

# CANOPY EXTRACTION USING AIRBORNE LASER SCANNING DATA IN FORESTRY AREAS

C. Y. Lo \*, L. C. Chen

Center for Space and Remote Sensing Research, National Central University, Jungli, Taoyuan 32001,  
Taiwan,China – (freezer, lcchen)@csrnr.ncu.edu.tw

Commission III, WG III/3

**KEY WORDS:** Forestry, Extraction, Estimation, Photogrammetry, Laser scanning, Radar

## ABSTRACT:

To understand the biomass condition in the forestry areas, the detection for individual trees becomes an important topic. For forestry management, the canopy may be described with several parameters, e.g., number of stems, stand position, tree height, timber volume, canopy cover, etc. Traditional methods rely on field measurements, imagery, and radar data to derive the canopy information. Following the development of airborne light detection and ranging (LIDAR) systems, this paper proposes a scheme to extract forestry canopy using raw data, i.e., discrete point clouds. In this study, the proposed scheme analyzes vertical profiles to detect elevation gaps and delineates canopy boundaries. The experimental results indicate that the proposed scheme is effective.

## 1. INTRODUCTION

For forestry management, the major parameters of biomass condition contain number of stands, tree height, timber volume, canopy cover, etc. They could be derived using several ways including of field measurement, aerial photos (Hirschmugl *et al.*, 2007), satellite imagery (Gougeon and Leckie, 2006), and radar data (Walker *et al.*, 2007). However, some limitations might exist in those approaches for canopy extraction, e.g., expensive cost, labor intensity, time-consuming, occlusion, blurred pattern, and so on. To extract individual trees is another issue when the focused canopies are mixed (Næsset *et al.*, 2004). Within spectral domain, a clump of tree crowns is hard to separate into individual object. It is more serious when those trees are covered with shadows. On the other hand, using radar data is also not an easy work for identification of individual tree position. Fortunately, based on the development of laser scanning system, an opportunity that estimates tree crowns from the planimetric coordinate to three-dimensional space becomes practical.

Several literatures focused on the process development about canopy extraction using light detection and ranging (LIDAR) data. During the last decades, Nelson *et al.* (1984) analyzed the characteristics of laser beam when foliage of canopies contacted the footprint. If laser beam contained enough power, it would penetrate leaves and derive further information under tree crown. Furthermore, the laser scanning system has the ability to detect targets within shadow sites, especially in forestry areas. For some species, e.g. boreal forests, laser scanning data even can provide as accurate results as field measurement-based models (Maltamo *et al.*, 2006).

Based on the raster processing, some approaches derived canopy information by employing image vision methods. Hyypä *et al.* (2001) rasterized discrete point clouds to grid-based data, i.e., digital surface models (DSM) and digital terrain models (DTM). When those vegetation points were classified

and compared with DTM, the authors also produced digital vegetation models (DVM) and digital tree height model (DTHM). Then they found local maximum to locate individual trees within the designed area using the segmentation method. Clark *et al.* (2004) proposed a scheme based on the grid data from laser scanning data. DSM and DTM were also produced for canopy identification. The elevation threshold between those two models was used to detect tree height and sub-canopy using inverse distance weighting and ordinary kriging techniques. Bortolot (2006) also used rasterized data, DSM, to derive the canopy model with digital elevation model (DEM). He proposed an idea to split and merge the segments. Other research fused different data to detect canopies. Wulder *et al.* (2007) combined LIDAR data and satellite imagery to identify the landscape change for tree crowns. Those images were used to locate the species in spectral domain. On the other hand, LIDAR data provided the vertical profiles to measure the changed areas. The study based on the grid data to analyze the canopy information, too. On account of the discrete data, Hosoi and Omasa (2006) used ground-based laser scanning system to compute the leaf area index (LAI) in an accurate way based on the voxel-based approach. However, this literature only focused on few targets to generate ground truth data, not for clumps of canopies.

According to those existing literatures, most of them based on the grid-interpolation based concepts to detect individual canopies from laser scanning data for forestry areas. Merged canopies, thus, were hard to explicit locate by oversmoothing and provided lower accuracy (Næsset *et al.*, 2004). Additionally, some approaches needed a seed point to proceed image segmentation. Different seed points and thresholds might cause different results. It is reasonable to assume that if the original data set with discrete point clouds is used, the higher accuracy might be expected. Thus, this study proposes a scheme to identify individual canopies using original laser scanning data, i.e. discrete point clouds. The method analyzes the vertical profiles using empirical mode decomposition (EMD) (Huang *et*

---

\* Corresponding author.

al., 1998) and identifies individual trees. Two test areas are selected with different canopy density and tree height.

