which the stand clustering in Algorithm A1 is performed. For example, if harvesting at certain marked stand has been interrupted and continued later, and a new starting time of harvesting is given when continuing, two separate objects will appear in the harvester data at that site. Often, the method results overlapping stand polygon geometries in such case. However, when embedding the stand into forest inventory stand network, the overlapping cases are easily solved by examining the harvesting types and starting times of operations. If only the harvesting type group is available for harvester data instead of the more detailed harvesting type, the method combines all the adjacent tree points of such object into same stand, without knowing if the marked stand contains same harvesting type or e.g., clear cut and seed tree cut next to each other. When such stand is embedded into forest inventory stand network, error in harvesting type will remain in the updated inventory, unless it could be removed with additional, external information.

The adjustment of the various parameters in Algorithms A1–A3 is a multifaceted issue. It is subject to discussion as "universal truth" rarely exists in numerical problems, or in stand delineation as well. The current parameter values were selected by examining the harvester data and based on expert knowledge of each topic. The parameters were, however, designed adjustable in the algorithms since the current dataset is limited both in length of sampling time and vastness of spatial coverage. For this reason, when more data are accumulating, the current selected values may need revision. There are several affecting factors behind the observed features of the harvester data, e.g., properties of marked stands, harvester assembly configuration including boom length, targets of wood procurement regarding the cut trees, etc. Changes in any of these factors may result slight changes to parameter values of algorithms in order to obtain the best possible outcome in stand delineation. Collecting many data is a recommended future research activity to study the detailed effects of parameter changes and ensure the appropriateness of used values. It is even possible that the adjustment of parameters may not solve correctly each single case that can appear when more operative data are available for putting through the automated method. This can be the case with, e.g., details of the point geometry.

Harvesting type was found to affect to Algorithm A1 based on the current work. Some thinning stands were found to have slight tendency to contain small empty slots, often longitudinal of shape, parallel and in between strip roads of the stand. The reasoning for this observation is obvious from the harvesting work point of view: at thinnings the strip road interval is aimed to keep as large as possible to avoid losing growth area of trees for strip roads. For the algorithm, this finding implies that increasing the sidewise distance parameter of Algorithm A1 for thinnings may be one of studied topics in the future. Other factors that affect Algorithm A1 include, e.g., density of cut trees, depending on the growing stock density before harvesting, tree-wise species or tree dimension data obtained or deduced from stem measurements within harvester data. However, the variation is generally speaking large, and it is possible that local correlations somewhere in the dataset may not be successfully extrapolated to entire harvesting data. Studying the mentioned factors and dependencies provides several recommended future topics to refine the parameters of Algorithm A1. It would also be intriguing to test more sophisticated methods than current methods, e.g., neural networks, to find weaker correlations of input variables in stand classification. Current method can serve to provide training data for such a study.

One issue in Algorithm A1 is that it does not recognize directly the longitudinal, one-strip road wide stands that are still a relevant feature of the harvester data. This issue was solved in this work by estimating the areal density of the cut trees at strip road parts. This approach, however, turned out rather challenging. Firstly, estimating the appropriate area for strip road part from which the trees were harvested around the location points was more uncertain than at "full-scale" harvested stand. After calculating the removal density values for strip road parts, visual inspection of the dataset indicated presence of notable variation. For example, of two strip road parts having the same removal density, one was deemed to be a longitudinal stand and the other a strip road. Therefore, the removal

density limit was seen necessary to raise to such level that above it no "strip road" cases further appeared i.e., only the "certain cases" are selected. In other hand, this directly leads to the fact that some of those cases which were longitudinal stands based on visual observations, were classified as strip roads. However, with current setup this is the best result, keeping in mind that the amount of sparser strip road parts grows rather rapidly when removal density lowers, and the strip road parts tend to include more miscellaneous harvested areas. Overall, the reverted strip road stands were found to improve the correspondence of the stands when included into the stand geometries. In the future, separation of the external strip roads and other harvested areas outside the marked stand directly in the harvester data is recommended to clarify these cases.

Algorithm A2 was found to create non-harvested empty slots within stands. From the premises of the harvester data and the method, certain minimum size limits prevail in forming of the slots. This phenomenon is interesting as the empty slots may represent habitats of special importance, but additional information is needed to ensure that. The topic is recommended for future study. Extraction of the empty slot geometries into a separate dataset from the stand polygons is nevertheless considerable before filling small slots for embedding the stand geometry into stand network. The appropriate limit area for filling the slot has to be selected depending on the case.

The appearance of isthmuses at results of Algorithm A2 is disadvantageous as it generates errors to the automated stand delineations. The issue demands that further studies of adjusting the parameters are needed, or alternatively a more detailed condition for selecting the triangles should be formulated. Obtaining the preferred outcome can be, however, a double-edged problem as reducing the value of parameters causing isthmuses would mean at least more empty slots appearing within the stands and possibly some other changes into the outcoming harvested stand delineation, e.g., at the edges of the stand. However, filling the empty slots is expected to be easier than dealing with the isthmuses in some other automated manner if areal and other changes near edges would be acceptable. In addition, it is expected that the isthmuses will become less significant when the stands are embedded into the forest inventory stand network.

The parameters in Algorithm A3 are used in setting geometrical details of the division line and help to identify the correct stand to which a split part of intersection is merged. Adjusting the numerical values of the parameters has rather slight effects on the result. Similar to Algorithm A2, empty slots appear in Algorithm A3 in certain intersections. Splitting them or not depends on the case-wise suitable limit area. Further research of the appropriate position of updated stand boundary is recommended. The stock properties at the intersection zone are suggested as one factor in adjusting the stand boundary.

Few geometry errors occurred in Algorithm A3 during geoprocessing in GIS program which required inclusion of several automated geometry cleaning stages. The appropriate level of cleaning the geometries varied case-by-case, making the successful geometry cleaning difficult. It was also observed, that geoprocessing operations did not occur as they should have, although no geometry error was programmatically found.

The accuracy of GNSS receivers was observed to depend on other factors in addition to the harvester properties. This can be seen from the results when the GNSS classification of harvesters was compared with the numerical estimation results. Of course, the numerical method presented in this work is very simple, rather a first attempt towards this direction, and it has drawbacks. Most importantly, the method does not take into account the density of the points along or perpendicularly to the strip roads which can distort the outcome in some degree for such stands where the removal density is lower. However, the removal density is aimed to be large enough when planning the harvester operations which probably lowers the amount of harvested sparser stands. Despite the simplicity of the numerical estimation method, the results clearly indicate that the harvesters with "less accurate" receivers can often obtain still the "more accurate" label in numerical estimation. A closer visual inspection supports the results on those stands where the classification differed in this way. Visually, it was confirmed that the local accuracy within harvested stands varies more, and indeed

a notable portion of the stand can have rather accurate positioning of harvester. This emphasizes the contribution of conditions on the observed GNSS accuracy. Typically, the lower accuracy took place when the harvester was in topographically challenging site, e.g., at lower side of a steep hill or cliff where the height differences are large. The landscape has temporarily shaded the view of positioning satellites. The harvesting type was also found to have some effect; thinnings had lower accuracy than the final fellings. This is reasonable as the remaining trees cause the same shading effect to the visibility of the satellites, thus lowering the net accuracy. Similar findings are presented in works [35,36]. Based on these considerations, it is recommended to study in further contexts the net effective GNSS accuracy using numerical methods. In addition, obtaining information of GNSS receivers and positioning conditions (satellite geometry, geometric and position dilution of precision, etc.) in standardized format from StanForD STM/HPR files is important for improving the accuracy estimate. The reason for this is that those objects which were harvested with a machine known to be more accurate in GNSS positioning, were also found "more accurate" in numerical estimation.

