Report to the Danish AgriFish Agency

February 2016

Research and development in the use of LiDAR elevation data for mapping of landscape elements including hedgerows, ditches and dikes

Forskning og udvikling indenfor anvendelse af LiDAR data til kortlægning af landskabselementer som hegn, grøfter og diger.

Radoslaw Malinowski, Department of Agroecology, Aarhus University

Contents:

Summary

1. Landscape elements – Trees
   1.1. Introduction
   1.2. Data used
   1.3. Methods
   1.4. Results

2. Landscape elements - ditches and dikes
   2.1. Ditches
   2.2. Earth or stone dikes

3. Tools for processing and visualisation of landscape elements

4. Major issues remaining and conclusions

Bibliography
SUMMARY

This document reports the analysis performed at the Department of Agroecology of Aarhus University as a response to the order titled: Forskning og udvikling indenfor anvendelse af LIDAR data til kortlægning af landskabselementer som hegn, grøfter og diger.

According to the agreement between the Danish AgriFish Agency and the Department of Agroecology, the task of the latter was to investigate possibility of using newly acquired (2014 - 2015) national-wide LiDAR dataset (DK-DEM/Point cloud, see details in Section 1.2) and its derivatives for identifying and delineating selected landscape elements that include: three hedgerows, single trees and groups of trees, stone and earth dikes, and ditches.

Availability of up-to-data LiDAR data with national coverage may serve as direct and straightforward source of information for maintaining other spatial databases such as Fieldblocks.

1. LANDSCAPE ELEMENTS – TREES

1.1. INTRODUCTION

Tree-related landscape elements such as single trees, groups of trees and tree hedgerows play important ecological and cultural functions in many European countries. In general, they serve as habitats for different flora and fauna species, and corridors for movement for various wildlife individuals. Additionally, being a kind of physical barrier they help protecting farm animals and crops from wind and prevent arable fields from wind soil erosion, which was found a very efficient successful mechanism in Denmark landscape (Veihe and Hasholt, 2006).

Mapping the tree-related landscape elements may be performed with various remote sensing technologies that enable analysis of large areas in a cost-efficient manner. Detection and delineation of areas covered with high vegetation is especially efficient with LiDAR data due to their 3-dimentional characteristic and ability of penetrating vegetation canopy. This provides possibility for analysing vegetation structure (vertical and horizontal distribution of wooden elements) and obtaining measurements from the ground, and as the result to measure vegetation height. Additionally, LiDAR assure high data accuracy and precision.

For the need of this report the tree/vegetation extent is understood as the area of the ground covered by the vertical projection of a tree crown/canopy on it (Jennings et al., 1999). Therefore, the final data presenting areas of the landscape overgrown by high vegetation (trees) represent information on the extent of the tree at its widest part (i.e. tree crown) and not the real extent of the tree at the ground level. Illustration of this term is presented in Figure 1.
1.2. DATA USED

The most important dataset used in this analysis is airborne laser scanning data acquired between years 2014 and 2015, for the whole country of Denmark, referred to as DK-DEM/Point cloud. The collection of this data was spread over three seasons i.e. spring 2014, autumn 2014 and spring 2015. Therefore, the newly collected data represents both spring leaf-off conditions and autumn conditions with often leaf-on state and possible presence of maize crop on the fields. The planned average point density (last return) of this dataset is about 4.5 points/m$^2$ and varies depending on the land cover. This is a substantial improvement as compared to the older LiDAR data acquired in 2005 – 2007 with average point density of circa 0.5 points/m$^2$. The accuracy of LiDAR measurements as stated by the data distributor (i.e. Agency for Data Supply and Efficiency) is 0.15 m for positional accuracy, and 0.05 m for vertical accuracy. The LiDAR data were collected in the full-waveform (FWF) mode and for areas flown during a daytime, with appropriate sun illumination, additionally RGB images were also collected. However, both FWF data and RGB images were not made available for public use during the time of work described in this report (March – November 2015). Similarly, data collected in autumn 2014 was not available and it was not possible to analyse differences between spring and autumn data in regards to trees (hedgerows) mapping. The data covering the whole Denmark was made available and ready for downloads from Kortforsyningen ([http://kortforsyningen.dk](http://kortforsyningen.dk)) only in January 2016.