## 2. CANOPY EXTRACTION

To indicate the individual stands, we propose a coarse-to-fine approach using laser scanning data to detect individual stands and delineate their boundaries. The sub-canopies are also separated from the main crown. Moreover, the proposed scheme may automatically merge those stands when they are too close to separate. Five conditions are pointed as follows. (1) The processes estimate crown edges with discrete point clouds. (2) The all returns of laser scanning data are used. (3) Explicit elevation changes may exist between two nearby canopies. (4) The locally highest point is probably the tree top. (5) Two nearby crowns may share the same edges. The proposed scheme includes two parts: pre-process and canopy extraction. In the first part, DEM and candidate points are produced. Those discrete point clouds are filtered out above-ground objects to derive the ground surface (Axelsson, 2001). Then we collect candidate canopy points from laser scanning data with an elevation threshold. Second, the extraction process analyzes vertical profiles to find the crown edge with discrete point clouds. We decompose the vertical profiles to locate the positions where change obviously. The sequential processes then delineate boundaries along the elevation gaps between two nearby canopies. The trial represents the results as polygons. The workflow is shown in Figure 1.

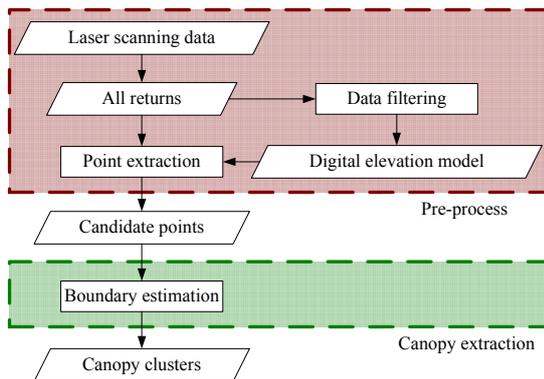


Figure 1. The workflow of the investigation

### 2.1 Pre-process

In this step, the main objective is to extract candidate canopy points from all returns of laser scanning data. The processes are shown in Figure 1. Because this kind of data has considerable discrete point clouds, directly analyzing it may be time-consuming and confusing the sequential processes to find the edges. Thus, the pre-process is to select candidates and remove others. Due to all points of LIDAR data represent the whole things in the area and the surface information. Following the general concepts, canopies grow above the ground with certain elevation. An intuitively way is to collect canopy points with an elevation threshold ( $\Delta h$ ) by comparing a reference surface, i.e., DEM. Therefore, original LIDAR data were filtered out those target points above ground for DEM generation. We then compare all points and DEM with  $\Delta h$  to collect canopy candidates. According to the above process, we use those candidate points for further processes to extract canopy edges.

### 2.2 Canopy extraction

In the canopy extraction, we examine the canopy profiles to distinguish individual trees using those candidate points. This stage contains three parts, pseudo-grid generation, profile decomposition, and polygon delineation. Following the authors' observations, those candidate points are noisy and hard to find the crown edges from all point data set. Therefore, the idea is to approximately locate the edges from coarse data, i.e., pseudo-grid data. The original points are then used to provide more detail to delineate boundaries. The flowchart for the extraction processes is shown in Figure 2.

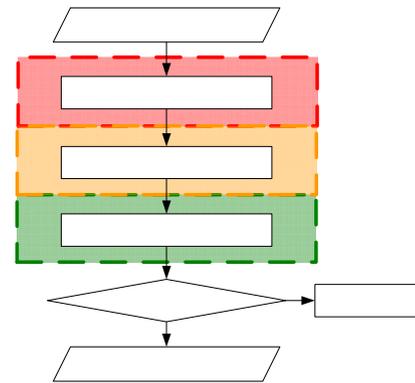


Figure 2. The workflow of canopy extraction

**2.2.1 Pseudo-grid Generation.** The pseudo-grid data provides a possibility to roughly indicate crown positions without interpolation step. This procedure is similar to the rasterization without interpolation. A virtual grid data is produced, and then this step inserts those candidate points into each cell. However, the non-data areas are keeping empty without interpolation so that the exterior shapes are extreme clear. Because those interior crown edges should have lower altitude within local areas, each cell then only selects the lowest point to be its value. Those data are produced to prepare for profile decomposition to find the inner edges. The process has illustrated in Figure 3.