The tree-wise removal information of stands, obtained, e.g., as explained in Appendix F, provides significant possibilities. If areally covering harvester data were available, the accurate removal could be determined for all harvested stands based on the stem measurements recorded by harvester, and further processing of the stem data such as obtaining the "ideal" stem curves [23] numerically. Accurate data of stand removal would enable using harvester data as a ground reference for calibrating ALS measurements, as piloted in [15]. That would lower the costs of performing forest inventory since less manual work would be needed. The tree-wise data may also be used in updating stand-wise forest inventory in the future. More detailed compilation of how to gather the tree-wise data for delineated stands is seen beneficial in supporting these purposes.

Updating of grid-level ( $16 \text{ m} \times 16 \text{ m}$ ) forest inventory may be relevant in the future as well. However, to reach and maintain the accuracy of grid-level inventory, the mere location of harvester is not adequate in representing position of harvested tree, and the crane position during cutting the tree would be needed. Naturally, that would also benefit the upkeeping of stand-wise forest inventory by providing spatially detailed and accurate tree-specific information.

### 5. Conclusions

An automated method to delineate operative stands from harvester location measurements is presented. The operative stands were produced by firstly clustering the harvester data points into stand tree groups in Algorithm A1. Here, the density of the harvested tree points was found to be an essential feature. Such stands which have only one longitudinal strip road (i.e., dense strip road), were separately recognized and included into the stands. Then, the stand tree groups were polygonised into stand geometries in Algorithm A2 with help of Delaunay triangulation. Based on the tests with the current dataset, one can conclude that the automated stand delineation works well and produces reasonable stand delineations for all harvesting types except hold-over cuttings, since there the density of cut trees often differs from other harvesting types.

To assist the automation of stand delineation, additional methods were developed. Firstly, estimation of the accuracy of the GNSS positioning of the harvester based merely on the location points was necessary. The method presented in this work is rather simple but works effectively. The results indicate how the net accuracy of the GNSS positioning varies within and between harvested objects, and conditions and environment clearly affect the accuracy. The topic would benefit from further development of the estimation tool.

Another method was developed to compare two stand boundaries thoroughly for validating how well the stands correspond with each other. The validation method allowed to adjust easily the buffer parameters of Algorithm A2 to obtain balance of areas between harvester and reference stands. After adjustment, the comparison results indicated that the overall correspondence to references is good, and the error of area is approximately 1% on average for stands over 0.75 ha. Therefore, a conclusion can be drawn that the automated harvested stands are suitable for updating the forest

inventory data. Finnish Forest Center prepares to use harvested stands as one source of information in updating of the forest inventory data which benefits the entire field of forest sector by more up-to-date forest resource information.

To convert the operative stand delineations towards the forest inventory stands, a method to balance the overlapping stands was introduced. Algorithm A3 creates a new, solid stand boundary between the stands. It has certain limitations due to the stand geometries, but if stand attributes were taken into account, for example in update of forest inventory data, the overlappings could be handled completely. Overall, Algorithm A3 turned out as a beneficial tool for updating stand networks and retaining balance of the previous and added stands.

Supplementary Materials: The following are available online at http://www.mdpi.com/2072-4292/12/17/2754/s1, Figure S1. The tree points of both intersecting stands are distributed in the intersection area. Harvesters were known to have a more accurate GNSS receiver. The ETRS-TM35FIN (EPSG:3067) coordinate system is used in the figure. (a) The tree points are rather evenly distributed in the intersection zone. (b) In this example, only few tree points are in the intersection, Figure S2. Examples of geometrically non-splittable harvested stand intersections. Intersections with four vertex points cannot be split in Algorithm A3. The colors indicate the harvested objects, and the points mark the locations of the harvester. Vertex points are shown with larger black dots. Harvesters were known to have a more accurate GNSS receiver. The ETRS-TM35FIN (EPSG:3067) coordinate system is used in the figure. (a) One stand overlaps another stand. (b) One stand hits an empty slot within other stand and overlaps outside that stand's boundary which makes the intersection non-splittable, Figure S3. The estimated net shifts between aerial image and GNSS references. The plot shows the observations for each four direction N-S, NE–SW, E–W and SE–NW for each compared stand pair in compass view. At radial direction is the estimated distance of the shift, and at circular axis is the bearing in degrees. The amounts of sampling points behind the individual shift observations vary from few to several tens, Figure S4. The estimated net shifts between automated harvested stands and references. The plot shows the observations for each four direction N–S, NE–SW, E–W and SE-NW for each compared stand pair in compass view. At radial direction is the estimated distance of the shift, and at circular axis is the bearing in degrees. The amounts of sampling points behind the individual shift observations vary from few to several tens. (a) The automated harvested stands compared to GNSS-recorded field references; (b) The automated harvested stands compared to aerial image references, Table S1. The details of the stands selected into adjustment of the area of automated harvested stands (AHS). The area and perpendicular displacement values (PD) to reference stands are determined using initial buffer width  $B_2 = 6.5$  m. The stands 1–9 belong to more accurate GNSS and stands 10–12 to less accurate GNSS in numerical classification, Table S2. The relative perimeters and the root mean square errors (RMSE) of perpendicular displacements for compared stands. The abbreviations: AI, aerial image references; AHS, automated harvested stands; GNSS, field references. The standard deviations are given in parentheses, and missing values indicate that there was only one stand in that area category, Table S3. The amounts of harvested stands where systematic shifts or partial size changes appeared with respect to reference stands. If stand was shifted or changed size with respect to both references, it is included in numbers of both references. In the results, two stands were shifted and other two stands changed size with respect to both reference stands-these occurred in each area category of stands. With respect to GŇSS reference, eight stands were shifted and six other stands changed size at stand sizes < 3 ha. When comparing with aerial image references, seven stands were shifted and ten other stands changed size. The GNSS accuracies of harvesters were determined numerically. In addition, stand size changes caused by the automated method are given separately. The results for mutual comparison of reference stands are also presented. For abbreviations, see Table S2.

**Author Contributions:** Conceptualization, T.M., K.R. and J.-A.S.; methodology, T.M., K.R. and J.-A.S.; software, K.R. and J.-A.S.; validation, T.M. and K.R.; formal analysis, K.R.; investigation, T.M. and K.R.; resources, T.M. and K.R.; data curation, T.M. and J.-A.S.; writing—original draft preparation, K.R. and T.M.; writing—review and editing, T.M. and K.R.; visualization, K.R.; supervision, Metsäteho Oy; project administration, T.M.; and funding acquisition, Metsäteho Oy. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Ministry of Agriculture and Forestry of Finland (grant number 173/03.02.02.00/2016).

