Taylor & Francis
Taylor & Francis Group

AUTOMATIC MAPPING OF FOREST DENSITY FROM AIRBORNE LIDAR DATA

Øivind Due TRIER

Norwegian Computing Center, Section for Earth Observation, P.O. Box 114 Blindern,
NO-0316 Oslo, Norway
E-mail: trier@nr.no

Abstract. This paper presents new methods for the automatic mapping of vegetation from airborne lidar data. The methods are developed specifically for orienteering maps, which are detailed maps in scale 1:15,000 or 1:10,000 of forested areas. However, the methods may be modified to be used for automatic mapping of vegetation for national topographic map series in various scales, e.g., 1:25,000 or 1:50,000.

We introduce the normalized difference vegetation density (NDVD) as an indicator of vegetation density in airborne lidar data. A modified version of NDVD is used for reduced runability mapping.

By comparing pixel-by-pixel the automatic mapping with the manual survey in four different forest areas in Oslo, Norway, the correct classification rate varies from 71% to 75%. However, close investigation reveals that the automatic mapping is better than manual survey for open areas. On the other hand, the automatic mapping of reduced runability remains a difficult problem. In many cases, the automatic method is able to identify the major areas of reduced runability, while in other areas the correspondence is low between the automatic mapping and manual survey of reduced runability. Still, the automatic method may be used to quickly produce an initial mapping of reduced runability, or in the production of orienteering maps in remote areas where a full manual survey cannot be afforded.

Keywords: orienteering map, airborne laser scanning, forest structure, normalized difference vegetation density, open land, map generalisation, remote sensing.

Introduction

Airborne lidar data, also called airborne laser scanning (ALS) data, are now being collected and used for a number of purposes. In order to fully utilize these datasets in a cost-effective manner, automatic mapping methods are needed.

One major application area is forestry, and methods exist for estimating mean tree height and timber volume (e.g., Næsset 1997a, 1997b), and detection of individual trees (e.g., Kwak *et al.* 2007; Ene *et al.* 2012). The methods are based on the fact that one individual laser pulse may result in multiple returns, typically several returns from different parts of a tree canopy, and if the tree vegetation is not too dense, one return from the ground. By constructing a triangulated irregular network (TIN) of all the ground points and converting it to a digital terrain model (DTM), a ground elevation may be estimated for every lidar vegetation point. Thus, for each vegetation point, the height above the ground may be estimated directly as the difference. Smith *et al.* (2009) compared estimates of forest canopy cover from multispectral optical data, from lidar data, and field measurements, and found that the lidar estimates were reasonably close to the field measurements, and much more accurate than the multispectral optical estimates. Smreček and Michňová (2014) detected individual trees and groups of trees automatically in lidar data with > 20 points/m^2.

Another important application area is urban mapping, including three-dimensional building extraction (Dorninger, Pfeifer 2008), automatic road mapping (Clode *et al.* 2007) and land cover classification (Yan *et al.* 2015).

An important aspect of automatic mapping is generalisation; this usually means to produce a less detailed map in small scale from an existing, more detailed

map in large scale. However, aspects of generalisation may also be used in the mapping process, as the source data may have more detail than the finished map. Papšienė and Papšys (2011) discuss various generalisation methods, and note that a cluster of point features at a large scale may be aggregated to an area feature at a smaller scale. Recently, a fully automatic workflow has been developed for generalizing 1:50,000 topographic vector maps from 1:10,000 maps in the Netherlands (Stoter *et al.* 2014). Tveite, Langaas (1999) developed a method to evaluate the accuracy of generalised line data sets.

The aim of this paper is to estimate forest density as perceived by a person running through the forest, outside of paths. This is an important aspect of orienteering maps, where normal runability is mapped as white areas, and three classes of increasingly reduced runability are mapped as light green, middle green and dark green. Open areas are mapped as yellow, and open areas with scattered trees are mapped with yellow with white dots in a regular pattern. Orienteering maps are used in orienteering competitions, where the runners have to visit controls, marked as purple circles on the map, in a prescribed order and as fast as possible. Between the controls, each runner may make individual route choices, often a combination of following paths and crossing through the forest outside of paths. The surveying of orienteering maps is currently done by walking through the terrain, mapping point features (e.g., distinct boulders), line features (e.g., paths) and area features (e.g., marshes), and adjusting the contour lines. Mapping of reduced runability is the most challenging and subjective task. So the potential benefits of an automatic method are to eliminate the subjective component and save time.

Automatic mapping of reduced runability from ALS data may be feasible, given that it is possible to measure the forest density, as perceived by a person trying to run through the forest outside of paths. The ALS pulses may be reflected from different parts of a tree. ALS pulse returns in the interval 0–2 m above the ground may be a direct indication that there are obstacles at those points. Pulse returns from higher than 2 m above the ground may or may not be correlated with forest runability, depending on the shape of the tree canopy. For Norway spruce (*Picea abies*), the dominating tree species in the study areas, the young trees have canopies all the way down to the ground, with the longest branches closest to the ground. In closed forest, mature trees typically have few or no branches at the bottom few meters of their height. However, in open areas, a single, mature Norway spruce tree has a similar canopy shape as young trees.

The shape and internal density of the canopies, as well as the fractional canopy cover, influences the proportion of ALS pulses that obtain ground returns. Norway spruce forest usually allows some pulses to hit the ground, and thus, also allows some pulses to hit the lower parts of the canopies. On the other hand, deciduous trees, in the leaf on period, may block the pulses high up in the canopies.

1. Data

1.1. Lidar data

ALS data was collected by TerraTec AS, Oslo, Norway. For the entire forested area of Oslo municipality (except forest patches in urban parks), data was collected with two emitted pulses per m^2. In addition, for the parts of the forest that are owned by Oslo City, ALS data was collected with 10 emitted pulses per m^2 (Table 1). The raw point data was automatically classified, visually inspected, edited and re-classified by TerraTec AS (Table 2). The data was delivered as LAS files (ASPRS 2010), with each (x, y, z) point labelled as 'ground' or 'other'. Each (x, y, z) point also has an intensity value, and a return number (1, 2, 3 or 4). So, each emitted pulse may result in up to four discrete returns. We have received some subsets of this data for the purpose of making orienteering maps.

Table 1. Laser scanning acquisition parameters

Pulse density	2 pulses / m^2	10 pulses / m^2
Acquisition dates	19 and 28 September 2010	8, 14, 15, 20 and 22 August 2010 1 September 2010
Laser scanner	ALS50-II, serial no 099	ALS50-II, serial no 099
Platform	FW, PA31, SE-LVF	RW, Eurocopter AS350, LN-OSL
Height above terrain	2000 m	250 m and 450 m
Velocity	135 kn	70 kn
Laser pulse rate	117 900 Hz	150 000 Hz and 105 900 Hz
Field of view	15 degrees	16 degrees
Strip overlap	20%	20% and 60%

Table 2. Software used for processing of ALS data

Software	Vendor	Functionality
GeoDig	TerraTec	Flight planning
AeroPlan Mission Planning	Leica Geosystems	Flight planning
TerraPOS	TerraTec	GNSS processing
IPAS Pro	Leica Geosystems	GNSS/INS integration
ALSPP	Leica Geosystems	Derive raw point clouds
WSKTrans	Statens Kartverk	Geoid
TerraScan	TerraSolid	Processing, visualization, editing and classification of laser point data
TerraMatch	TerraSolid	Sensor and stripe corrections
Terra-Modeler	TerraSolid	Visualization of terrain surface
MicroStation	Bentley	CAD platform for T erraSolid software

1.2. GIS data

For the same areas, we have received GIS data, including water polygons, buildings, roads, etc.