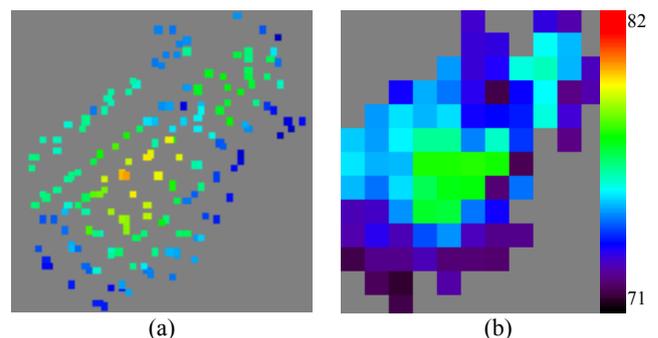


Figure 3. Pseudo-grid data generation: (a) all points, (b) pseudo-grid data (Unit: m)

**2.2.2 Profile Decomposition.** In order to find the crown edges from noisy canopy surfaces, this study proceeds the EMD approach to decompose the information and extract elevation change sites. EMD has been proven as a powerful tool to analyze the perturbed signals or images (Chen *et al.*, 2007; Wu *et al.*, 2007; Chen *et al.*, 2008). To test the performance of the

decomposition, we include EMD as the first part in canopy analyses. EMD finds the local maximum and local minimum on the original curve and fits them to produce two envelopes with B-spline (Figure 4a). The Mean curve is then calculated with those two envelopes and used to get the difference from the original curve (Figure 4b). The iteration stops when the final curve has the same numbers of zero-crossing and extrema. The stage result is the decomposed component, so called “Intrinsic Mode Function (IMF)”. The equation is shown below.

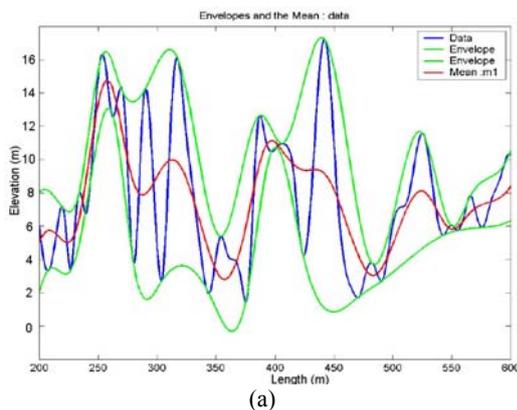
$$\begin{aligned} X(l) - m_1 &= h_1 \\ h_1 - m_2 &= h_2 \\ &\vdots \\ h_{n-1} - m_n &= h_n \Rightarrow h_n = c_1 \end{aligned} \quad (1)$$

where  $X(l)$  = original curve

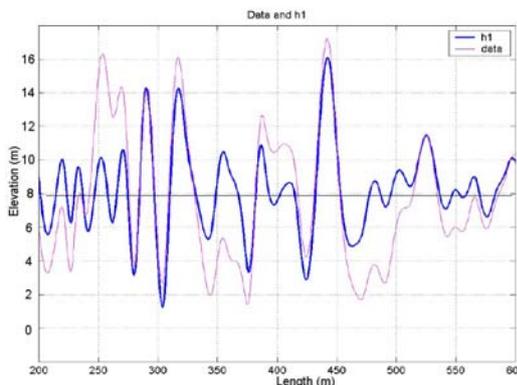
$m_n$  = mean curve

$h_n$  = difference of sequential curves

$c_n$  = intrinsic mode function



(a)



(b)

© Norden *et al.*, 1998

Figure 4. Calculation of IMF: (a) mean profiles, (b) difference of original data and  $h_1$  curve

After calculating one IMF, the original data minors the IMF to derive the first residual ( $r_1$ ) (Equation 2). The iterative procedure then calculates each IMF from the original data. The

stop criterion of the second part is also the same as the IMF calculation. Notice that the above procedure is reversible. That means the sum of all IMFs which equals to the original curve. The illustrations of other IMFs are shown below. Therefore, the results of decomposition still equal the original canopy information. Following the results of the decomposition process, we can derive several IMF components as Figure 5. The IMFs are sorted from high frequency to low frequency.