Acknowledgments: The authors thank Metsäteho's owners—StoraEnso Oyj (Helsinki, Finland), Metsä Group (Espoo, Finland) and United Paper Mills UPM Corporation (Helsinki, Finland)—and their entrepreneurs for their help with logging preparations and their excellent cooperation during the logging. Ponsse, Komatsu Forest and John Deere are acknowledged for their assistance with the data collection. In addition, the authors thank Heikki Pajuoja, Jukka Malinen, Jarmo Hämäläinen and Tapio Räsänen from Metsäteho Oy for all the support to the work. Juho Heikkilä and Juha Inkilä from Finnish Forest Center are acknowledged for successful cooperation within the grant by Ministry of Agriculture and Forestry of Finland.

Conflicts of Interest: The authors declare no conflict of interest.

## Appendix A

A method for estimating GNSS positioning accuracy of harvester is presented below. It is based on averaging the distances between individual felled trees and utilizes the StemNumber -variable from the harvester data. Figure A1 below shows the properties of the data that are behind the method. The notation // indicates the comment lines within the method descriptions in all appendices.

Estimation method of GNSS accuracy:

Input: harvester location points P of one harvested object

Convert the  $N_p$  input points into line which is ordered by StemNumber. This ordering is based on the cutting order of the trees (see Figure A1).

Explode the line geometry into separate line segments.

// Measure the total length of the segments which are not too long (some segments can be rather // long due to the cutting order and harvester movements, see Figure A1c). Initialize  $L_{GNSS} = 0$ .

Loop over line segments *s*, and exclude segment if it is too long:

If L(s) < 20 m:

 $L_{GNSS} = L_{GNSS} + L(s).$ End if.

End loop.

A.1. Obtain the average length per total amount of segments:  $L_{GNSS} = L_{GNSS}/(N_p - 1)$ .

// Set the GNSS accuracy labels.

Initialize quality variable  $Q_{GNSS}$  = "More accurate".

// Change the label if needed based on the obtained average length  $L_{GNSS}$  and the amount of data // points.

```
If (L_{GNSS} > 3.5 \text{ m} \text{ and } N_p > 120):

Q_{GNSS} = \text{"Less accurate"}.

End if.
```

It would be possible to calculate the average length in Stage A.1. per amount of included segments as well, and slightly adjust the limit parameters correspondingly. However, often the amount of long segments is small.



**Figure A1.** Location data points from two different harvesters at clear cutting objects. The line is drawn on the points of harvested object in cutting order of the trees, based on StemNumber. Note the scales of the subfigures. The ETRS-TM35FIN (EPSG:3067) coordinate system is used in the figure. (a) The harvested object at basemap. (b) The harvested object at aerial image. (c) The overall cutting order of one object. Harvester was known to have less accurate GNSS receiver. (d) At the upper right corner of Figure a, the recorded locations of consequent trees have distances of several meters, showing a presence of GNSS positioning error. (e) At the left part of Figure a the positioning accuracy is visually better than in Figure b. (f) For comparison, data of object where harvester was known to have a more accurate GNSS positioning. Basemap and aerial image © National Land Survey of Finland.

## Appendix B

Algorithm A1 classifies the harvester data location points into stand tree groups and separates external tree points that are harvested from the strip roads leading to stands.

Algorithm A1: Stand clustering and strip road separation

```
Input: harvester location points P of one harvested object
```

Set the parameter values based on the accuracy of GNSS positioning.

 // Currently, only one parameter is adjusted by the GNSS accuracy: number of points for averaging

 // Nave. For more accurate GNSS,  $N_{ave} = 11$  and for less accurate GNSS,  $N_{ave} = 31$ .

 // Calculate average locations for all input points.

 Loop over points  $p_i \in P$ :

 Request  $N_{ave}$  nearest points  $P_N$  from P with respect to the point  $p_i$  by spatial query.

 // Find points which are closer than stand buffer  $B_S = 23$  m to  $p_i$ .

 Loop over  $p_n \in P_N$ :

 If distance  $D(p_i, p_n) > B_S$ :

 Remove  $p_n$  from  $P_N$ .

 End if.

 End loop.

Obtain averaged coordinates, noted as  $\langle p_i \rangle$ , from remaining points  $p_n \in P_N$ .

End loop.

// Make preliminary classification of the points into two categories, stand or strip road leading to // stand. If point has adjacent points as suitable distance, it belongs to stand. See Figure A2. Loop over points  $\langle p_i \rangle \in P$ :

For point  $\langle p_i \rangle$ , obtain the closest point  $\langle p_c \rangle$  where  $D(\langle p_i \rangle, \langle p_c \rangle) > 5$  m. Determine the bearing of the found point  $\langle p_c \rangle$  from the point  $\langle p_i \rangle$ . Convert the bearing into scale  $\beta \in [0^\circ, 180^\circ]$ .

Create sector polygons  $G_{sec}$  perpendicularly to the converted bearing direction  $\beta$ , using sector angle  $\gamma = 22^{\circ}$ :

$$G_{sec} = sector(\alpha, r = 30 \text{ m}) - sector(\alpha, r = 8 \text{ m}),$$
(A1)

where  $\alpha = [\beta_- - \gamma, \beta_- + \gamma] \cup [\beta_+ - \gamma, \beta_+ + \gamma]$  and  $\beta_{\pm} = \beta \pm 90^\circ$ . // Find if the sector polygons intersect with other averaged points.

B.1. // Find if the sector polygons intersect with other average If  $G_{sec}$  intersect with other point locations  $\langle p_i \rangle \in P$ ,  $j \neq i$ :

			-
Set index1	= 1	for	$< p_i >$ .

Else:

```
Set index1 = 0 for \langle p_i \rangle.
```

End if.

End loop. // Smoothen the classification by averaging the preliminary classification results (see Figure A3a

// and Figure A3b.) For that, define strip road buffer value  $B_{sr} = 8$  m.

```
B.2. Loop over all points \langle p_i \rangle \in P:
```

Request points  $\langle p_a \rangle \in P_A$  from P such that  $D(\langle p_i \rangle, \langle p_a \rangle) \langle B_{sr}$ . Average the values of *index1* of points  $P_A$ :  $\langle index1 \rangle = \Sigma(index1)/N(P_A)$ . // Study if  $\langle index1 \rangle$  for  $P_A$  exceeds limit value 0.4. If  $\langle index1 \rangle \ge 0.4$ : Set *index2* = 1 for  $\langle p_i \rangle$ . Else: Set *index2* = 0 for  $\langle p_i \rangle$ . End if.

End loop.

```
B.3. // Aggregate stand and strip road points into multipoint "motifs" M<sub>k</sub> where k denotes the index of // the motif. Each motif relates to either stand points or strip road points, depending on the value // of index2 of the tree points assigned to the motif. Initialize k = 0.
Loop over points [<p<sub>1</sub>>, <p<sub>N(P)-1</sub>>] ∈ P:
If <p<sub>i</sub>>(index2) == <p<sub>i+1</sub>>(index2):
```

Switch  $\langle p_i \rangle$  (index2): Case 1: Label  $\langle p_i \rangle$  to (stand) motif  $M_k$ . Case 0: If  $D(\langle p_i \rangle, \langle p_{i+1} \rangle) \langle B_{sr}$ : Label  $\langle p_i \rangle$  to (strip road) motif  $M_k$ . Else: Increase motif index k to k+1. End if. End switch.