Because of the above-mentioned lack of access to dataset acquired in autumn 2014 another dataset was used to check whether there might be any problems for landscape elements mapping with data representing leaf-on conditions. This work was conducted on the LiDAR data, referred to as LiDAR-2012, purchased by the Department of Agroecology, Aarhus University for performing a research
study on mapping localized flooding. This data was acquired on 23rd September 2012 for an area of approx. 9 km² along a 10 km reach of the Nørreå River. The average point density of this data is 10 – 15 points/m² with 0.15 m for horizontal and 0.1 m for vertical accuracy.

The performed analysis also utilized Digital Terrain Models (DTM) referred to as DK-DEM/Terrain that was derived from the DK-DEM/Point cloud. DK-DEM/Terrain was provided in a GeoTIFF format with 0.4 m cell size and was mainly utilized for calculation of a normalized Digital Surface Model (nDSM) or so-called vegetation height model.

Another dataset used in the current study was DDO®Land2014 orthophoto product of COWI A/S (Cowi, 2014) provided with national coverage. This data was acquired between 15 May and mid July 2014, therefore it represent the most up-to-data image dataset in Denmark. It is provided by the data supplier with ground sampling distance (GSD) of 0.16 m and was used for the purpose of visual accuracy assessment of extracted vegetation mask. The horizontal accuracy of the orthorectified images is estimated being around 0.35 m (www.cowi.dk).

The last set of data utilized in the analysis is a vector dataset called Fieldblocks (Markblokke2014) and represents Danish implementation of the Land Parcel Identification System (LPIS). The Fieldblocks dataset is a reference system of land parcels composed of blocks, where each block contains a coherent set of 1 to 10 field parcels delimited by permanent physical borders. The mapping scale of Fieldblocks is 1:10,000.

1.3. METHODS

1.3.1. Study areas

The process of mapping spatial extent of landscape features realted to trees such as hedgerows, single trees and groups of trees was performed within four locations spread out on Zealand, Funen and Jutland (Figure 2). The study areas (areas of interest – AOIs) were selected in the way so that they provide examples of different settings of trees in the landscape and concerns other important aspects of mapping trees from LiDAR data. AOI 1 (1x1 km size) represents agricultural fields separated by hedgerows with small forest included (Figure 3a). Additionally, there are two middle-size powerlines running throughout the analysed area. AOI 2 (1x1 km size) (Figure 3b), besides the numerous hedgerows growing around the agricultural files, also contains many orchards that could pose problems while mapping hedgerows. AOI 3 (1x1 km sieve) (Figure 3c) shows relatively large fields sourounded by hedgerows, small forest and very tall electricity pylons with numerous densely fixed overhead powerlines (also visible on Figure 3c). Finally, AOI 4 (2.5 x 0.5 km size) (Figure 3d) represents a location with not harvested maize fields. For the AOIs 1, 2 and 3 the DK-DEM/Point cloud was used (spring dataset), while AOI 4 utilized LiDAR-2012 dataset (autumn dataset).
1.3.2. Vegetation mask extraction

The DK-DEM/Point cloud is provided for users with classification of the laser points that include among other the following classes: low vegetation, medium vegetation and high vegetation. However, after initial check of the classified points several erroneous classifications were found (Figure 4). Because the magnitude of misclassification was not known beforehand, the following data processing was performed on the DK-DEM/Point ignoring the existing classification. Therefore the analysis was based only on the geometric information (3D coordinates) provided by LiDAR data.

All different data layers used for extracting the extent of trees were created with a fixed cell size of 1 meter. The mentioned cell size was selected to ensure high level of details derived from the LiDAR data and to avoid, at the same time, apperance of cells with NoData value that would correspond to areas where no LiDAR pulse was registered. The selection of cell size is directly dependent on the LiDAR point density and in this study it was empirically selected through trial and error.

