The European Information Society
Taking Geoinformation Science One Step Further
The Association of Geographic Information Laboratories for Europe (AGILE) was established in early 1998 to promote academic teaching and research on GIS at the European level. AGILE seeks to ensure that the views of the geographic information teaching and research community are fully represented in the discussions that take place on future European research agendas and it also provides a permanent scientific forum where geographic information researchers can meet and exchange ideas and experiences at the European level.

In 2007 AGILE provided - for the first time since its existence - a book constituting a collection of scientific papers that were submitted as full-papers to the annual AGILE conference and went through a competitive and thorough review process. Published in the *Springer Lecture Notes in Geoinformation and Cartography* this first edition was well received within AGILE and within the European Geoinformation Science community as a whole. Thus, the decision was easily made to establish a Springer Volume for the 11\textsuperscript{th} AGILE conference held 2008 in Girona, Spain, and led to what you now hold in your hands.

The 11\textsuperscript{th} AGILE call for full-papers of original and unpublished fundamental scientific research in all fields of Geoinformation Science resulted in 54 submissions, of which 23 were accepted for publication in this volume (acceptance rate 43\%). These figures are similar to those of the 2007 volume and indicate that having full-paper submissions leading to an annual high-quality scientific edition is a promising model for following AGILE conferences.

The scientific papers published here, cover a number of basic topics within Geoinformation Science. The papers included in this book span fundamental aspects of geoinformation processing: Measuring spatiotemporal phenomena, quality and semantics of geoinformation, spatiotemporal analysis, spatiotemporal modelling and decision support, and spatial information infrastructures. We believe that the papers comprise innovative research and take Geoinformation Science one step further.

Organising the programme of an International Conference and simultaneously editing a volume of scientific papers necessarily requires time and effort. We therefore would like to gratefully acknowledge the efforts of the
authors and reviewers of this book, who in adhering to a strict timetable, helped to finalise this book so that it could be delivered at the AGILE 2008 conference. We thank the local chair Irene Compte (University of Girona) and her team in Girona for giving all kind of local support to make the AGILE 2008 conference happen. Veronica Schemien (TU Dresden) put a lot of effort by supporting us with the editing of this book. Additionally, we would like to thank Definiens AG, ESRI Inc., and Intergraph Corporation who for years have sponsored the annual AGILE conferences. Last but not least, our thanks go to Agata Oelschläger (Springer) who helped us to finalise this book in time.

*Lars Bernard, Anders Friis-Christensen, Hardy Pundt*

February 2008
Programme Committee

Programme Chair Lars Bernard,
TU Dresden (Germany)

Programme Co-Chair Anders Friis-Christensen,
European Commission, Joint Research Centre (Italy)

Programme Co-Chair Hardy Pundt,
University of Applied Sciences Harz (Germany)

Local Committee

Local Chair Irene Compte
University of Girona (Spain)

Scientific Committee

Itzak Benenson, Tel Aviv University (Israel)
Michela Bertolotto, University College Dublin (Ireland)
Lars Bodum, University of Aalborg (Denmark)
Arnold Bregt, Wageningen University (Netherlands)
Christoph Brox, University of Münster (Germany)
Gilberto Camara, INPE (Brazil)
Nicholas Chrisman, Laval University (Canada)
Christophe Claramunt, Naval Academy Research Institute (France)
Arzu Çöltekin, University of Zurich (Switzerland)
Helen Couclelis, University of California - Santa Barbara (USA)
Max Craglia, European Commission, Joint Research Centre (Italy)
Isabel Cruz, University of Illinois - Chicago (USA)
Leila De Floriani, University of Genova (Italy)
Pasquale diDonato, University of Rome - La Sapienza (Italy)
Sara Fabrikant, University of Zurich (Switzerland)
Peter Fisher, University of Leicester (United Kingdom)
Anders Friis-Christensen, European Commission, Joint Research Centre (Italy)
Fred Toppen, University of Utrecht (Netherlands)
Peter van Oosterom, Delft University of Technology (Netherlands)
Danny Vandenbroucke, Katholieke Universiteit Leuven (Belgium)
Monica Wachowicz, Wageningen University (Netherlands)
Rob Weibel, University of Zurich (Switzerland)
Stephan Winter, University of Melbourne (Australia)
Andreas Wytzisk, 52 North (Germany)
May Yuan, University of Oklahoma (USA)
Javier Zarazaga-Soria, University of Zaragoza (Spain)
Contributing Authors