Else:

Increase motif index k to k+1.

End if.

End loop.

// Handle last point  $\langle p_{N(P)} \rangle \in P$ .

Switch  $\langle p_{N(P)} \rangle$  (*index2*):

Case 1:

Label  $\langle p_{N(P)} \rangle$  to (stand) motif  $M_k$ .

Case 0:

Label  $\langle p_{N(P)} \rangle$  to (strip road) motif  $M_k$ .

End switch.

Merge all motifs  $M_k(index2 == 1)$  to  $M_{\cup}$ , and set index3 = 1 for points  $\langle p \rangle \in M_{\cup}$ . If  $M_{\cup}$  has type "multipolygon", convert it to single polygons. Buffer  $M_{\cup}$  and all  $M_k(index2 == 0)$  with respective buffer distances  $B_S$ ,  $B_{sr}$  (see Figure A3c).

Merge the overlapping  $m_{k,B} \in M_{k,B}$  (*index*2 == 0) to  $M_{\cup,B}$  and their points (see Figure A3c). // To identify spur trails, limit value of relative area  $A_r$  is obtained from geometric considerations // of parameter  $L_{st} = 15$  m which describes the maximum spur trail length which should be still // merged into the stand (*i.e.* it is too short for being a strip road leading to stand):

$$A_r = (\pi (B_{sr})^2 / 2 + 2(B_S - L_{st})B_{sr}) / (\pi (B_{sr})^2 + 2B_{sr}L_{st}).$$
(A2)

// Find if  $m_{\cup,B} \in M_{\cup,B}$  overlap by area  $A_r$  with  $M_{\cup}$ . Loop over  $m_{\cup} \in M_{\cup}$ : Loop over  $m_{\cup,B} \in M_{\cup,B}$  (index2 == 0): If  $A(m_{\cup} \cap m_{\cup,B}) > A_r$ : Set *index3* = 1 for points  $\langle p \rangle \in m_{\cup,B}$ . // The strip road motif  $m_{\cup,B}$  is a spur trail and it is returned into the stand. See // Figure A3c,d. Else: Set *index*3 = 0 for points  $\langle p \rangle \in m_{\cup,B}$ . End if. End loop. End loop. Aggregate points P(index2 == 1) again into motifs  $M_{k,uvd}$  similarly as in Stage B.3. but using distance condition  $D(\langle p_i \rangle, \langle p_j \rangle) \langle B_S$ . Buffer motifs  $M_{k,upd}$  with  $B_S$  to motifs  $M_{k,upd,B}$ . Merge the points of overlapping  $m_{k,\mu pd,B} \in M_{k,\mu pd,B}$  to form groups of stand points  $P_S$ . Give unique labels StandID for the stand-wise point groups  $P_S$ .



**Figure A2.** The averaged point data is used to find the directions (shown with black lines) which resemble the strip road network at the stand. The sector polygons are created based on those directions and used to classify the points into stands or external strip roads leading to stands. The ETRS-TM35FIN (EPSG:3067) coordinate system is used in the figure.



**Figure A3.** An example of averaged data points in stand recognition in Algorithm A1. The data is from thinning cut object harvested by a machine with a more accurate GNSS. The ETRS-TM35FIN (EPSG:3067) coordinate system is used. (a) The stand indices *index1* after Stage B.1. of Algorithm A1. (b) The locally averaged stand indices *index2* after Stage B.2. of Algorithm A1. (c) The stand motifs have been merged and buffered (blue), and the overlapping buffered strip road motifs have been merged (orange), both based on *index2*. The purpose is to find, using the intersection of buffer areas, which buffered strip roads are either within stand, spur trails or external strip roads leading to stands. (d) The result of Algorithm A1: tree points have been classified into two categories, stand and external strip road.

## Appendix C

Sometimes a harvested stand can comprise only one longitudinal strip road. Algorithm A1 cannot recognize them as a stands, as no neighboring points lie aside the strip road. Such cases are called as dense strip roads in this work, and are identified based on the removal density of the strip roads in the following manner.

#### Recognizing the dense strip roads:

Input: strip road threes  $P_{SR}$  from Algorithm A1 of one harvested object

// Calculate average location for the strip road points in similar way as in Algorithm A1 but with // different parameter values. Here a single value for  $N_{ave} = 15$  is used for all strip roads

// independently from GNSS accuracy, as the variance in strip road point densities is larger than

// in typical stands, especially towards the sparse values, and therefore the role of GNSS accuracy // is not as pronounced.

Loop over points  $p_i \in P_{SR}$ :

Request  $N_{ave}$  nearest points  $P_N$  from  $P_{SR}$  with respect to the point  $p_i$  by spatial query. // Find points which are closer than strip road distance  $D_{sr} = 10$  m to  $p_i$ .

Loop over  $p_n \in P_N$ :

If  $D(p_i, p_n) > D_{sr}$ :

Remove  $p_n$  from  $P_N$ .

End if.

End loop.

Average the coordinates of remaining points of  $P_N$  and store the averaged location for  $p_i$ .

End loop.

// Buffer the averaged strip road points.

Loop over points  $p_i \in P_{SR}$ :

Buffer the averaged coordinates of point  $p_i$  to geometry  $g_i$  by buffer width  $B_{temp} = W_{sr} + D_{sr}/2$ , where strip road half width  $W_{sr} = 2.25$  m.

End loop.

// Merge the buffered point geometries into identifiable strip road parts in the following manner.

Merge the overlapping  $g_i$  into one polygon geometry  $G_{\cup,sr}$  and the respective points into group of points  $P_{\cup,sr}$ . Obtain the amount of points  $N_{sr} = N(P_{\cup,sr})$  in each strip road part.

// Find those strip road parts that should rather be stands than strip roads leading to stands.

// They have high enough removal density of trees per unit area,  $D_A$ .

Loop over strip road parts  $P_{\cup,sr}$ :

Calculate removal density  $D_A = N_{sr}/A(G_{\cup,sr})$ .

// Check the removal density of strip road part.

If  $(D_A > 0.05 \text{ points/m}^2 \text{ and } N_{sr} \ge 15)$ :

Give unique StandID for points  $p_i \in P_{\cup,sr}$  to label the points as a stand.

End if.

End loop.

## Appendix D

The stand polygon is delineated from grouped stand tree points in Algorithm A2. Outlier points are removed during the delineation. Harvested objects are handled one at a time and the dense strip road polygons are included into the resulting stand polygons at the end of the algorithm.

#### Algorithm A2: Stand polygon formation

Input: stand points P<sub>stands</sub> for one harvested object (including dense strip road points)

Find unique values of StandID from *P*<sub>stands</sub> and add them to list *Stands*.

Loop over  $SI \in Stands$ :

Select points of one stand  $P_S = p_s \in P_{stands}$ ,  $p_s(StandID) == SI$ .

// Check that the group of stand points  $P_S$  consists of at least three points at different locations // which are not farther than outlier distance  $D_{max} = 19$  m from each other.

If  $(\nexists p_1, p_2, p_3 \in P_S \text{ such that } p_i(x_i, y_i) \neq p_i(x_i, y_i) \text{ where } i, j \in [1,3] \text{ and } D(p_i, p_i) < D_{max})$ :

Stop the geoprocessing here. No resulting stand delineations will be produced for  $P_S$ . End if.