1.3. Field data

Base maps for field survey (Fig. 1) were produced as follows. Lakes, rivers, roads, buildings, power lines, etc. were imported from ESRI® Shape files. Contour lines with 1 m elevation interval and index contours at 5 m and 25 m intervals were generated from the ALS point clouds. Only the points labelled as 'ground' were included. The ENVI® routines TRIANGULATE and CONTOUR were used to generate vector contour lines. TRIGRID was used to produce a digital terrain model (DTM) of the 'ground' points. Next, TRIGRID was used to produce a digital surface model (DSM) of all the first returns (including both 'ground' and 'other' points). To compute the vegetation height image (often called normalized DSM or nDSM) the DTM was subtracted from the DSM. The vegetation height was color-coded (Table 3).
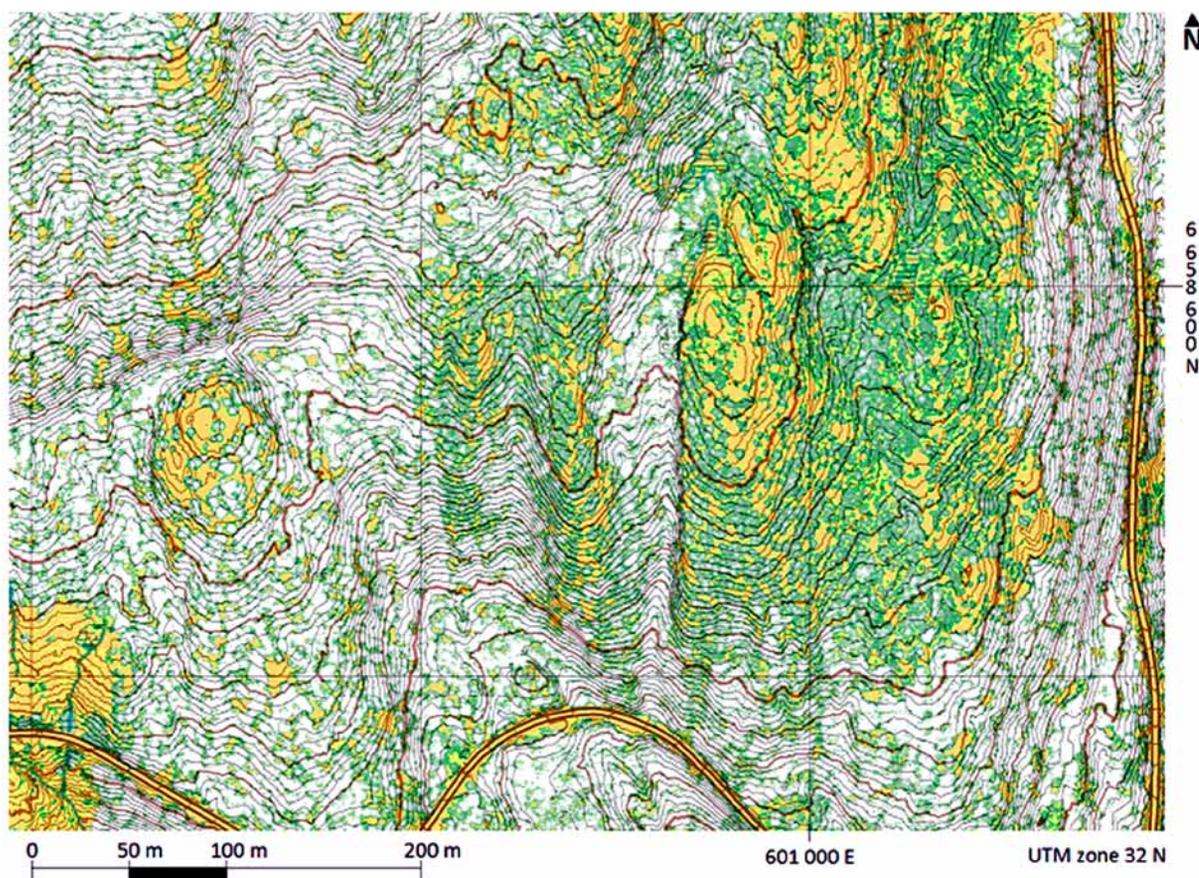
Fig. 1. Base map for field survey at Tømte. Vegetation height is color-coded (yellow-green-white; see Table 3). Grey lines = elevation contour lines with 1 m height interval and no smoothing. Thin and thick brown lines are index contour lines at 5 m and 25 m height intervals, respectively. The parallel grey lines with yellow in the middle are forest roads

Table 3. Colour codes for vegetation height in base maps

Vegetation height	Inter-preta-tion	Colour	Red	Green	Blue
0.0 m < h ≤ 0.1 m	no vege-tation	Dark yellow	255	191	0
0.1 m < h ≤ 1.0 m	low vege-tation	Yellow	255	255	0
1.0 m < h ≤ 2.0 m	small trees	Bright green	0	255	0
2.0 m < h ≤ 12.0 m	medium to tall trees	Dark green to white	0 ... 255	127 ... 255	0 ... 255
12.0 m < h	very tall trees	White	255	255	255

Table 4. The four areas surveyed in 2013. Tryvann is covered by two datasets with different point densities

Bergendal	3.3 km²	9–16 May 2013	0.95
Bjørnholt	2.6 km²	13–23 July 2013	0.84
Tømte	4.6 km²	16–26 May 2013	0.84
Tryvann	2.8 km²	23 July – 3 August 2013	0.74 | 3.17

Four separate areas were surveyed in 2013 (Table 4, Fig. 2), using similar base maps as in Fig. 1.

2. Methods

This section describes the motivation for the different processing steps in the method, and summarizes the method at the end of the section. The main processing steps are:

– Mapping of open land;
– Mapping of reduced runability.

Mapping of open land is quite straightforward, this is essentially to map areas with vegetation height below a threshold, and then to apply some generalisation. Mapping of reduced runability is more challenging, as the signal may be quite weak in some situations. We may assume that reduced runability is related to the vegetation density between 0 and 2 m above the ground. However, if the tree canopies are very dense above 2 m above the ground, then very few lidar pulses may reach the 0–2 m vegetation height interval. So, we need to compensate for the possibly weak signal.

When analysing ALS data, one may either work directly on the (x, y, z) point cloud, or convert the ALS data to images. In order to exploit the capabilities provided by image analysis in terms of convolution and morphological operators, we converted the ALS data
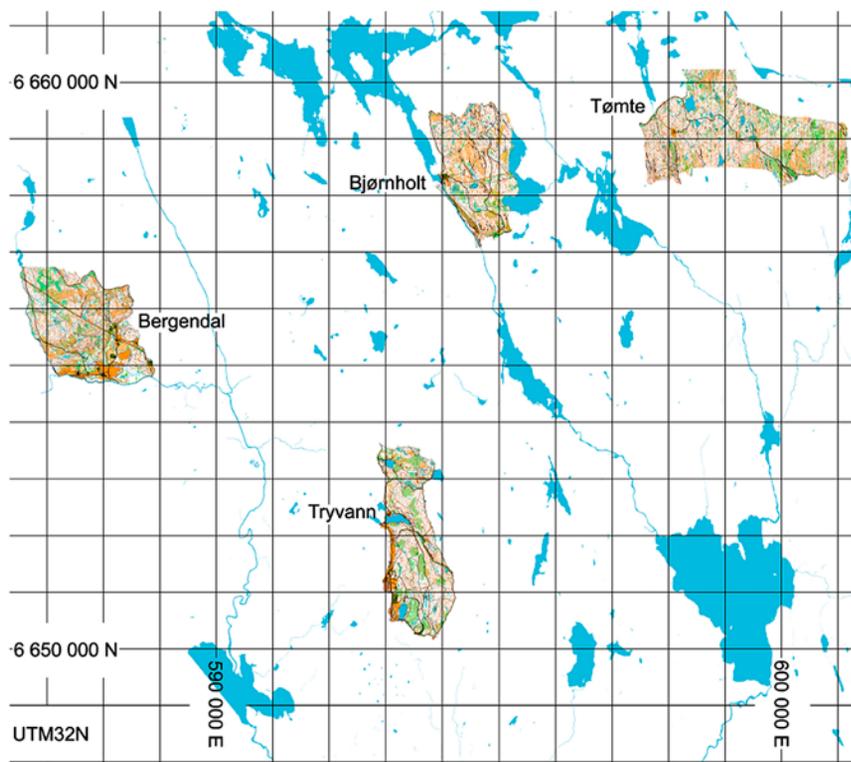
Fig. 2. Location of areas surveyed in 2013. The grid lines are 1 km apart and the coordinates are in meters in UTM zone 32 N, datum = Euref-89

to images. With two emitted pulses per m^2, 0.5 m pixel size was considered adequate to preserve the detail of the lidar data.

In order to measure the density of the vegetation, we propose a quantity that we call the normalized difference vegetation density (NDVD):

$$NDVD = \frac{V - G}{V + G}.$$ (1)

Here, *V* and *G* are the number of vegetation and ground hits, respectively, within a small neighbourhood around each location, and for *V*, within a height interval. For our purpose, we will need to determine three parameters for computing *NDVD*:

– The minimum height above the ground for a vegetation hit, ideally = 0.0 m;

– The maximum height above the ground for a vegetation hit, ideally = 2.0 m;

– The size of the neighbourhood should ideally be large enough to include several ground hits.

Based on trial and error, we use a circular neighbourhood with radius = 2.0 m, giving equal weight to all hits within a 1,0 m radius from the centre, and linearly decreasing weight from 1.0 m to 2.0 m from the centre (Fig. 3).

The NDVD of the 0.0–2.0 m above ground vegetation returns contains some artefacts in the form of clearly visible stripes (Fig. 4), related to overlaps of data from different flight strips. These artefacts were reduced by using the returns from 0.2–2.0 m above ground.