Alexander Almer
Joanneum Research, Austria

Arda Alp
PhD Student: Artificial intelligence Laboratory - Ecole Polytechnique Federale Lausanne, Switzerland

Peter Bachhiesl
School of Telematics/Network Engineering, Carinthia University of Applied Sciences, Austria

Alberto Belussi
University of Verona, Italy

Michela Bertolotto
School of Computer Science & Informatics, University College Dublin, Ireland

Elena Camossi
School of Computer Science & Informatics, University College Dublin, Ireland

Claudio Carneiro
PhD Student: GIS LAB - Ecole Polytechnique Federale Lausanne, Switzerland

Martin Charlton
National Centre for Geocomputation, National University of Ireland Maynooth, Co Kildare, Ireland

Diego Díaz Doce
The University of Edinburgh, United Kingdom

Jürgen Döllner
Hasso-Plattner-Institute, University of Potsdam, Germany

Martin Espeter
University of Münster, Germany

Peter Foley
National Centre for Geocomputation, National University of Ireland Maynooth, Co Kildare, Ireland

A Stewart Fotheringham
National Centre for Geocomputation, National University of Ireland Maynooth, Co Kildare, Ireland

Andrew U. Frank
Vienna University of Technology, Institute for Geoinformation and Cartography, Austria

Mauro Gaio
LIUPPA, France

François Golay
Swiss Federal Institute of Technology, Switzerland

Klaus Granica
Joanneum Research, Austria
Gerald Gruber
Fachhochschule Kaernten, Austria

Alex Hagen-Zanker
Urban Planning Group, Technische Universität Eindhoven, Netherlands

Henning Sten Hansen
Aalborg University, Denmark

Brent Hecht
University of California, Santa Barbara, United States

Maarten Hilferink
Object Vision BV, CIMO-Vrije Universiteit Amsterdam, Netherlands

Manuela Hirschmugl
Joanneum Research, Austria

Hartwig Hochmair
University of Florida, United States

Jens Ingensand
Swiss Federal Institute of Technology, Switzerland

Krzysztof Janowicz
University of Münster, Germany

Markus Jobst
Vienna University of Technology, Austria

Farid Karimipour
Vienna University of Technology, Institute for Geoinformation and Cartography, Austria

Tahar Kechadi
School of Computer Science & Informatics, University College Dublin, Ireland

Carsten Kessler
University of Munster, Germany

Eric Koomen
Vrije Universiteit Amsterdam FEWEB/SPINlab, Netherlands

Martin Krch
School of Geoinformation, Carinthia University of Applied Sciences, Austria

Federica Liguori
GIS Expert, Italy

Willem Loonen
Netherlands Environmental Assessment Agency (MNP), Netherlands

Haik Lorenz
Hasso-Plattner-Institute, University of Potsdam, Germany

Pierre Loustau
LIUPPA, France

Jose Macedo
PosDoc researcher: Database Laboratory - Ecole Polytechnique Federale Lausanne, Switzerland

Jody Marca
Politecnico di Milano, Italy

Christian Menard
Fachhochschule Kaernten, Austria
Hossein Mohammadi  
The University of Melbourne,  
Geomatics Department, Australia

Abbas Rajabifard  
The University of Melbourne,  
Geomatics Department, Australia

Gerhard Navratil  
Vienna University of Technology,  
Institute for Geoinformation and  
Cartography, Austria

Martin Raubal  
University of California, Santa  
Barbara, United States

Mauro Negri  
Politecnico di Milano, Italy

Bernhard Schachinger  
Fachhochschule Kärnten, Austria

Thierry Nodenot  
LIUPPA, France

Thomas Schnabel  
Joanneum Research, Austria

Tonny Oyana  
Southern Illinois University, United  
States

Johannes Scholz  
School of Geoinformation, Carinthia  
University of Applied Sciences, Austria

Ilija Panov  
University of Münster, Germany

Mirco Schwarz  
University of Münster, Germany

Genevieve Patenaude  
The University of Edinburgh,  
United Kingdom

Kara Scott  
Southern Illinois University, United  
States

Gernot Paulus  
School of Geoinformation, Carinthia  
University of Applied Sciences, Austria