// Remove all duplicate coordinates from the stand points.

```
If p_i(x_i, y_i) = p_i(x_i, y_i), i \neq j:
```

Remove  $p_i$  from  $P_S$ .

End if.

D.1. // Remove all points which are farther than outlier distance  $D_{max}$  away from other points. Loop over  $p_i \in P_S$ :

> Request one nearest neighbour  $p_{N1}$  for point  $p_i$ . // Study how far the point  $p_{N1}$  is from other points. If  $D(p_i, p_{N1}) > D_{max}$ : Remove  $p_i$  from  $P_S$ . End if.

End loop.

Create a Delaunay triangulation T for the remaining points in  $P_S$ .

Define maximum triangle edge value  $T_{max} = 35$  m. Set selected triangles  $T_{sel}$ :

 $t \in T_{sel}$  if  $edge(t, i) < T_{max}$  and  $i \in [1,3]$ .

Dissolve all selected triangles  $t \in T_{sel}$  to geometry  $G_D$ .

Buffer the dissolved polygon geometry  $G_D$  outwards by value  $B_2$  to obtain geometry  $G_{D,B}$ .

```
D.2. Check geometry type of G_{D,B}. In case of type "multipolygon", convert it to single polygons.
Add G_{D,B} to list G_{object}.
```

End loop.

// Obtain stand geometries  $G_S$  for harvested object by dissolving possibly intersecting stand and // strip road polygons by ObjectID.

If  $g_m \cap g_n$ ,  $m \neq n$  and  $m, n \in N(G_{object})$ : Set  $G_{D,IS} = g_m \cap g_n$ ,  $m \neq n$  and  $m, n \in N(G_{object})$ . Add  $G_{S,\cup} = \cup G_{D,IS}$ ,  $g \in G_{D,IS}$ , to stand geometries  $G_S$ . End if. Loop over  $g \in G_{object}$  for which  $\nexists (g_m \cap g_n, m \neq n$  and  $m, n \in N(G_{object}))$ : Add g to stand geometries  $G_S$ . End loop. D.3. Label the obtained stand delineations  $G_S$  (harvested stands) uniquely within object by StandID2.

Then stand is identified based on ObjectID and the stand label.

## Appendix E

The boundaries of overlapping adjacent polygons are balanced by Algorithm A3 as follows. The algorithm can be used to compile arbitrary adjacent polygons into a stand network, or to maintain an existing stand network with updated stand geometries. Algorithm A3: Balancing overlapping stand boundaries