2.1. Open land

Mapping of open land, i.e., areas with no tree vegetation, may be done directly from the vegetation height. However, to make the map more readable, some generalization is needed. Small and/or thin details must either be exaggerated or removed. Morphological operators (Soille 2002) are ideal for automatic exaggeration and/or removal of details in an image. In The international specification for orienteering maps (ISOM) (Zentai 2000), there are two area symbols for open land: '401 open land' for grassland etc. with easy running, to be mapped using yellow full colour, and '403 rough open land', with rough ground vegetation, and runability similar to open runnable forest, to be mapped using yellow halftone (Fig. 5). ISOM specifies the smallest area of colour, when printed in halftone, to be 1 mm^2 in scale 1:15 000, which is 225 m^2 in the terrain. For full colour, the smallest area is 0.5 mm^2, which is 112.5 m^2 in the terrain.
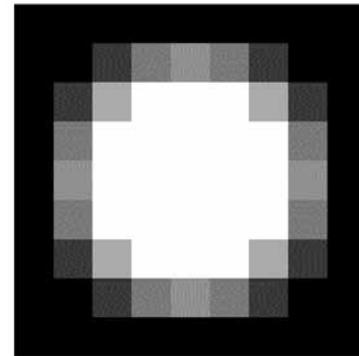
Fig. 3. Neighborhood kernel. Pixel size = 0.5 m. Inner radius = 1.0 m, outer radius = 2.0 m

a)   b)
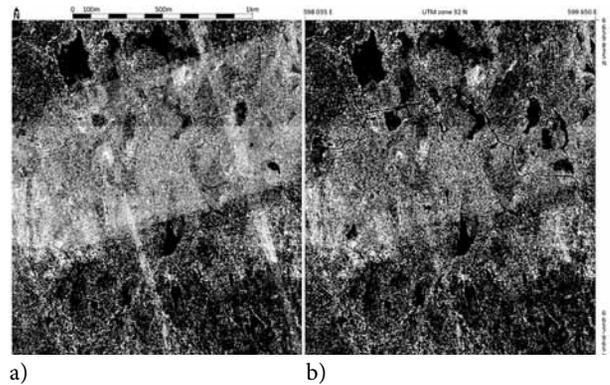
Fig. 4. NDVD of a part of the forest at Tømte. (a) NDVD for vegetation returns 0.0–2.0 m above the ground. (b) Similarly, for 0.2–2.0 m above the ground

401 open land
402 open land with scattered trees
403 rough open land
404 rough open land with scattered trees
405 forest: easy running
406 forest: slow running
407 undergrowth: slow running
408 forest: walk
409 undergrowth: walk
410 vegetation: fight
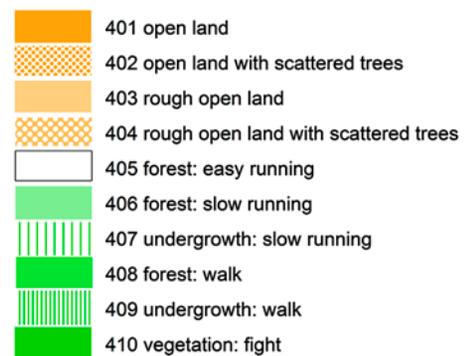
Fig. 5. Legend for vegetation mapping

Mapping of open land with scattered trees is more challenging. Again, there are two different area symbols for this in ISOM: '402 open land with scattered trees' with grass etc. which allows easy running, and '404 rough open land with scattered trees', with rough ground vegetation, and runability similar to open runnable forest. Symbol '402 open land with scattered trees' is mapped with a coarse yellow dot raster, and the smallest area allowed is 10 mm^2 in scale 1:15,000,

which is 2,250 m² in the terrain. Symbol '404 rough open land with scattered trees' is mapped using an area symbol with yellow halftone, similar to the one used for rough open land, but with additional large white dots to symbolize the presence of scattered trees. The smallest area allowed with this symbol is 16 mm² in scale 1:15,000, which is 3,600 m² in the terrain. As an alternative, each cluster of trees may be surveyed as a white small area inside an area of rough open land. For manual surveying, the latter alternative would be very time-consuming, but for an automatic method, this alternative is much more straightforward than to try to construct rules for rough open land with scattered trees. So, one automatic method will be used for all varieties of open or rough open land, with or without scattered trees.

Mapping of open areas is done by using the following steps (Fig. 6):

- Aggregate the vegetation height image to 1.0 m pixel size (Fig. 6a);
- Threshold this image at 0.75 m vegetation height, labelling vegetation higher than 0.75 m as 'tree vegetation' (white in Fig. 6b), and areas with vegetation lower than 0.75 m as 'open land' (yellow in Fig. 6b);
- Remove small patches and thin strips of 'open land' by using binary morphological opening with a 3×3 pixels disk kernel (Fig. 6c);
- Exaggerate small trees:
  - Remove single white pixels (Fig. 6d), this means that very small trees are ignored;
  - Remove white objects smaller than 6 pixels (Fig. 6e);
  - Extract small trees as the difference of this (Fig. 6e) and the previous (Fig. 6d). In the difference image (Fig. 6f), small trees are dark green;
  - Enlarge the small trees by using binary morphological dilation with a 3×3 square (Fig. 6g);
  - Do morphological closing with a 3×3 square on the enlarged small trees (Fig. 6h);
  - Subtract the small trees (Fig. 6h) from the open areas (Fig. 6d). In this way, no (white) vegetation holes in the (yellow) open areas are less than 5 pixels (Fig. 6i);
- Remove small patches and thin strips of 'open land' by using binary morphological opening with a 3×3 pixels disk kernel (Fig. 6j);
- Remove (yellow) open areas smaller than 22.5 m² (Fig. 6k). This image is called the 'open areas mask', and will be used in the method for

mapping reduced runability;
- Remove (yellow) open areas smaller than 225 m² (Fig. 6l).

2.2. Reduced runability

The mapping of reduced running speed is done using three classes:

- Light green, 'slow run', 20–50% reduced running speed;
- Middle green, 'walk', 50–80% reduced running speed;
- Dark green, 'fight', 80–100% reduced running speed.

The same method is used for all three classes. The only difference is the threshold on NDVD of the 0.2 m to 2.0 m vegetation layer. The following steps are used:

- Aggregate the NDVD of the 0.2 m to 2.0 m vegetation height, to 1.0 m resolution (Fig. 7a);
- Threshold this (Fig. 7a) using the appropriate threshold value (Table 5). The resulting image (Fig. 7b) of dense vegetation (in the 0.2–2.0 m vegetation height interval) is a weak signal, which will be exaggerated through the next steps;
- Use binary morphological closing with a 7×7 disk kernel (Fig. 7c);
- Use binary morphological opening with a 3×3 disk kernel (Fig. 7d);
- Use binary morphological closing with a 9×9 disk kernel (Fig. 7e).
- Use binary morphological opening with a 5×5 disk kernel (Fig. 7f).
- Use binary morphological closing with a 11×11 disk kernel (Fig. 7g);
- Use binary morphological opening with a 7×7 disk kernel (Fig. 7h);
- Remove the 'open areas mask' (Fig. 6k) from

Table 5. Threshold values used in the estimation of forest runability from the NDVD of the 0.2 to 2.0 m vegetation layer