$$\begin{aligned} X(l) - c_1 &= r_1 \\ r_1 - c_2 &= r_2 \\ &\vdots \\ r_{n-1} - c_n &= r_n \end{aligned} \quad (2)$$

where  $r_n$  = residual

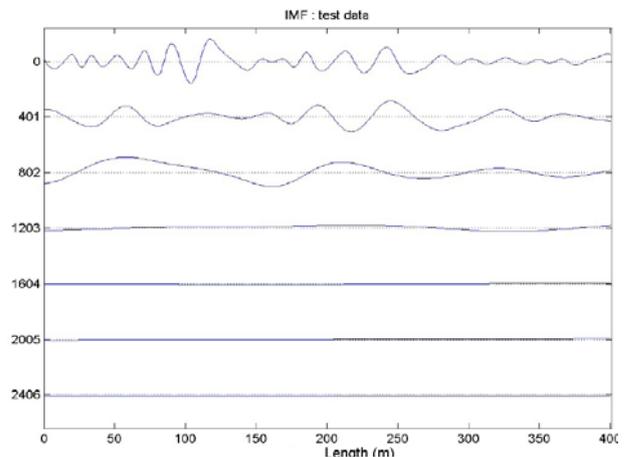


Figure 5. All IMFs of the decomposed data

In this investigation, we assume that the obvious elevation changes belong to noisy parts of canopy profiles so that may exist in the first IMF, i.e., the high frequent component. We, thus, re-produce the pseudo-grid data to one dimensional data in four directions, i.e., row-column and diagonal directions, and implement EMD to extract the first IMF in each direction. Then, those IMFs are reorganized to a new 2-D data as Figure 6 shown. This data represents the canopy edges in the pseudo-grid image.

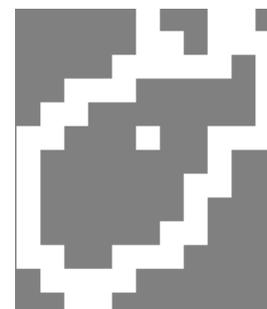


Figure 6. Example of the approximated canopy edge

Finally, we are back to the pseudo-grid data to extract point clouds which are within the edge areas (Figure 7a). The proposed scheme also finds the locally highest point of original data to be the tree center and connects those selected points around this reference spot. Furthermore, the linking step is constrained with the angles in order to ensure the canopy areas are the closed polygon. In Figure 7b, the colours mean different stands, and white cross is the tree top of each canopy.

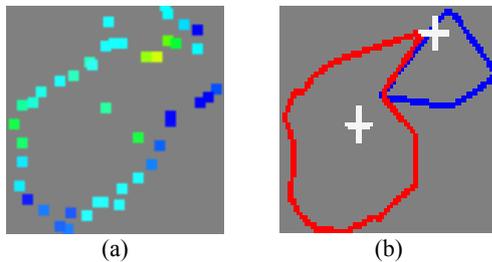


Figure 7. The results of proposed scheme: (a) Selected edge points, (b) delineated boundary

### 3. EXPERIMENTAL RESULTS

This investigation includes two test areas which locate in Espoo of the south Finland with pine forest. It's a pure forest with many pines containing different tree heights and canopy density. The areas of two cases have different areas and canopy densities, 660m<sup>2</sup> and 602m<sup>2</sup>. The case I and case II are the forestry areas with medium and high density canopy closures, respectively. The laser scanning data was delivered by Optech ALTM 2033 in June, 2004. The point density of all points reaches 8 pts/m<sup>2</sup>. Two test sites are selected, and those candidate points extracted from original point clouds are illustrated in Figure 8. This study focuses on those points above ground 2m and regards them as canopy candidates. In the process of canopy extraction, we would design the minimum canopy area to 1.5 m<sup>2</sup>. To verify the results, the results are compared with field-measurement data. The validation calculates how many stands are detected, merged, and missed. Thus, the individual check would estimate the quality of canopy extraction.

In Figure 9, we delineate the boundaries to represent the canopy cover of each stand. Each colour indicates different individual trees. The crosses and spots represent the positions of analyzed treetops and reference data, respectively. The results of canopy extraction are also listed in Table 1. By comparing the results with reference data, we can estimate that the stands are detected within the test areas. Some tree crowns may be too close to others so that the process automatically merges them to be a canopy boundary. Furthermore, some sub-canopies are also detected and extracted by the proposed scheme.