Input: a set of stand delineation polygons *S* to be balanced (note that geometry type "multipolygon" needs to be converted into single polygons)

```
Put the first input stand S_1 from input stands S as such into list UpdatedStands. In case of
    previous stands existing, put only them to the list.
     Loop over rest of the input stands S_{0i}, i = [2, N(S)]: (if previous stands exist, use i = [1, N(S)])
       Initialise IScounter = 0.
E.1. Loop over updated stand polygons S_u \in UpdatedStands:
          Initialise list DP to store the division points.
         // Check whether stand polygons intersect. Ensure that if the stand S_i had previous
         // intersections, the latest version of the stand geometry is used.
         Set S_i = S_{0i}.
          If (IScounter > 0):
              Set S_i = S_{i,F}.
         End if.
         // If stand polygons do not intersect, go to next stand polygon.
         If not S_i \cap S_u:
              Go to E.1.
         End if.
         Convert S_i, S_u into boundary line geometries L_i, L_u.
          Find the intersections of lines L_i, L_u and call them vertex points P_V. Add p_V \in P_V into
         list DP.
         // See Figure A4a for example of determining the vertex points.
         Obtain buffered geometries G_V by buffering P_V with small value. Here 0.4m was used.
         Take intersection geometry G_{IS} = S_i \cap S_u and ensure that it has geometry type
          "polygon" (in case of "multipolygon", convert to single polygons).
          Select G_{IS,s} = g_{IS} \in G_{IS} such that A(g_{IS}) > A_{min}.
E.2.
          Initialise boolean variable CannotSplit = FALSE.
          Loop over intersection polygons g_{IS} \in G_{IS,s}:
              // Check how many of the vertex points p_V \in P_V intersect with g_{IS}.
              If N(p_V \cap g_{IS}) \neq 2:
                   Set CannotSplit = TRUE.
              End if.
              // Check whether the polygon g_{IS} has holes in its geometry.
              If holes exist in g_{IS}:
                   Set CannotSplit = TRUE.
              End if.
         End loop.
E.3.
         // In case of CannotSplit == TRUE, intersection G_{IS} cannot be split unambiguously.
          If CannotSplit == TRUE:
              // It is possible that the stand has been added into the updated stands, if it
              // had other intersections before the non-splittable intersection, see Stage E.5.
              // Here non-splittable stands are removed from the resulting stand network.
              If IScounter > 0:
                   Remove S_{i^*} from from list UpdatedStands where S_{i^*}(StandID2) ==
                   S_i(StandID2).
              End if.
              Set IScounter = IScounter + 1.
              Go to E.1.
          End if.
```

```
// From here on the cases CannotSplit == FALSE are handled. Start finding the
// possible empty slots between the intersecting stand polygons. For that, the stands
// cannot have holes i.e. internal rings in their geometries.
If S_i or S_u has holes:
     Fill the holes and obtain filled stand polygons S_{if} and S_{uf}.
End if.
Obtain combined geometry S_{\cup} = S_{i,f} \cup S_{u,f}.
// The empty slot polygons that lie between the intersecting stands will appear as
// internal rings in S_{\cup}.
Extract the hole polygons G_H from S_{\cup}.
Select G_{H,s} = g_H \in G_H such that A_{min} < A(g_H) < A_{max}.
Add all polygons of G_{IS,s} and G_{H,s} to G_{split} for splitting.
// Form the division line points by geoprocessing the polygon geometries G<sub>split</sub>.
Loop over g_{split} \in G_{split}:
     Convert g_{split} into its boundary lines L_g.
     Erase buffered vertices: L_{g,E} = L_g - G_V.
     Manipulate L_{g,E} to form two, continuous lines L_1 and L_2 at opposite sides of the
     polygon g<sub>split</sub>, between vertex points.
     Create equal amount N_r sample points for lines L_1, L_2 at equal fractions of their
     lengths from one vertex point p_V \in P_V onwards.
     // Points r_{i,1}, r_{i,2} \in N_r, i \in [1, N_r] form counterpart pairs, and their average determines
     // the location of the division line point. A few tens of points is a recommendable
     // amount of sample points. Lengths of L_1, L_2 can be used to adjust the amount of
     // points. See Figure A4a,b.
     Loop over i \in [1, N_r]:
          Calculate average \langle r_i \rangle from the coordinates of r_{i,1}, r_{i,2}.
          If \langle r_i \rangle \cap g_{split}:
               Insert \langle r_i \rangle into list DP.
          End if.
     End loop.
End loop.
// Examine division points DP in order to sort them.
Take a bounding box of S_i and corner points P_{BB} of bounding box.
Loop over j \in [1,4]:
     Set loop variable Point1 = P_{BB,j}.
     Set p \in DP into list AvailablePoints.
     Initialise OrderLength = 0 to measure cumulative distance of ordered DP's.
     Loop over n \in [0, N(DP)]:
          Find p_c such that distance D(Point1, p_c) = \min(D(Point1, p)) for p \in
          AvailablePoints. Set Point2 = p_c.
          Add p<sub>c</sub> into list OrderedPoints.
          If n > 0:
               Remove Point1 from AvailablePoints.
               Set OrderLength = OrderLength + D(Point1, Point2).
          End if.
          Set loop variable Point1 = Point2.
     End loop.
     Store OrderedPoints and respective values of OrderLength.
End loop.
Find order index j which has shortest OrderLength.
Create division line geometry L<sub>D</sub> from OrderedPoints of order j.
Buffer L_D with 0.3 m to obtain geometry G_L.
Split polygons G_{split} (B.II.14) by division line L_D to obtain halved geometries G_{div}.
// The split can be done e.g. using polygon-line-intersection tool in GIS.
Erase buffered line from halved geometries: G_{div,E} = G_{div} - G_L.
```

```
Buffer polygons G_{split} with 0.1 m to obtain G_{split,B}.
          Buffer the polygons G_{div,E} with 0.15 m to obtain G_{div,E,B}.
          Buffer the G_{div} (B.II.20) with 0.4 m to obtain G_{div,B}.
          Erase intersection from stand polygons: S_{i,E} = S_i - G_{split,B} and S_{u,E} = S_u - G_{split,B}.
E.4.
           // To obtain a proper new stand boundary (Figure A4d), examine which half of G_{div}
          // belongs to which stand. This is done by intersecting S_{i,E} and S_{u,E} with G_{div,E,B}.
          Loop over g_{div,E,B} \in G_{div,E,B}:
                If N(G_{div}) == N(G_{div,E,B}):
                     // The following is possible since the ordering of features remains in GIS
                     // program during geoprocessing.
                     If g_{div,E,B} \cap S_{i,E}:
                          Record that g_{div} \in G_{div}, g_{div}(id) == g_{div,E,B}(id) belongs into S_i.
                     End if.
                     If g_{div,E,B} \cap S_{u,E}:
                          Record that g_{div} \in G_{div}, g_{div}(id) == g_{div,E,B}(id) belongs into S_u.
                     End if.
                End if.
               If N(G_{div}) \neq N(G_{div,E,B}) or (g_{div,E,B} \cap S_{i,E} \text{ and } g_{div,E,B} \cap S_{u,E}):
                     If A(g_{div,E,B} \cap S_{i,E}) > A(g_{div,E,B} \cap S_{u,E}):
                          Loop over polygons g_{div} \in G_{div}:
                               If A(g_{div,E,B} \cap g_{div}) > 0.3 * A(g_{div,E,B}):
                                     Record that g_{div} belongs into S_i.
                                End if.
                          End loop.
                     End if.
                     If A(g_{div,E,B} \cap S_{i,E}) < A(g_{div,E,B} \cap S_{u,E}):
                          Loop over polygons g_{div} \in G_{div}:
                                If A(g_{div,E,B} \cap g_{div}) > 0.3 * A(g_{div,E,B}):
                                     Record that g_{div} belongs into S_u.
                                End if.
                          End loop.
                     End if.
                End if.
          End loop.
          Select small polygons G_m = g \in \{G_{IS}, G_H\} such that A(g) < A_{min}.
          // Handle g_m \in G_m in alternating manner without splitting as A(g_m) \ll A(typical)
          // stand).
          Loop over l \in [0, N(G_m)]:
                If l is even number:
                     Record that g_m \in G_m belongs to S_i.
                Else
                     Record that g_m \in G_m belongs to S_u.
                End if.
          End loop.
          List the halved and small polygons: G_F = G_{div} + G_m.
          // To obtain the resulting stand polygons with proper stand boundaries, erase g_F \in G_F
          // from the stands based on recorded information to which stand, S_i or S_u, they
          // belong.
          Obtain S_{i,F} = S_i - g_F where g_F \in G_F and g_F \in S_u.
          Obtain S_{u,F} = S_u - g_F where g_F \in G_F and g_F \in S_i.
          Change geometry S_u \rightarrow S_{u,F} in UpdatedStands.
E.5.
          // Store the resulting geometry of S<sub>i,F</sub> depending on whether this was its first
          // intersection with stands in UpdatedStands.
          If IScounter == 0:
                Add S<sub>i,F</sub> into list UpdatedStands as a new entry.
          Else:
                Find stand S_{i^*} from list UpdatedStands where S_{i^*}(StandID2) == S_i(StandID2) (it
                should be the latest entry of the list). Change geometry S_{i^*} \rightarrow S_{i,F}.
          End if.
          Set IScounter = IScounter + 1.
     End loop.
     // Check if the stand did not have intersections with updated stands.
     If IScounter == 0:
          Add S<sub>i</sub> as such into list UpdatedStands.
     End if.
   End loop.
   The list UpdatedStands contains now all the stands for which the balanced delineation could be
   formed. Those stands which cannot be split in this way with the previously updated stands, can
   be separated from Stage E.3.
```



**Figure A4.** Detailed stages of Algorithm A3. Original polygons are shown with blue and red colors, and their intersection zone with yellow color. The ETRS-TM35FIN (EPSG:3067) coordinate system is used in the figure. (a) The two black dots mark the intersections of the polygon boundary lines (vertex points). Smaller blue and red dots mark the points generated to the side lines of the intersection polygon. (b) The division points (orange) are averaged from the point pairs at the boundaries of the intersection polygon. Few grey lines show examples of the averaging. (c) The division line is created from the ordered points. (d) The intersection polygon is split with the division line and the halves are addressed to the original polygons (from which the intersection polygon was erased). The updated stand boundary is formed between the stands.

## Appendix F

The finalization stage is performed to update tree removal from stands by rearranging the tree-wise data for the forest inventory-style balanced stands obtained from Algorithm A3. The main idea is the following. Some of the external strip road trees are located within some other stand's area although they were harvested along some other harvested object(s). If that stand geometry where the tree points are located should be originated from harvester data within reasonable time window of starting times of harvest, examination of the tree-wise cutting removal data from the stand geometries would be possible. In a more general case, if a significant fraction of harvester data of entire country would be available for stand network generation, the strip road tree points of objects would most often hit some stand geometry of other object, unless they are cut from areas which are in non-forest use.

In this work, the process was tested with simple approach. The starting point were the balanced stand delineations and the remaining external strip road tree points from Algorithm A1, after the dense strip roads were identified. In order to obtain cut trees from within delineation of certain stand, the remaining strip road tree points are transferred into stand by re-labelling them with help of

intersecting the tree point geometries with the stand polygons (Figure A5). Then the trees of the stand can be gathered from the tree-wise data. As a consequence of the approach, resulting group of trees may originate from different harvested objects. In other words, growing stock of stand is originally formed by harvested object identifiers and complemented by the location of the additional tree points that obviously belong to that stand.

In current dataset, the strip road trees that were not within any stand, remained as point-wise locations from which trees were cut. The volume removal of cut trees is still available for those points for possible later use.



**Figure A5.** The trees of external strip roads (shown with diamond symbols) that locate within other stand, are transferred into that stand's tree-wise removal data. The stand points and the polygons are colored to indicate different harvested objects. The ETRS-TM35FIN (EPSG:3067) coordinate system is used in the figure.