Class	Reduced running speed	ALS pulse density per m²	Threshold for NDVD 0.2–2.0m
Slow run	20–50%	2	–0.2
Walk	50–80%	2	0.20
Fight	80–100%	2	0.60
Slow run	20–50%	10	0.00
Walk	50–80%	10	0.35
Fight	80–100%	10	0.70
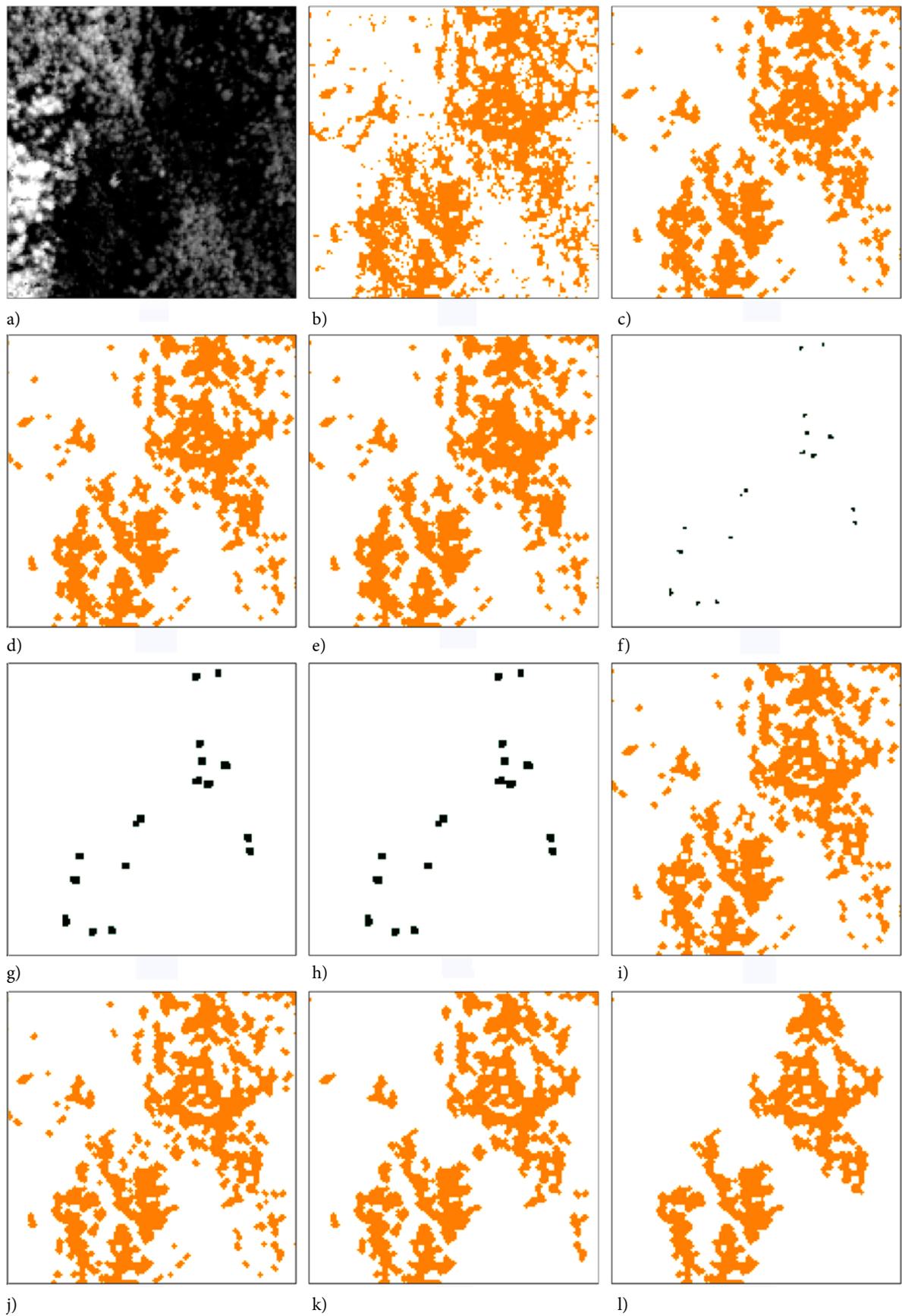
Fig. 6. Results of each processing step in the mapping of areas with no tree vegetation.
Pixel size = 1.5 m, image portion size = 225 m × 225 m. For the explanation of each step (a-l), see Section 2.1
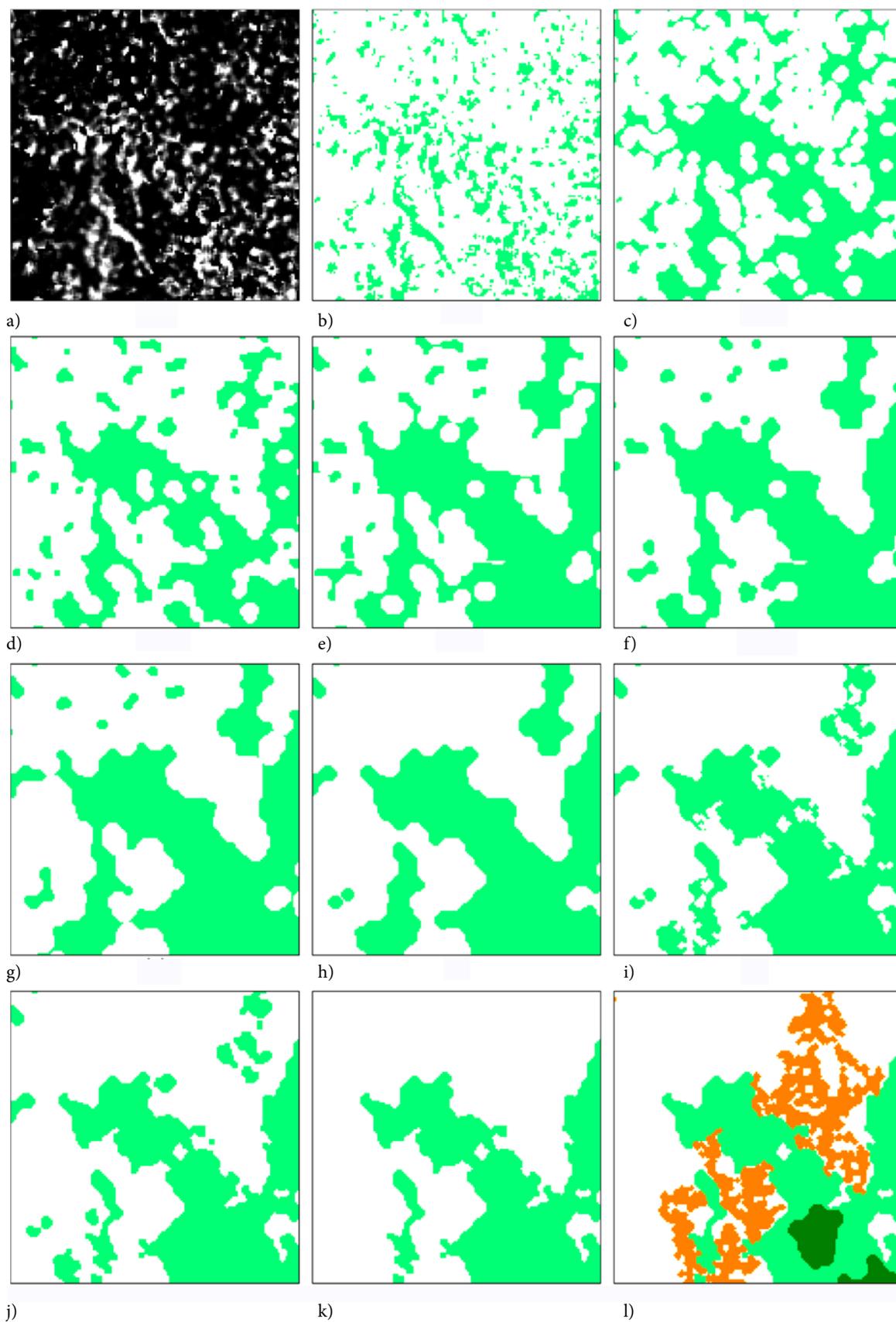
Fig. 7. Results of processing steps in the mapping of areas with reduced runability due to forest tree vegetation (a–k). The combined mapping of open areas and reduced runability (l). Pixel size = 1.5 m, image portion size = 225 m × 225 m. For the explanation of each step (a–l), see Sections 2.2 and 2.3

the dense vegetation (Fig. 7h) to obtain a masked dense vegetation (Fig. 7i);

– Remove small patches and thin strips of dense vegetation by using binary morphological opening with a 3×3 pixels square kernel (Fig. 7j);

– Remove dense vegetation objects smaller than the smallest allowed area (Fig. 7k). For 'slow run' and 'walk', this is 225 m^2, and for 'fight', this is 112.5 m^2.

Morphological closing extracts areas with relatively high vegetation density by filling in the space between areas of dense vegetation that area closer than the kernel diameter, but without removing anything. Therefore, morphological closing is followed by morphological opening, which removes objects and parts of objects that are thinner than the kernel diameter, and keeps everything else. If a too large kernel is used in morphological closing, then areas with quite low vegetation density area also included. The solution is to use gradually increasing kernels iteratively (steps 3–8 above).

2.3. Combine maps

To arrive at the final vegetation map, combine the results from the three dense vegetation classes and the open areas (Fig. 7l).

2.4. Parameter values

There are a number of parameters that have to be set in the method. We have investigated the effect of varying them, one at a time. Several of the parameters affect the amount of area mapped as dense vegetation, but this may be reversed by adjusting the NDVD thresholds.

Of particular interest is the minimum size of open areas. The proposed method uses a minimum area of 225 m^2 for halftone areas, to provide a more readable and generalized result, and to adhere to the ISOM. By reducing this limit, many small open areas appear, but only a few more areas of reduced runability appear (Fig. 8). The motivation for reducing this limit is to be able to map open areas with scattered trees more in line with current practice in manual survey.

3. Results

The method is applied on four areas (Table 4). All four areas are covered by the 2/m^2 dataset. Only the Tryvann area is in addition covered by the 10/m^2 dataset.

By comparing pixel-by-pixel the automatic classification with the manual survey, the correct classification rate varies from 71% to 75% (Table 6, Table 7). This is actually not as good as it sounds. If we simply assigned 'forest: easy running' (which is also called 'normal' forest) to all pixels, then the correct classification rates would be from 54.5% to 66% (Table 8). So, the improvement in classification rates by using the new method is 7% to 20% of the total number of pixels, compared to assigning 'forest, easy running' to all pixels.