In the first step of data processing the DK-DEM/Terrain was resampled from 0.4 m cell size to 1 m cell size. This was done to assure that all pixels from the different raster datasets used in analysis are perfectly aligned. For resampling the DK-DEM/Terrain was converted to a point shapefile with elevation information assigned to each point located in the center of initial raster data. In the following step a new raster data was created with a 1 m cell size by interpolating elevation values from the points by inverse distance weighted (IDW) technique (in ArcGIS Desktop).

After performing the DTM resampling the OPALS software (Pfeifer et al., 2014) was used to generate a digital surface model (DSM) from laser points in the way that for each raster cell the elevation value was assigned of the highest laser point located within the area of a given raster cell. Following that step a nDSM was derived by subtracting the interpolated DTM from the DSM.
Figure 3. Landscape settings in the analysed areas.

Figure 4. Example of incorrect classification of points in DK-DEM/Point cloud. Yellow, dark green and bright green colours represents points classified as low, middle and high vegetation, respectively. Brown colour represents points classified as ground and grey colour shows unclassified points. Many laser points reflected from powerlines and electricity pylon are incorrectly classified as vegetation.
The extracted nDSM layer contained information about vegetation objects but also on other elevated off-terrain objects such as buildings, powerlines etc. Therefore, it was necessary to filter them out in order to create a real vegetation mask. For removing buildings object the echo ratio (ER) measure was used. The ER is a measure, implemented in the OPALS software, which provides indication of the local surface permeability or transparency. It is defined as ratio between the number of points located in a predefined 3D fixed search radius (a sphere) to the number of points located in the same radius but measured in 2D (vertical search cylinder with infinite height) (Opals, 2016). The values of ER are close to 100% for impervious surfaces (i.e. asphalt, bare earth, building roofs) and decreases for penetrable surfaces such as vegetation due to penetration of laser points through the vegetation structure (i.e. tree crown). Areas were recognized as buildings and removed from the nDSM in case when the values of a raster layer representing ER measure were below a threshold of 90% (empirically determined). Subsequently, the nDSM was processed with morphological opening filter that helped in removing isolated single pixels (too small for being a tree) and misclassified objects originating, for example, from powerlines. Finally, the resulting nDSM was thresholded with a value of 2 meters for deriving the vegetation mask. The threshold of 2 meters is frequently used for filtering trees (Kaartinen et al., 2012).

The vegetation mask was a dataset representing the spatial extent of objects classified as trees with height above 2 m. Another optional step was applied on that mask and was performed in object-oriented eCognition software (v. 9.0, Trimble Germany GmbH, Munich, Germany, 2014). In this process all pixels (objects) with height above 1.5 m were also added to the vegetation mask if they were adjacent to another vegetation object classified in the previous steps. In this way also small trees and bushes were found as well as lower lying branches of bigger trees. Additionally, this step allowed to retrieve small trees elements that were removed by the morphological filter.

1.3.3. Accuracy assessment

To assess the accuracy of the extracted vegetation mask a point-based method was used. The validation was performed separately for two study areas, AOI 1 and AOI 2. For each of them 600 sampling points were used in total. The samples were generated randomly in two ways. A sample set of 400 points (for each study area) was generated by stratified sampling, assuring 200 samples being selected for areas classified as trees and 200 samples for areas classified as non-vegetation (the remaining area). Additional 200 samples were generated randomly for each of the two study areas without considering land cover classes. In the following step the samples were checked by visual interpretation with the DDO®Land2014 orthophoto product and DK-DEM/Point cloud. After visual interpretation basic statistics were derived that estimate the accuracy of the extracted vegetation mask. These statistics include overall accuracy (OA) as well as completeness and correctness of the trees class.