# Appendix G

A method for detailed comparison of stand boundaries is presented below. The method studies the boundary of reference stand perpendicularly and measures the distance to the other stand's boundary (see Figure 10 in Section 2.4.). Statistics of the distances is then obtained. Direction shifts are also estimated.

#### Determining perpendicular distances between stand boundaries:

Input: two stand polygons S and  $S_R$  without inner rings

Convert the compared two stand polygons into boundary line geometries. Denote boundary of stand polygon S by  $L_0$  and boundary of the reference polygon  $S_R$  by  $L_B$ . // Sample the stand boundary of reference stand. Obtain the length of  $L_B$ . Calculate amount of samples as  $N_{SB} = int(L_B/L_{segm})$  where  $L_{segm} = 5m$ . Split  $L_B$  into lines  $L_S$  of length  $L_{segm}$ . // The split can be started *e.g.* from the point which appears first in the list of boundary // coordinates of polygon's presentation in GIS. One segment shorter than  $L_{segm}$  typically occurs // at the "end" of the boundary line. // Find the perpendicular distance values. Initialize lists Distances and NormalAngles. Loop over  $l_S \in L_S$ : Approximate  $l_s$  by a segment s between the start point  $l_{s,1}$  and end point  $l_{s,2}$  of the line  $l_{s,2}$ Obtain bearing  $\beta_S$  of  $l_{S,2}$  from  $l_{S,1}$ . Obtain normal angle  $\alpha_N = \beta_S - 90^\circ$ . If needed, convert  $\alpha_N \in$  $[0^{\circ}, 360^{\circ}]$ . Insert  $\alpha_N$  into list *NormalAngles*. Obtain center point  $s_c$  of s. Create normal lines perpendicular to *s* with help of  $\alpha_N$ . The lines are at both sides of the stand boundary:  $L_{N,1}$  outwards (angle  $\alpha_N$ ) and  $L_{N,2}$  inwards (angle  $\alpha_N + 180^\circ$ ) of the stand and have adequate length each (e.g. 80 m). // Determine the distance value at s<sub>c</sub>. Make agreement of positive and negative direction // with respect to the progress of stand boundary coordinates list. Find the possible intersection point(s)  $P_1 = L_{N,1} \cap L_0$ . If  $N(P_1) > 0$ : Take point  $p_1$  which has min $(D(p_1, s_c)), p_1 \in P_1$ . Measure  $d = D(p_1, s_c)$  and give it positive sign. Add *d* into list *Distances*. End if. Find the possible intersection point(s)  $P_2 = L_{N,2} \cap L_0$ . If  $N(P_2) > 0$ : Take point  $p_2$  which has min $(D(p_2, s_c)), p_2 \in P_2$ . Measure  $d = D(p_1, s_c)$  and give it negative sign. Add *d* into list *Distances*. End if. End loop. // Pick the distance values for different normal angles covering the main directions. This is done // to find out if the compared stand has shifted with respect to the reference stand. Initialize separate distance lists for all eight main directions. Loop over NormalAngles: Loop over main directions (N, NW, E, SE, S, SW, W, NW): Denote bearing of main direction as  $\gamma$  (e.g.  $\gamma_{SE} = 135^{\circ}$ ), and if needed, convert all angle values to  $[0^\circ, 360^\circ]$ .

If  $\alpha_N$  in  $[\gamma - 45^\circ, \gamma + 45^\circ]$ :

Insert *Distances*[index( $\alpha_N$ )] into the relevant list of main direction distances. End if.

End loop.

End loop.

Calculate mean and standard deviation of all distances. Calculate means of distances at main directions.

Calculate the net shifts at four directions N–S, NE–SW, E–W and SE–NW by subtracting the respective mean values in those directions.

## References

- 1. Holopainen, M.; Vastaranta, M.; Hyyppä, J. Outlook for the next generation's precision forestry in Finland. *Forests* **2014**, *5*, 1682–1694. [CrossRef]
- Siipilehto, J.; Lindeman, H.; Vastaranta, M.; Yu, X.; Uusitalo, J. Reliability of the predicted stand structure for clear-cut stands using optional methods: Airborne laser scanning-based methods, smartphone-based forest inventory application Trestima and pre-harvest measurement tool EMO. *Silva Fenn.* 2016, *50*, 1568. [CrossRef]
- 3. Malinen, J.; Kilpeläinen, H.; Ylisirniö, K. Description and evaluation of Prehas software for preharvest assessment of timber assortments. *Int. J. For. Eng.* **2014**, *25*, 66–74. [CrossRef]
- 4. Davis, L.S.; Johnson, K.N. Forest Management, 3rd ed.; McGraw-Hill Book Company: New York, NY, USA, 1987.
- Poso, S. Kuviottaisen arvioimismenetelmän perusteita. [Basic features of forest inventory by compartments]. Silva Fenn. 1983, 17, 313–349. [CrossRef]
- 6. Leckie, D.G.; Gougeon, F.A.; Walsworth, N.; Paradine, D. Stand delineation and composition estimation using semi-automated individual tree crown analysis. *Remote Sens. Environ.* **2003**, *85*, 355–369. [CrossRef]
- 7. Radoux, J.; Defourny, P. A quantitative assessment of boundaries in automated forest stand delineation using very high resolution imagery. *Remote Sens. Environ.* **2007**, *110*, 468–475. [CrossRef]
- Næsset, E.; Gobakken, T.; Holmgren, J.; Hyyppä, H.; Hyyppä, J.; Maltamo, M.; Nilsson, M.; Olsson, H.; Persson, Å.; Söderman, U. Laser scanning of forest resources: The Nordic experience. *Scand. J. For. Res.* 2004, 19, 482–499. [CrossRef]
- 9. Maltamo, M.; Eerikäinen, K.; Packalén, P.; Hyyppä, J. Estimation of stem volume using laser scanning-based canopy height metrics. *Forestry* **2006**, *79*, 217–229. [CrossRef]
- Holopainen, M.; Vastaranta, M.; Rasinmäki, J.; Kalliovirta, J.; Mäkinen, A.; Haapanen, R.; Melkas, T.; Yu, X.; Hyyppä, J. Uncertainty in timber assortment estimates predicted from forest inventory data. *Eur. J. For. Res.* 2010, 129, 1131–1142. [CrossRef]
- 11. Næsset, E. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sens. Environ.* **2002**, *80*, 88–99. [CrossRef]
- 12. White, J.C.; Wulder, M.A.; Varhola, A.; Vastaranta, M.; Coops, N.C.; Cook, B.D.; Pitt, D.; Woods, M. A best practices guide for generating forest inventory attributes from airborne laser scanning data using an area-based approach. *For. Chron.* **2013**, *89*, 722–723. [CrossRef]
- Valonen, M.; Haltia, E.; Horne, P.; Maidell, M.; Pynnönen, S.; Sajeva, M.; Stenman, V.; Raivio, K.; Iittainen, V.; Greis, K.; et al. Finland's Model in Utilising Forest Data. PTT Reports 261. Pellervo Economic Research PTT. 2019. Available online: https://www.metsakeskus.fi/sites/default/files/ptt-report-261-finlands-model-inutilising-forest-data.pdf (accessed on 23 June 2020).
- 14. Rasinmäki, J. Management of Multi-Scale Forest Resource Data over Time. Dissertationes Forestales 49. Ph.D. Thesis, Department of Forest Resource Management, Faculty of Agriculture and Forestry, University of Helsinki, Helsinki, Finland, 2007. Available online: http://www.metla.fi/dissertationes/df49.htm (accessed on 23 June 2020).
- Saukkola, A.; Melkas, T.; Riekki, K.; Sirparanta, S.; Peuhkurinen, J.; Holopainen, M.; Hyyppä, J.; Vastaranta, M. Predicting Forest Inventory Attributes Using Airborne Laser Scanning, Aerial Imagery, and Harvester Data. *Remote Sens.* 2019, 11, 797. [CrossRef]
- 16. Lindroos, O.; Ringdahl, O.; La Hera, P.; Hohnloser, P.; Hellström, T. Estimating the Position of the Harvester Head-a Key Step towards the Precision Forestry of the Future? *Croat. J. For. Eng.* **2015**, *36*, 147–164.
- 17. StanForD/StanForD 2010-Standard for Forest Machine Data and Communication. Skogforsk, Sweden, 2020. Available online: http://www.skogforsk.se/english/projects/stanford/ (accessed on 15 June 2020).
- 18. Bhuiyan, N.; Möller, J.J.; Hannrup, B.; Arlinger, J. Automatisk Gallringsuppföljning-Areal Beräkning Samt Registrering av Kranvinkel för Identifiering av Stickvägsträd och Beräkning av Gallringskvot. [Automatic Follow-up of Thinning-Stand Area Estimation and Use of Crane angle Data to Identify Strip Road Trees and Calculate Thinning Quotient.] Arbetsrapport 899–2016. Skogforsk, Sweden, 2016. In Swedish with Abstract in English. Available online: https://www.skogforsk.se/cd\_20190114161523/contentassets/ 6eda4307138347648171f87e3d9a3885/automatisk-gallringsuppfoljning-arealberakning-samt-registreringav-kranvinkel-for-identifiering-av-sticktrad-och-berakning-av-gallringskvot-arbetsrapport-899-2016.pdf (accessed on 23 June 2020).