Considering one class at a time, the correct classification rate for 'open' varies from 81% to 88%, 'normal' from 76% to 80%, 'slow run' from 22% to 52%, 'walk' from 9% to 40%, and 'fight' from 0% to 4% (Table 7). Based on this, it is tempting to use only one class for reduced runability, in which case the correct classification rate for reduced runability varies from 26% to 60% (Table 7).
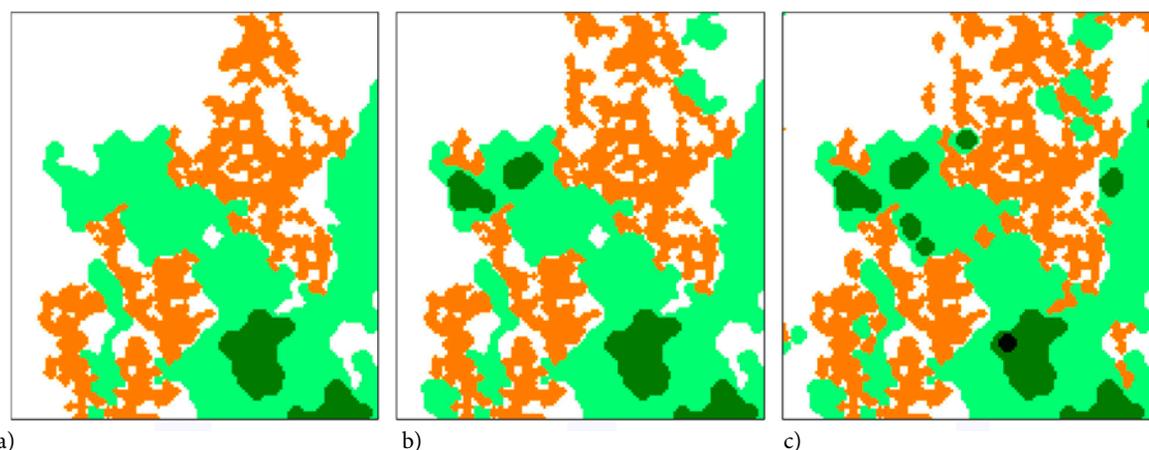
Fig. 8. The effect of reducing the minimum size of half tone areas. (a) Reference result: minimum half tone area is 225 m^2. (b) Minimum half tone area is reduced to 112.5 m^2. (c) Minimum half tone area is reduced to 22.5 m^2

Table 6. Confusion matrix and classification rates for the Tømte area

Tømte normal		Manual survey						
		open	slow-run	walk	fight	total	correct %	
Automatic classification	normal	2423503	393924	264708	14477	387	3096999	78.25%
	open	96610	541327	19362	2922	40	660261	81.99%
	slow-run	388824	64618	256565	36427	254	746688	34.36%
	walk	6909	3615	31624	20760	116	63024	32.94%
	fight	0	5	5	919	0	929	0.00%
	total	2915846	1003489	572264	75505	797	4567901	70.98%
	correct %	83.11%	53.94%	44.83%	27.49%	0.00%	70.98%	

Table 7. Classification rates

		$2/m^2$ data				$10/m^2$ data
		Bjørnholt	Bergendal	Tømte	Tryvann	Tryvann
Automatic classification	normal	79.87%	76.62%	78.25%	80.18%	82.38%
	open	79.62%	88.36%	81.99%	81.58%	82.79%
	slow-run	22.04%	31.69%	34.36%	52.43%	47.56%
	walk	9.54%	40.23%	32.94%	10.07%	8.99%
	fight	0.00%	3.76%	0.00%	0.00%	0.00%
	total	74.35%	73.89%	70.98%	75.02%	73.85%
Combined class 'dense'		25.69%	48.35%	42.76%	59.98%	57.05%

However, some generalization is made in the manual survey, and also in the automatic method, so some pixel-by-pixel differences are expected. When comparing the results of manual and automatic mapping, it may be more meaningful to check if the general distribution of open areas and reduced runability is captured, instead of focusing on pixel-by-pixel differences.

For the Tømte area, the agreement is very good regarding open areas. The main difference is that a large number of small open areas, smaller than the minimum size allowed by ISOM, have been included in the manual mapping. Also, semi-open areas have been mapped by the automatic method by inserting holes in open areas, each hole having the exact location and shape of individual trees or groups of trees. There are some examples where the manual survey has mapped semi-open areas larger than the automatic method has. This is typically in old clear cuts with young emerging trees. It is a matter of definition: when does an open area with scattered trees have too many trees? The ISOM states that open areas with scattered trees may be mapped by either (1) surveying small areas of white or green, or (2) using a generalised raster with regularly spaced white dots. The automatic method has been designed to use the first alternative. So, when the automatic method does not map an area as open with

scattered trees in this way, this may be due to one of three reasons: (1) the automatic method makes mistakes, (2) the scattered trees are too many for the area to be mapped as open with scattered trees, or (3) trees have been cut between the time of the lidar acquisition and the field survey. By comparing some portions of the mapping results with air photos acquired in the summer of 2013 (Fig. 9), we see that there are in fact many small trees in the areas that were mapped as forest by the automatic method but open with scattered trees by the manual survey. So, there has been no cutting of trees between the lidar acquisition and the manual survey for these selected portions. Field inspection on 13 September 2014 confirms this: although some thinning has been done, possibly in 2013, the areas do not appear as open with scattered trees.

Another issue is the large number of small open areas mapped by the manual survey. By changing the lower bound for open areas in the automatic method from 225 m^2 to 22.5 m^2 (Fig. 9f), a number of small open areas are added to the automatic mapping result. Most of these are within areas mapped as either open or open with scattered trees by the manual survey. However, many of the small open areas mapped by the manual survey are not mapped as open by the automatic method. Also, large parts of the areas mapped

by manual survey as open with scattered trees within the two rectangles are still mapped as forest by the automatic method. All in all, we may conclude that the manual survey has included too much as open or open with scattered trees, and that the automatic mapping of open areas is more accurate than the manual method.

For the Bergendal area, open areas are also mapped very well by the automatic method. However, there are some quite large open areas that have been missed by the automatic method, due to clear cuts between the time of the lidar acquisition in 2010 and the field survey in 2013 (Fig. 10).
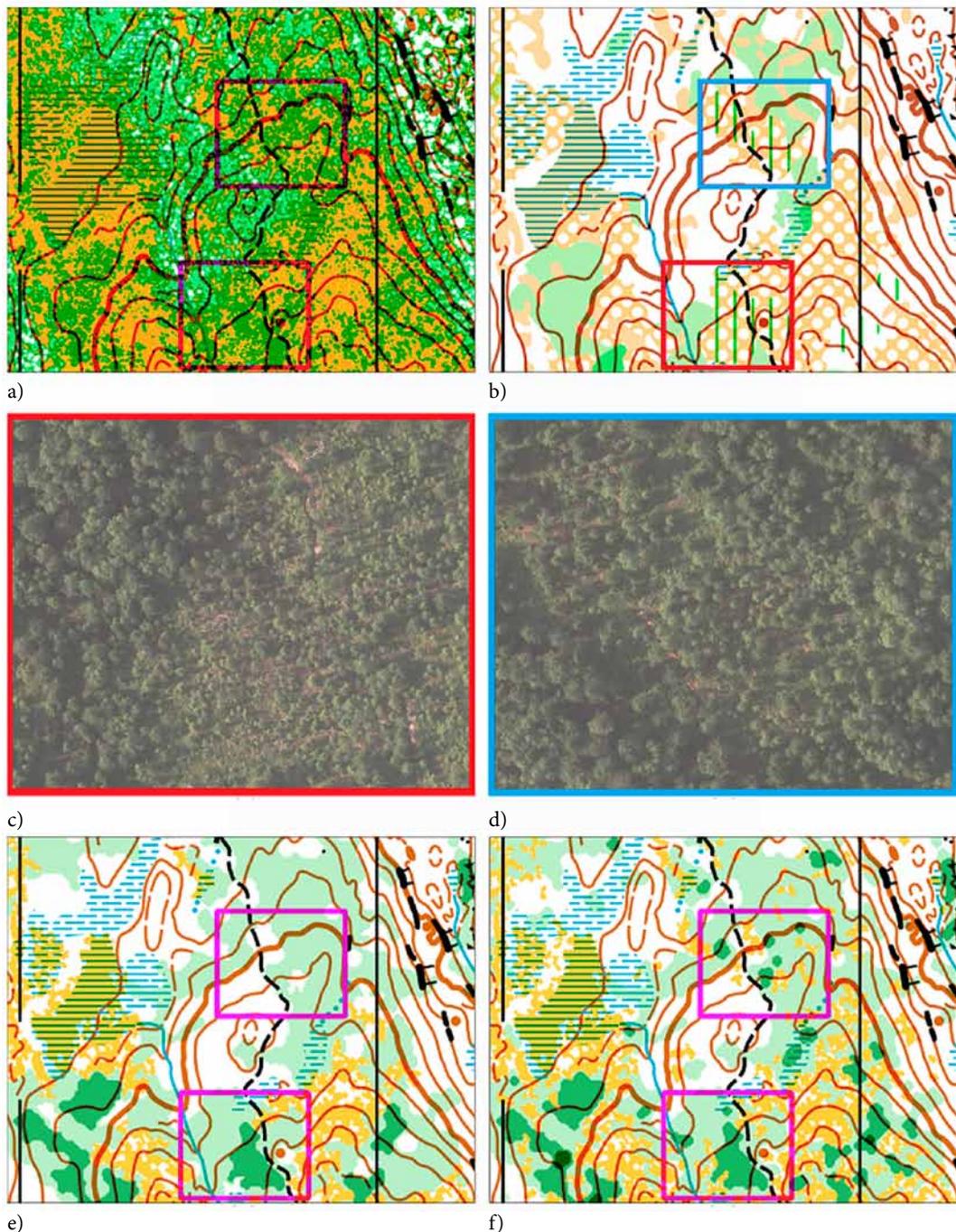
Fig. 9. Inspection of mapping results for a small portion of the Tømte area. (a) Vegetation height: dark yellow = 0.0–0.1 m, light yellow 0.1–1.0 m, dark green = 1–7 m, light green = 7–12 m, white > 12 m. (b) Manual survey. (c–d) Small portions of air photo of summer 2013, for the (c) southern rectangle and the (d) northern rectangle. (e) Automatic vegetation mapping, with 225 m^2 minimum size of half-tome areas. (f) As (e), but with 22.5 m^2 limit