Another method for accuracy check would be comparison of coverage of the classified vegetation mask and manually delineated tree extent. However, such method is very time consuming and more subjective regarding the quality of delineation of the trees extent by the operator. This is due to the very irregular tree structure and often small size of branches. Such method would be also much more dependent on the shadowing issue apparent in the summer orthophoto. Finally, manual delineation of the reference data would result in a relatively smooth boundary of trees, while the
boundary of the classified trees would follow the shape of pixels in the raster data (Figure 5). This would result in additional areal difference between the reference and analysed datasets.

![Reference data](image1.png)

**Figure 5.** Pixelated border of classified hedgerow compared to straight boundary of reference data.

### 1.3.4. Application of vegetation mask for Fieldblocks verification

Once a vegetation mask was extracted it was possible to analyse it together with the Fieldblocks dataset and check the potential overlap of these two datasets. Intersection of vegetation mask and Fieldblocks datasets resulted in a new dataset presenting the extent of areas where tree crowns overlap with aerial fields.

### 1.4. RESULTS

#### 1.4.1. Trees extent mapping

Figure 6 presents the result of tree extent extraction for the four analysed areas. The area classified as trees includes hedgerows, single trees and group of trees growing within the analysed areas.

The presented data show that the method provides precise estimate for the location and extent of trees. A visual inspection indicates only problem with the powerlines mapped as trees for the study area AOI 3 (Figure 6c). No other apparent error might be seen from the maps presented.

The results of accuracy assessment are presented in Table 1 and shows very high accuracy scores for both analysed areas (AOI 1 and AOI 2). The overall accuracy represents the ratio between the correctly classified samples (both classes) and the total number of samples. The completeness measure represents the level of false negative classifications (under-estimation of the Tree class). It means that 98% (for AOI 1) of all samples of trees from the reference dataset were also trees in the classified data. The correctness score shows the level of false positives (over-estimation of the Tree class). It shows that nearly 99.6% (for AOI 1) of samples classified as trees were classified correctly.
Table 1. The accuracy assessment measures derived for the two study areas.

<table>
<thead>
<tr>
<th>AOI #</th>
<th>Overall accuracy (%)</th>
<th>Completeness of Tree class (%)</th>
<th>Correctness of Tree class (%)</th>
<th>Correctly classified samples (out of 600)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AOI 1</td>
<td>99.0</td>
<td>97.9</td>
<td>99.6</td>
<td>594</td>
</tr>
<tr>
<td>AOI 2</td>
<td>98.2</td>
<td>96.1</td>
<td>99.1</td>
<td>589</td>
</tr>
</tbody>
</table>

However, a closer look at the results combined with knowledge of the analysed areas indicates some problems.

**Powerlines**

First, as mentioned before power lines were incorrectly classified as trees and it can be observed at Figure 6c as double lines running from the top to bottom of the study area. There are powerlines present also within the area AOI 1, however, the once from AOI3 are much bigger with electricity pylons 50 m tall and are composed of powerlines attached at four levels (Figure 4). In such a way there are many laser points reflected from wires and compose object on rasterised data (nDSM) that are too big to be filtered with morphological operations. The magnitude of this issue is difficult to estimate at the national scale and similar problem is expected to be monitored also for windmills. If
this error is to be removed the point cloud data would need to be filtered out. This would result in points reflected form powerlines, pylons and windmills (and other similar structures) being removed or classified into some predefined class and not included within the tree classification process. Such filtering would require further data processing and could be supported by existing vector data such as KORT10 that provides up-to-date information on various infrastructure objects such as powerlines, etc.

Orchards

Another important classification issue is related to orchards areas and Christmas tree plantations. The analysed areas did not provide examples of Christmas tree plantations, however, similar problems are expected as encountered with orchards. Figure 6b represents area with many orchards. For most cases orchards were intentionally classified as non-trees class as they belong to the group of permanent crops (pers. comm. Peter Viskum Jørgensen, 2015). However, to get that result it required additional application of morphological filter. This relatively efficient method for removing single orchard trees (failed only with more mature trees taller than 2 m) would lead to removing many real trees (e.g. from hedgerow) while performing on areas without orchards. Therefore using morphological analysis requires additional caution.