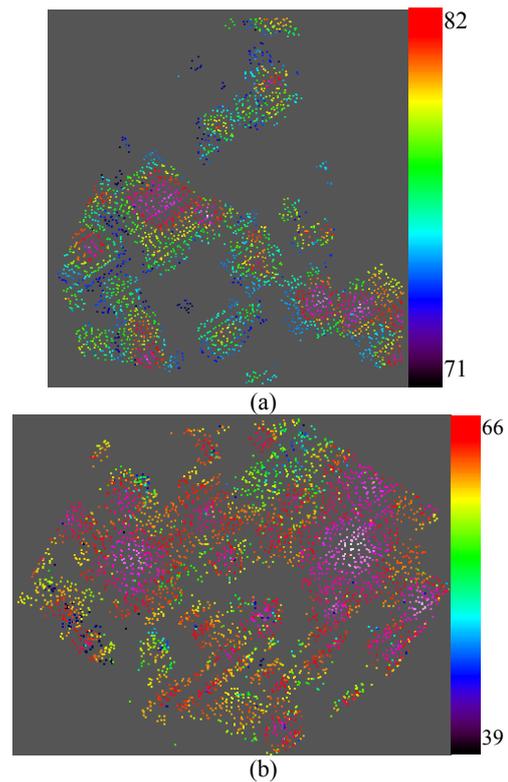
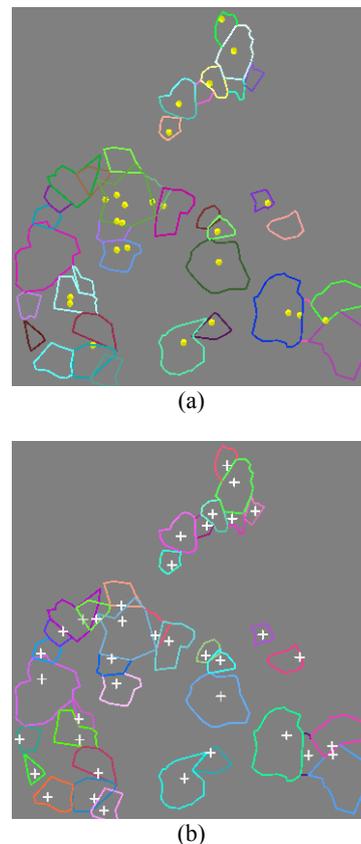


Figure 8. Candidate points in test areas: (a) case I, (b)case II (Unit: m)



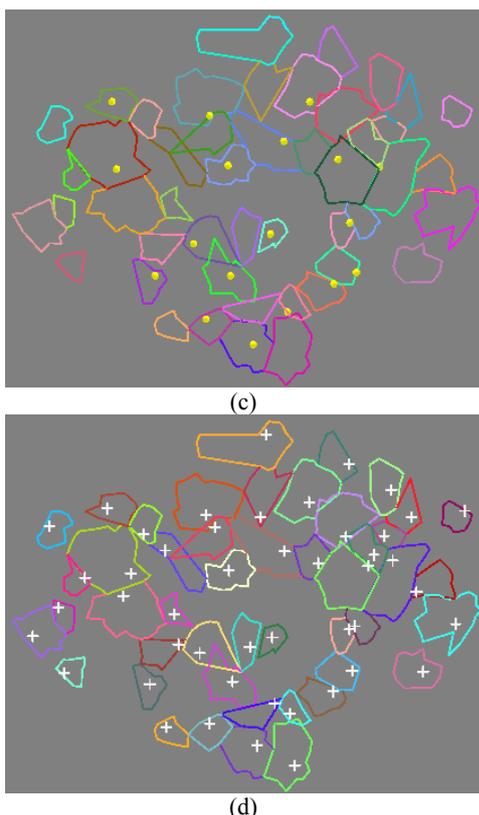


Figure 9. Extracted results with reference data:  
 (a) boundaries and ground truth in case I,  
 (b) boundaries and treetops in case I,  
 (c) boundaries and ground truth in case II,  
 (d) boundaries and treetops in case II.

		Case I	Case II
Ground truth		25	18
Detected results	Individual stands	19	18
	Merged stands	6	0
	Missed stands	0	0
Total stands		25	18
Unit: stands			

Table 1. Results of canopy extraction

Following the results of the canopy extraction, the proposed scheme can detect the individual trees as well as possible if the elevation changes are specific. Therefore, EMD can detect and locate those perturbations to extract the individual trees. Those sub-crowns are also identified by the proposed scheme as shown in Figure 9d.

#### 4. CONCLUSIONS

In the forestry management, the estimation of timber volume or biomass is an important task. This investigation proposes the scheme to extract individual canopies and shows the results that have the capacity to indicate the canopies and sub-canopies. The tree position, tree height, and canopy cover are identified.