a)

b)
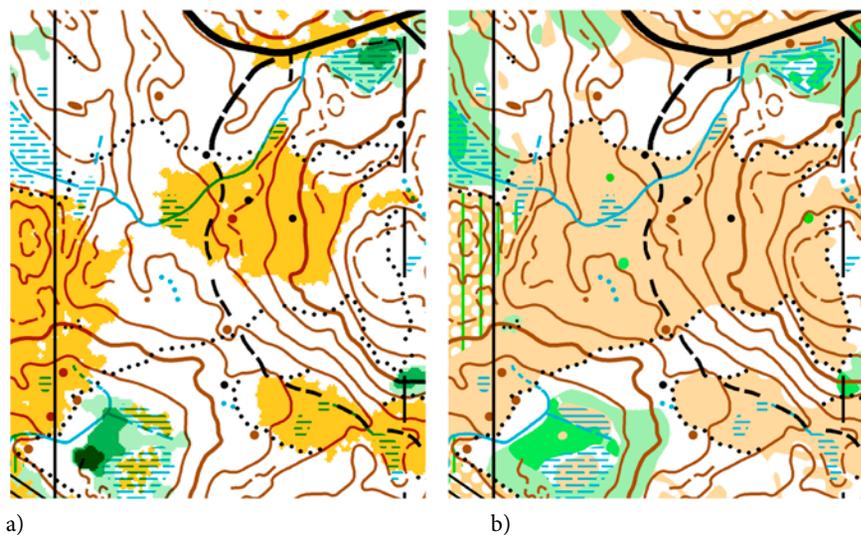
Fig. 10. New clear cuts have been made between the lidar acquisition (2010) and the manual survey (2013). (a) The new clear cuts are missed by the automatic method. (b) They are mapped by the manual survey

a)

b)

c)

d)
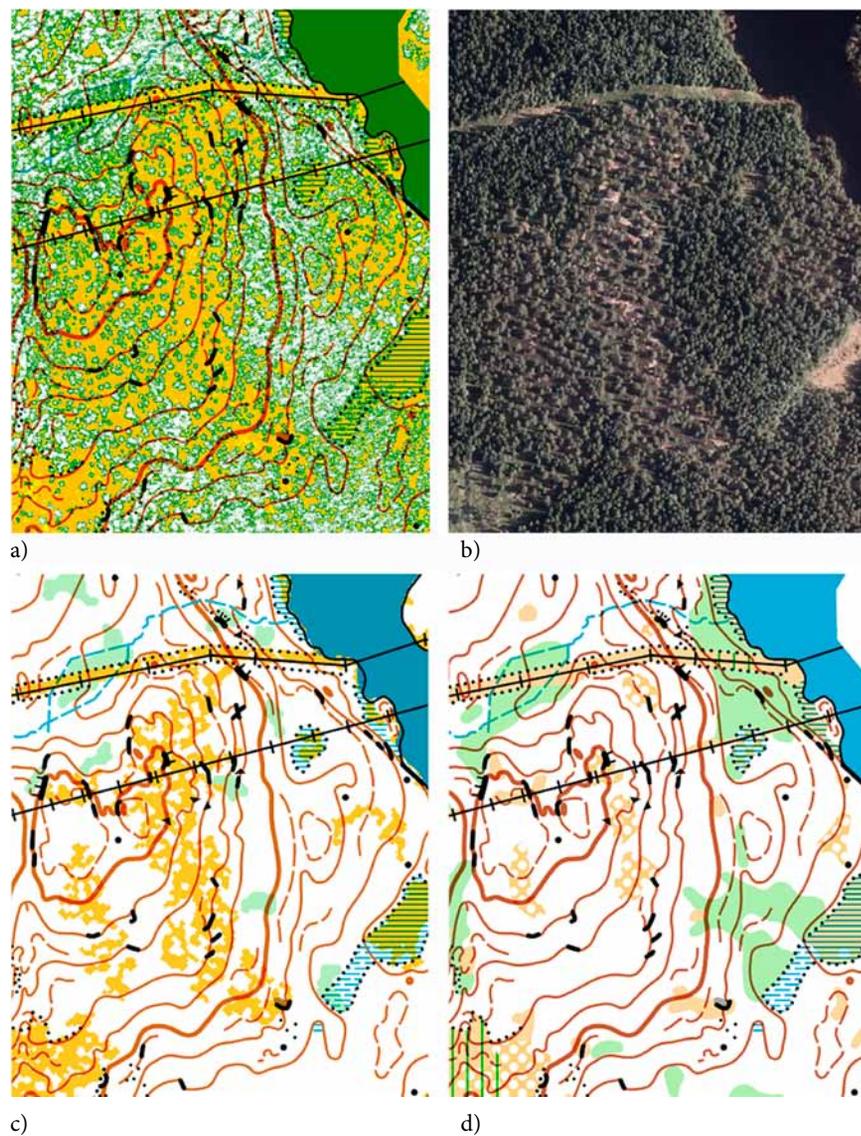
Fig. 11. The automatic method has mapped a larger area as 'open with scattered trees' than the manual survey. (a) Colour-coded vegetation height. (b) Air photo of 2011. (c) Automatic vegetation mapping. (d) Manual survey

For the Bjørnholt area, some of the open areas with scattered trees are mapped as larger by the automatic method than by manual survey. So, when the automatic method maps larger areas than manual survey as open, this could be for one of three reasons: (1) rapid regrowth between the time of lidar acquisition (2010) and manual survey (2013), (2) the automatic method makes mistakes, or (3) the scattered trees are too many for the area to be perceived as 'open with scattered trees' by the surveyor. By close inspection (Fig. 11), it appears that the automatic method has mapped the area correctly according to the 2010 lidar data and a 2011 air photo. Field inspection on 10 September 2014 confirms this.