Maize crops

Probably the most important problem in precise mapping of landscape features related to trees is misclassification of maize crops. Due to the fact that maize plants easily reach height above 2 m they may be in many cases classified as groups of trees. Additionally, because maize fields cover relatively big areas the morphological filtering may be not as efficient as in the case of orchard trees. This problem may be observed in Figure 7 that shows the extract of the eastern part of the AOI 4 study area. The yellow polygons in that figure represent not harvested maize, which was classified as trees. This problem would concern the part of the DK-DEM/Point cloud data acquired in autumn 2014, and especially data captured during September and October months. With such data there might be considerable amount of maize still not harvested. The magnitude of potential misclassification will be directly dependent on the size of the area for which LiDAR data was collected at the mentioned months and what kind of landscape was concerned (i.e. agricultural, forested or urban areas). This kind of information was not available, however, at the time of data processing and report preparation.

Figure 7. Example of misclassification of maize crops as trees. Green colour represents correct identification of trees and yellow colour indicates maize wrongly classified as trees.
1.4.2. Trees overlap with Fieldblocks data

Figure 8 presents results of spatial analysis, in which the Fieldblocks (Figure 8b) dataset was intersected with data presenting the tree extent (Figure 8a). Resulting dataset (Figure 8c) represents the overlapping areas and indicates locations and size of encroachment of trees (i.e. tree crown) on arable lands. This type of information may be used for updating Fieldblocks dataset if relevant spatial requirements are fulfilled.

Figure 8. Results of analysis of existing Fieldblocks data with newly derived tree extent mask.
2. LANDSCAPE ELEMENTS - DITCHES AND DIKES

Detection of landscape objects such as ditches or earth/stone dikes are highly dependent on the quality and properties of the data to be used. Due to their possibly relatively small size in the landscape scale adequate data have to be acquired that makes possible their detection. An analysis of ditches and dikes with the newly collected DK-DEM/Point cloud started with an on-screen initial visual inspection of these landscape elements and verification of possible analyses approaches.

As opposite to landscape elements such as trees (hedgerows) that have quite distinct structure and their efficient detection is possible regardless of the season (at both leaf-on and leaf-off state), as was shown in Section 1, detection of ditches and dikes may be hampered due to several factors.

2.1. Ditches

One of the most important constraint for successful ditches delineation is possible high water stage in ditches. Depending on the local hydrological conditions it might be difficult to extract precise information on the ditch location and extent from the LiDAR data acquired during high water stages in the ditch network. Figure 9 presents examples of ditches without standing water (a) and with high water stage in a ditch (b). While in the situation without water or with low water level in a ditch it would be possible to extract information of the ditch geometry, it would be impossible in the second case.

Similar case is expected when ditch and its surrounding is overgrown with dense herbaceous or even bushy vegetation. An example is shown in Figure 9c where the ditches is completely masked with herbaceous vegetation.

2.2. Earth or stone dikes

The situation of dikes is very similar to that of ditches. They are not dependent on hydrology but are often overgrown with different vegetation including herbs, bushes or even trees (Figure 10). Their very irregular shape also makes it more difficult to delineate their extent (Figure 10). In such a case the quality of dikes mapping is not only depended on the method used but also on the quality of separation of the laser points reflected from the ground and those reflected from vegetation.

In both cases of ditches and dikes, more work is needed with applied geomorphometry methods to extract the mentioned landscape elements. Such methods were not tested widely in the presented work. This among other would require precise field work performed during the LiDAR data acquisition which was not possible (unknown data acquisition time) and because of late delivery (late January 2016) of LiDAR dataset for the known locations (e.g. Nørreå valley). Due to the hydrologic and overgrowing vegetation issues mentioned above and presented in Figure 9 and 10, in some cases the manual interpretation of spatial data, including LiDAR data and orthophotos, would be the only method to derive the required information.
Figure 9. Examples of ditches seen from LiDAR point cloud (point cloud profile) and orthoimage. (a) – a ditch in dry condition (spring dataset), (b) – a ditch during high water stage (autumn dataset), (c) – a ditch with overgrowing vegetation.
3. TOOLS FOR PROCESSING AND VISUALISATION OF LANDSCAPE ELEMENTS