- Ester, M.; Kriegel, H.-P.; Sander, J.; Xu, X. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining (KDD-96), Portland, OR, USA, 2–4 August 1996; AAAI Press: Paolo Alto, CA, USA, 1996; pp. 226–231.
- 20. Duckham, M.; Kulik, L.; Worboys, M.; Galton, A. Efficient generation of simple polygons for characterizing the shape of a set of points in the plane. *Pattern Recognit.* **2008**, *41*, 3224–3236. [CrossRef]
- Yu, X.; Hyyppä, J.; Karjalainen, M.; Nurminen, K.; Karila, K.; Vastaranta, M.; Kankare, V.; Kaartinen, H.; Holopainen, M.; Honkavaara, E.; et al. Comparison of Laser and Stereo Optical, SAR and InSAR Point Clouds from Air- and Space-Borne Sources in the Retrieval of Forest Inventory Attributes. *Remote Sens.* 2015, 7, 15933–15954. [CrossRef]
- 22. Wallenius, T.; Laamanen, R.; Peuhkurinen, J.; Mehtätalo, L.; Kangas, A. Analysing the agreement between an airborne laser scanning based forest inventory and a control inventory—A case study in the state owned forests in Finland. *Silva Fenn.* **2012**, *46*, 111–129. [CrossRef]
- 23. Laasasenaho, J. Taper Curve and Volume Functions for Pine, Spruce and Birch. Ph.D. Thesis, Communicationes Instituti Forestalis Fenniae, University of Helsinki, Helsinki, Finland, 1982.
- 24. Maanmittauksen Ortokuva. [Description of Orthoimages by National Land Survey of Finland]. In Finnish. Available online: https://www.maanmittauslaitos.fi/kartat-ja-paikkatieto/asiantuntevallekayttajalle/tuotekuvaukset/ortokuva (accessed on 23 July 2020).
- 25. Äijälä, O.; Koistinen, A.; Sved, J.; Vanhatalo, K.; Väisänen, P. (Eds.) Metsänhoidon Suositukset. [Best Practices for Sustainable Forest Management.] Tapion Julkaisuja. Tapio Oy 2019 In Finnish. Available online: https://www.metsanhoitosuositukset.fi/wp-content/uploads/2019/09/Metsanhoidon\_suositukset\_ Tapio2019\_verkko\_1.2.pdf (accessed on 24 June 2020).
- 26. Næset, E. Effects of Delineation Errors in Forest Stand Boundaries on Estimated Area and Timber Volumes. *Scand. J. For. Res.* **1999**, *14*, 558–566. [CrossRef]
- 27. Hyppänen, H.; Pasanen, K.; Saramäki, J. Päätehakkuiden kuviorajojen päivitystarkkuus. [Accuracy of updating stand boundaries at final fellings]. *Folia For.—Metsätieteen Aikakauskirja* **1996**, *4*, 321–335. [CrossRef]
- 28. Courteau, J.; Darche, M.-H. *A Comparison of Seven GPS Units under Forest Conditions*; Special report 120; Forest Engineering Research Institute of Canada: Vancouver, BC, Canada, 1997.
- 29. Darche, M.-H. *A Comparison of Four New GPS Systems under Forest Conditions*; Special report 128; Forest Engineering Research Institute of Canada: Vancouver, BC, Canada, 1998.
- 30. Næsset, E. Effects of differential single- and dual-frequency GPS and GLONASS observations on point accuracy under forest canopies. *Photogramm. Eng. Remote Sens.* **2001**, *67*, 1021–1026.
- 31. Meneguzzo, D.M.; Liknes, G.C.; Nelson, M.D. Mapping trees outside forests using high-resolution aerial imagery: A comparison of pixel- and object-based classification approaches. *Environ. Monit. Assess.* 2013, *185*, 6261–6275. [CrossRef]
- 32. Gasparovic, M.; Dobrinic, D.; Medak, D. Geometric accuracy improvement of WorldView-2 imagery using freely available DEM data. *Photogramm. Rec.* **2019**, *34*, 266–281. [CrossRef]
- 33. Hughes, M.L.; McDowell, P.F.; Marcus, W.A. Accuracy assessment of georectified aerial photographs: Implications for measuring lateral channel movement in a GIS. *Geomorphology* **2006**, *74*, 1–16. [CrossRef]
- 34. Topographic Database, National Land Survey of Finland. Available online: https://www.maanmittauslaitos.fi/ en/maps-and-spatial-data/expert-users/product-descriptions/topographic-database (accessed on 23 July 2020).
- 35. Valbuena, R.; Mauro, F.; Rodriguez-Solano, R.; Manzanera, J.A. Accuracy and precision of GPS receivers under forest canopies in a mountainous environment. *Span. J. Agric. Res.* **2010**, *8*, 1047–1057. [CrossRef]
- Ordóñez Galán, C.; Rodríguez-Pérez, J.; Martínez, J.; Garcia Nieto, P.J. Analysis of the influence of forest environments on the accuracy of GPS measurements by using genetic algorithms. *Math. Comput. Model.* 2011, 54, 1829–1834. [CrossRef]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).