For the Tryvann area, with two different ALS pulse densities, the automatic mapping of open areas is very similar for the two datasets. The few differences are mainly due to removal of areas smaller than 225 m^2 (Fig. 12). Some of the open areas are elongated, and small differences in the two datasets may result in small elongated areas that are contiguous in one dataset and fragmented in the other. However, some open areas with scattered trees are mapped as larger by the automatic method than by the manual survey (Fig. 12).
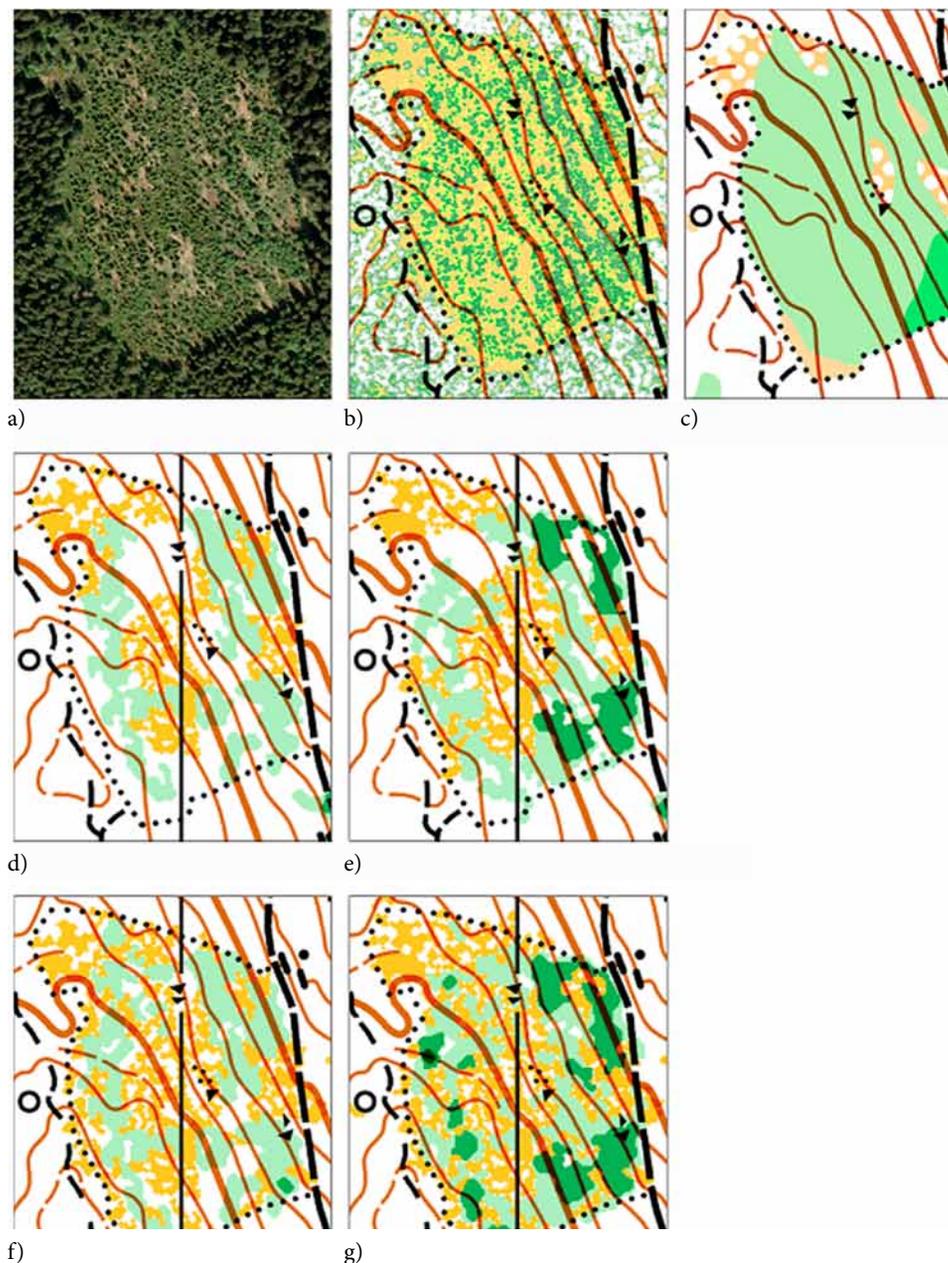
Fig. 12. Inspection of a small area from the Tryvann area. (a) Air photo of 13 June 2011. (b) Vegetation height. (c) Manual survey. (d) Automatic mapping on 2/m^2 data, minimum area of colour tone = 225 m^2. (e) Automatic mapping on 10/m^2 data, minimum area = 225 m^2. (f) Automatic mapping on 2/m^2 data, minimum area = 22.5 m^2. (g) Automatic mapping on 10/m^2 data, minimum area = 22.5 m^2

Field inspection on 2 November 2013 confirms that the automatic mapping is more correct.

The mapping of reduced runability is in less agreement with the manual survey than the mapping of open areas. The best agreement is on the Tryvann area, with about 50% of the area mapped as 'slow run' by the automatic method is also mapped as 'slow run' by the manual survey. However, by visual comparison of the automatic and manual survey on the Tryvann area, it appears that the major areas with reduced runability according to the manual survey have been identified by the automatic method. The difference is mainly in the exact shape of each mapped area with reduced runability. Regarding the different levels of reduced runability, there is little correspondence between the automatic method and the manual survey for areas mapped as 'slow run' or 'walk', both from the pixel-by-pixel comparison and visual comparison. It appears to be more meaningful to use one common class for reduced runability in the automatic method. By doing so, the correct classification rate for 'dense' vegetation is about 60% for the Tryvann area. However, there are differences locally between the automatic mapping on the $2/m^2$ data and the automatic mapping on the $10/m^2$ data (Fig. 12). Some of these differences are reduced by lowering the minimum area size. The largest remaining difference is that the automatic mapping of the $10/m^2$ dataset results in more areas of 'walk'.

For the Tømte area, the visual comparison indicates that the mapping of reduced runability is in relatively good agreement with the manual survey for the eastern half of the terrain, in that the general distribution of reduced runability is relatively similar in the two methods, but with many differences in the details. However, for the western part of the terrain, the automatic method maps much more as reduced runability than the manual survey. By merging the three reduced runability classes into one, the correct classification rate for reduced runability is now about 43%, compared to 60% for the Tryvann area. The difference is expected, since all of the Tryvann area had a good visual agreement regarding the general distribution of reduced runability, whereas this was only the case for about half of the Tømte area.

For the Bjørnholt area, several of the areas mapped manually as reduced runability are identified by the automatic method, but not all. In addition, the automatic method includes many small areas with

reduced runability, which were not mapped as such by the manual survey. When using just one class for reduced runability, only 26% of the area mapped as reduced runability by the automatic method is mapped as that by the manual survey. The pattern with many small 'slow run' areas mapped by the automatic method is similar to the western part of Tømte.

For the Bergendal area, most of the areas mapped as reduced runability by the manual survey are also captured by the automatic method, but not all. The pixel-by-pixel correspondence between the manual survey and the automatic mapping of reduced runability is almost 50%.

Merging the three reduced runability classes 'slow run', 'walk' and 'fight' to one 'dense' class increases the correct classification rate of the automatic method by 1.5% overall, and between 0.5% and 2.0% for the four map areas. Assuming that the manual survey is correct, the overall correct classification rate for the automatic method is about 75%, and varies from 72% to 77% for the four map areas (Table 8).

By further reducing the number of classes to two, by merging the 'normal' and 'dense' classes to a common 'forest' class, and still insisting that the manual survey is correct, increases the correct classification rate to 87% overall. However, we have already argued that, apart from new clear cuts between the time of the lidar acquisition and the manual survey, the automatic mapping of open areas is better than the manual

Table 8. Classification rates for different number of classes. For 5 classes, all five classes in Table 7. Classification rates. are used. For 3 classes, 'slow-run', 'walk' and 'fight' are merged to 'dense'. For 2 classes, 'normal' and 'dense' are merged to 'forest'

| | Number of classes | | | |
| | default* | 2 | 3 | 5 |
	normal	open, forest	open, normal, dense	open, normal, slow-run, walk, fight
Bergendal $2/m^2$	54.46%	87.69%	75.99%	73.89%
Bjørnholt $2/m^2$	65.99%	84.55%	74.74%	74.35%
Tømte $2/m^2$	63.83%	87.28%	72.50%	70.98%
Tryvann $2/m^2$	63.35%	88.50%	77.17%	75.02%
Overall $2/m^2$	61.83%	87.10%	74.78%	73.21%
Tryvann $10/m^2$	63.35%	88.49%	77.73%	73.85%

*Note: for default, everything is classified as 'normal', which means 'forest, easy running', and compared to the classes 'normal', 'open' and 'dense'.

survey, so that in the ideal case that the lidar acquisition is done immediately before the manual survey, and with no ongoing logging activities, the correct classification rate of the automatic method is 100%, and that any deviations are mistakes done by the manual survey.

By assuming that the automatic mapping of open land is 100% correct, we may recompute the classification rates for 'normal' forest and 'dense' forest. With this assumption, the average correct classification rate is about 87%, and for 'normal' forest it is almost 92%, but for 'dense' forest it remains at 44% overall, with large variations from 25% to 60% for the four mapped areas (Table 9).

Table 9. Stipulated classification rates with the assumption that the automatic classification of open land is correct

		Correct classification rates				
		Bergendal	Bjørn-holt	Tryvann	Tømte	Overall
Auto-matic classifica-tion	nor-mal	90.47%	94.91%	91.78%	90.97%	91.79%
	open	100%	100.00%	100.00%	100.00%	100.00%
	dense	48.35%	25.69%	59.98%	42.76%	44.36%
	total	87.43%	89.45%	87.92%	83.72%	86.65%

4. Discussion

The results of this study show that the automatic method is better than manual survey of open areas and open areas with scattered trees. What remains to be done by manual survey of these areas is to discriminate between areas with rough ground vegetation versus areas with easy running.

However, the results also show that automatic mapping of reduced runability is difficult. For some areas, the automatic method is able to identify the major patterns of reduced runability, whereas for other areas, the method is less successful. One potential problem is that many lidar pulses are blocked before they reach the 0–2 m vegetation height interval, which is the part of the vegetation data that the automatic method is using to estimate reduced runability. Deciduous trees block the signals almost completely in the leaf-on period. So, using lidar data from the leaf off period would be better. Also, one may think that increasing the point density could be beneficial. However, for the Tryvann area, which featured a high point density dataset in addition to the standard one, the high point density data did not result in better results in automatic reduced runability mappings. There could be several reasons for

this, including: (1) since the coniferous forest vegetation seldom blocks all lidar pulses from reaching the ground and the 0–2 m vegetation height interval, the percentages of pulses reaching these parts remains relatively constant with increasing pulse density; thus the NDVD also remains relatively constant. (2) The manual survey might not be correct (see below).