There are some small sub-canopies around the main crown. The authors would like to improve the investigation to merge them into the main canopies. In addition, to add the test data set for further estimation is also one of our future tasks.

#### ACKNOWLEDGEMENT

The laser scanning data was derived from European Spatial Data Research (EuroSDR). The data sets were the test material of one of the projects, Tree Extraction. The authors would thank this institution to provide the data for the preliminary test.

#### REFERENCES

Axelsson, P., 2001. Ground estimation of laser data using adaptive TIN models, *Proceedings of OEEPE workshop on airborne laserscanning and interferometric SAR for detailed digital elevation models*, Royal Institute of Technology, Stockholm, Sweden, Vol. 40, CD-ROM, pp. 185-208.

Bortolot, Z.J., 2006. Using tree clusters to derive forest properties from small footprint LIDAR data, *Photogrammetric Engineering & Remote Sensing*, 72(12), pp.1389-1397.

Chen, J.S., Yu, D.J., and Yu, Y., 2007. the application of energy operator demodulation approach based on EMD in machinery fault dianosis, *Mechanical Systems and Signal Processing*, 21(2), pp.668-677.

Chen, S.H., Su, H.B., Zhang, R.H., Tian, J.,2008. Fusing remote sensing images using a trous wavelet transform and empirical mode decomposition, *Pattern Recognition Letters*, 29(3), pp.330-342.

Clark, M.L., Clark, D.B., and Roberts, D.A., 2004. Small-footprint lidar estimation of sub-canopy elevation and tree height in a tropical rain forest landscape, *Remote Sensing of Environment*, 91, pp.68-89.

Gougeon, F.A., and Leckie, D.G., 2006. The individual tree crown approach applied to Ikonos images of a coniferous plantation area, *Photogrammetric Engineering & Remote Sensing*, 72(11), pp. 1287-1297.

Hirschmugl, M., Ofner, M., Raggam, J., and Schardt, M., 2007. Single tree detection in very high resolution remote sensing data, *Remote Sensing of Environment*, 110, pp.533-544.

Hosoi, F., and Omasa, K., 2006. Voxel-based 3-D modeling of individual trees for estimating leaf area density using high-resolution portable scanning lidar, *IEEE Transactions on Geoscience and Remote Sensing*, 44(12), pp.3610-3618.

Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N.C., Tung, C.C., and Liu, H.H., 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, *Proceedings of the royal society of london series a-mathematical physical and engineering sciences*, 454(1971), pp. 903-995.

Hyypää, J., Kelle, O., Lehtikoinen, M., and Inkinen, M., 2001. A segmentation-based method to retrieve stem volume estimates from 3-D tree height models produced by laser scanners, *IEEE Transactions on Geoscience and Remote Sensing*, 39(5), pp. 969-975.

Maltamo, M., Hyypä, J., and Malinen, J., 2006. A comparative study of the use of laser scanner data and field measurements in the prediction of crown height in boreal forests, *Scandinavian Journal of Forest Research*, 21, pp.231-238.

Næsset, E., Gobakken, T., Holmgren, J., Hyypä, H., Hyypä, J., Maltamo, M., Nilsson, M., Olsson, H., Persson, A., and Söderman, U., 2004. Laser scanning of forest resources: the Nordic experience, *Scandinavian Journal of Forest Research*, 19, pp.482-499.

Næsset, E., Gobakken, T., Holmgren, J., Hyypä, H., Hyypä, J., Maltamo, M., Nilsson, M., Olsson, H., Persson, A., and Söderman, U., 2004. Laser scanning of forest resources: the Nordic experience, *Scandinavian Journal of Forest Research*, 19, pp.482-499.

Nelson, R., 1984. Determining forest canopy characteristics using airborne laser data, *Remote Sensing of Environment*, 15, pp.201-212.

Walker, W.S., Kellndorfer, J.M., LaPoint, E., Hoppus, M., Westfall, J., 2007. An empirical InSAR- optical fusion approach to mapping vegetation canopy height, *Remote Sensing of Environment*, 109: pp. 482-492.

Wu, Z.H., Huang, N.E., Long, S.R., Peng, C.K., 2007. On the trend, detrending, and variability of nonlinear and nonstationary time series, *Proceedings of the National Academy of Sciences of the United States of America*, 104(38), pp.14889-14894.

Wulder, M.A., Han, T., White, J.C., Sweda, T., and Tsuzuki, H., 2007. Integrating profiling LIDAR with Landsat data for regional boreal forest canopy attribute estimation and change characterization, *Remote Sensing of Environment*, unpaginated.