There was different software used for processing LiDAR dataset either as point cloud data or raster data. This included OPALS (Orientation and Processing of Airborne Laser Scanning data), eCognition and ArcGIS Desktop. Details are provided in Section 1.3. Although the most professional tools are provided by the LiDAR data dedicated software such as OPALS, the more commonly used GIS software like ArcGIS also enable analysis of LiDAR data, including analysis of the data in a point cloud format. Especially visualisation of LiDAR point cloud is easily handled in ArcMap application and may be performed using so-called LAS Dataset. Such tool allow for creating LiDAR database that stores a number of datasets in LAS format. Using LAS Dataset the user may visualize point cloud while rendering it according to elevation, classification, number of returns etc. Additionally, different measurements may be taken on the laser points and they can be also manually classified.
More possibilities may be seen from a seminar titled ‘Working with Lidar Data in ArcGIS 10.1’ (http://training.esri.com.gateway/index.cfm?fa=catalog.webCourseDetail&courseID=2497).

Figure 11 presents a screenshot with the LAS Dataset toolbar and two available viewers that might be used for screening and editing LiDAR point cloud, Profile viewer (Figure 11b) and 3D viewer (Figure 11c).

Figure 11. A screenshot presenting tools available in ArcGIS for visualizing and editing point cloud datasets. (a) – LAS Dataset toolbar, (b) Profile viewer, (c) 3D viewer.

All the options available for visualization, analysis and editing of LiDAR point cloud are included in the LAS Dataset toolbar or in the Manu bars places at the top of the Profile and 3D viewers.
4. MAJOR ISSUES REMAINING AND CONCLUSIONS

This report presents results of analysis which aimed in verification whether national-wide LiDAR dataset acquired in Denmark in recent years (2014 – 2015) may serve as data source for extraction spatial data on various landscape features. It has been showed that this data provides very useful information and especially for the tree-related features that could be extracted in a semi-automated manner.

Although some constraints were presented on mapping dikes and ditches, which are related among other to hydrological conditions met during data acquisition, there are several issues that might be addressed in a further research and could improve the results of landscape elements mapping.

- Optional parameters modifications – some thresholds used in the study were derived from literature or by practical reasoning. These values might be adjusted if necessary according to data parameters, specific study location or other requirements. This may include, for example, the pixel size of raster dataset derived from the LiDAR data (1 m pixel size was used for mapping trees). Similarly, the threshold of 2 m used for filtering trees from lower vegetation may need to be changed according to other specifications (the same applies to 1.5 m threshold used and described at the end of Section 1.3.2)

- Application of available DSM data – in this study the DSM was derived for extracting the vegetation mask. Currently with data available in different formats, including point cloud, DSM and DTM, for the whole country it would be desirable to test how the methods would perform with data that is delivered by the Agency for Data Supply and Efficiency. By utilizing ready to use data (DSM) some steps from the process presented above could be omitted. However, more information is necessary on the methods how the mentioned DSM was created and on its accuracy.

- Erroneous objects – some objects in the landscape such as powerlines or windmills are frequently misclassified as vegetation due to their shape and structure. Removal of errors resulting from these misclassifications might be performed while using ancillary data, for example the KORT 10 vector dataset. This data provide information on location and size of different infrastructural objects like powerlines.

- Separation of hedgerows from forest - additional analysis would be required to separate trees that were classified as hedgerows but in reality are parts of forested area. Spatial analysis supported with data presenting forest stands could facilitate such process.

- Automatization of landscape elements mapping – the whole process of extracting landscape elements, for example hedgerows, could be made more automated by connecting processes done by different software. By appropriate programing, for example with Python programing language, separate tools available with OPALS, ArcGIS and eCognition could be connected and used without the need for manual operating. This would enable processing of data covering bigger areas with only minor engagement of the data analysts.
BIBLIOGRAPHY