A potential problem when comparing the manual survey of reduced runability with the automatic mapping is that the manual survey might not always be correct. We have seen that for some parts of the terrain, there is a fairly good correspondence between what is mapped automatically and by manual survey, whereas for other parts the correspondence is not good. This could be related to different mapping styles by different mapmakers. If this is the case, then automatic mapping could be a tool in harmonizing the mapping styles of several mapmakers.

However, the lack of correspondence might also be related to different types of tree vegetation. Tree height, tree species and growing conditions could all affect the number of laser beams that are blocked before reaching the 0–2 m above the ground interval. We have already seen that deciduous trees block the laser pulses quite effectively in the leaf on season. In this case, there will be no or very few vegetation returns in the 0–2 m interval above the ground; however, there could be occasional ground returns in small openings between the trees, which may result in a very low NDVD even if the vegetation is actually dense 0–2 m above the ground. There could also be differences between spruce and pine; however, these are probably minor relative to the difference between coniferous and deciduous trees. Species or species groups classification may be possible from hyperspectral data (Dalponte *et al.* 2013); however, this assumes that hyperspectral data is available, which was not the case for the areas in this study.

Another possible source of error could be the spacing of individual trees. The method for estimating reduced runability is measuring an average density. If the spacing is relatively even, then the perceived density is probably quite close to the estimated density. However, if the tree spacing is uneven, which is often the case in the areas in this study, then there are small gaps and openings which increases the runability. One could investigate if texture measurements (e.g., Haralick *et al.* 1973) could be used to discriminate between even and uneven spacing of trees. For example, the textures could be computed on the thresholded NDVD image.

Although the automatic mapping of reduced runability produces results of varying quality, the method may still be useful. It may be used to quickly produce an initial mapping of reduced runability. This may be useful for the planning of manual survey, since dense forest is much more time consuming to survey than forest with normal runability. It may also be used to select which areas to map and which to discard. An initial mapping may also be useful for course setters who wish to plan orienteering competitions before the final map is ready. Further, the automatic method may also be used in the production of orienteering maps in remote areas where a full manual survey cannot be afforded. Finally, automatic mapping of reduced runability is an obvious candidate to replace tree height in base maps for manual survey; the feedback from surveyors is that the open areas in the colour-coded vegetation height maps (Fig. 1) are very useful, but the tree height in these maps are less useful.

One may discuss whether an automatic method for estimating forest density or runability should rather have produced values on a continuous scale rather than four categorical classes (white, light green, middle green, and dark green). This could be obtained by making a smoothed version of the NDVD, and either skipping the morphological steps or using gray scale morphological operators. However, a map with a continuous scale is difficult to read for humans, both in terms of perceiving the exact values and in comparing values from different parts of the map (Mak, Coulson 1991). Also, the use of four categorical classes is mandatory in the ISOM, presumably for the same reason.

The methods may be adapted to automatic mapping of vegetation for topographic maps in smaller scales, e.g., 1:25,000 or 1:50,000. The mapping of open land versus forest may be generalized using standard generalization methods for categorical data in raster format, e.g., majority voting inside aggregated pixels. For the mapping of dense vegetation, one must first consider which property of the forest that may be desirable to map in topographic maps: forest density for the 0–2 m vegetation height interval (from modified NDVD), forest density for the entire canopy (from original NDVD), or forest tree height (from nDSM). Also, one must decide if a continuous colour mapping is desired or how many categorical classes to use.

Conclusions

In conclusion, the automatic mapping of open land and open land with scattered trees is better than manual survey. The automatic method for mapping of reduced runability may be useful as an initial mapping, but additional manual survey is recommended.

References

ASPRS (The American Society for Photogrammetry and Remote Sensing) 2010. *LAS Specification: Version 1.3 – R11. 24 October 2010* [online], [cited 29 December 2014]. The American Society for Photogrammetry and Remote Sensing. Available from Internet: http://www.asprs.org/a/society/committees/standards/LAS_1_3_r11.pdf

Clode, S.; Rottensteiner, F.; Kootsookos, P.; Zelniker, E. 2007. Detection and vectorization of roads from lidar data, *Photogrammetric Engineering & Remote Sensing* 73(5): 517–535.

Dalponte, M.; Økra, H. O.; Gobakken, T.; Gianelle, D.; Næsset, E. 2013. Tree species classification in boreal forests with hyperspectral data, *IEEE Transactions on Geoscience and Remote Sensing* 51(5): 2632–2645. http://dx.doi.org/10.1109/TGRS.2012.2216272

Dorninger, P.; Pfeifer, N. 2008. A comprehensive automated 3D approach for building extraction, reconstruction, and regularization from airborne laser scanning point clouds, *Sensors* 8(11): 7323–7343. http://dx.doi.org/10.3390/s8117323

Ene, L.; Næsset, E.; Gobakken, T. 2012. Single tree detection in heterogeneous boreal forests using airborne laser scanning and area based stem number estimates, *International Journal of Remote Sensing* 33(16): 5171–5193. http://dx.doi.org/10.1080/01431161.2012.657363

Haralick, R. M.; Shanmugam, K.; Dinstein, I. 1973. Textural features for image classification, *IEEE Transactions on Systems, Man, and Cybernetics* 3(6): 610–621. http://dx.doi.org/10.1109/TSMC.1973.4309314

Kwak, D.-A.; Lee, W.-K.; Lee, J.-H.; Biging, G. S.; Gong, P. 2007. Detection of individual trees and estimation of tree height using lidar data, *Journal of Forest Research* 12(6): 425–434. http://dx.doi.org/10.1007/s10310-007-0041-9

Mak, K.; Coulson, M. R. C. 1991. Map-user response to computer-generated choropleth maps: comparative experiments in classification and symbolization, *Cartography and Geographic Information Systems* 18(2): 109–124. http://dx.doi.org/10.1559/152304091783805527

Næsset, E. 1997a. Determination of mean tree height of forest stands using airborne laser scanning data, *ISPRS Journal of Photogrammetry & Remote Sensing* 52(2): 49–56. http://dx.doi.org/10.1016/S0924-2716(97)83000-6

Næsset, E. 1997b. Estimating timber volume of forest stands using airborne laser scanning data, *Remote Sensing of Environment* 61(2): 246–253. http://dx.doi.org/10.1016/S0034-4257(97)00041-2

Papšienė, L.; Papšys, K. 2011. Possibilities of updating small-scale basic spatial data in Lithuania using generalization methods, *Geodesy and Cartography* 37(4): 143–148. http://dx.doi.org/10.3846/13921541.2011.645310

Smreček, R.; Michňová, Z. 2014. Identification of individual trees and groups of trees in the landscape using airborne laser scanning data, *Geodesy and Cartography* 40(3): 110–115. http://dx.doi.org/10.3846/20296991.2014.962736

Smith, A. M. S.; Falkowski, M. J.; Hudak, A. T.; Evans, J. S.; Robinson, A. P.; Steele, C. M. 2009. A cross-comparison of field, spectral, and lidar estimates of forest canopy cover, *Canadian Journal of Remote Sensing* 35(5): 447–459. http://dx.doi.org/10.5589/m09-038

Soille, P. 2002. *Morphological image analysis: Principles and applications.* 2nd ed. Berlin: Springer.

Stoter, J.; Post, M.; van Alten, V.; Nijhuis, R.; Bruns, B. 2014. Fully automated generalization of a 1:50k map from 1:10k data, *Cartography and Geographic Information Science* 41 (1): 1–13. http://dx.doi.org/10.1080/15230406.2013.824637

Tveite, H.; Langaas, S. 1999. An accuracy assessment method for geographical line data sets based on buffering, *International Journal of Geographical Information Science* 13(1): 27–47. http://dx.doi.org/10.1080/136588199241445

Yan, W. Y.; Shaker, A.; El-Ashmawy, N. 2015. Urban land cover classification using airborne LiDAR data: A review, *Remote Sensing of Environment* 158: 295–310. http://dx.doi.org/10.1016/j.rse.2014.11.001

Zentai, L. (Ed.). 2000. *International Specification of Orienteering Maps* [online], [cited 29 December 2014]. International Orienteering Federation. Available from Internet: http://orienteering.org/wp-content/uploads/2010/12/International-Specification-for-Orienteering-Maps-2000.pdf

Øivind Due TRIER. Dr Scient, senior research scientist at the Norwegian Computing Center, Section for Earth Observation, P.O. Box 114 Blindern, NO-0316 Oslo, Norway. Phone +47 22852698. E-mail: trier@nr.no. Dr Scient degree in informatics in 1996 at the University of Oslo. Research interests: automatic and semi-automatic processing of remote sensing data for various applications, including cultural heritage detection, vegetation and vegetation change mapping, object detection, snow and ice monitoring, and oil spill detection. Remote sensing data includes lidar, optical and SAR.