Stefano Spaccapietra  
Senior researcher (LAB's responsible)  
and full professor: Database  
Laboratory - Ecole Polytechnique  
Fédérale Lausanne, Switzerland

Giuseppe Pelagatti  
Politecnico di Milano, Italy

Emmanuel Stefanakis  
Harokopio University of Athens,  
Greece

Michal Petr  
Forest Research, United Kingdom

Juan Suárez  
Forestry Commission, United  
Kingdom

Johann Raggam  
Joanneum Research, Austria
Harry Timmermans
Urban Planning Group, Technische Universität Eindhoven, Netherlands

Matthias Trapp
Hasso-Plattner-Institute, University of Potsdam, Germany

Michael van Dahl
VCS Aktiengesellschaft, Austria

Alexander C. Walkowski
Institute for Geoinformatics, University of Münster, Germany

Marc Wilkes
University of Münster, Germany

Ian Williamson
The University of Melbourne, Geomatics Department, Australia
Table of Contents

Measuring Spatiotemporal Phenomena

Forest Stand Volume of Sitka Spruce Plantations in Britain: Can Existing Laser Scanning Methods Based on the Conventional One Provide Better Results, a Comparison of Two Approaches
    Michal Petr, Genevieve Patenaude, Juan Suárez .............................................1

Assessing Stand and Data Variability Using Airborne Laser Scanner
    Diego D. Doce, Juan C. Suárez, Genevieve Patenaude...........................................27

Model Based Optimization of Mobile Geosensor Networks
    Alexander C. Walkowski ..........................................................................................51

Quality and Semantics of Geoinformation

Evaluation of the Geometric Accuracy of Automatically Recorded 3D – City Models Compared to GIS-Data
    Gerald Gruber, Christian Menard, Bernhard Schachinger.................................67

Lifting Imprecise Values
    Gerhard Navratil, Farid Karimipour, Andrew U. Frank..................................79

GeoSR: Geographically Explore Semantic Relations in World Knowledge
    Brent Hecht, Martin Raubal..................................................................................95

A Study on the Cognitive Plausibility of SIM-DL Similarity Rankings for Geographic Feature Types
    Krzysztof Janowicz, Carsten Keßler, Ilija Panov, Marc Wilkes,
    Martin Espeter, Mirco Schwarz ............................................................................115
Spatiotemporal Analysis

A Geospatial Implementation of a Novel Delineation Clustering Algorithm Employing the $K$-means
Tonny J. Oyana, Kara E. Scott ........................................................................ 135

DBSCAN-MO: Density-Based Clustering among Moving Obstacles
Emmanuel Stefanakis ...................................................................................... 159

A Metric of Compactness of Urban Change Illustrated to 22 European Countries
Alex Hagen-Zanker, Harry Timmermans ..................................................... 181

Advanced Data Mining Method for Discovering Regions and Trajectories of Moving Objects: “Ciconia Ciconia” Scenario
Claudio Carneiro, Arda Alp, Jose Macedo, Stefano Spaccapietra ................. 201

Mining Spatio-Temporal Data at Different Levels of Detail
Elena Camossi, Michela Bertolotto, Tahar Kechadi .................................... 225

Automated Boundary Creation: Atomic Small Areas in Ireland
A Stewart Fotheringham, Peter F Foley, Martin Charlton .............................. 241

Spatiotemporal Modelling and Decision Support

Climate-Change Adaptations in Land-Use Planning; A Scenario-Based Approach
Eric Koomen, Willem Loonen, Maarten Hilferink ........................................ 261

Quantifying and Analysing Neighbourhood Characteristics Supporting Urban Land-Use Modelling
Henning Sten Hansen .................................................................................. 283

Interactive Multi-Perspective Views of Virtual 3D Landscape and City Models
Haik Lorenz, Matthias Trapp, Jürgen Döllner, Markus Jobst ....................... 301
Scenario-Based Spatial Decision Support for Network Infrastructure Design
   Gernot Paulus, Martin Krch, Johannes Scholz, Peter Bachhiesl .................. 323

Grouping of Optimized Pedestrian Routes for Multi-Modal Route Planning: A Comparison of Two Cities
   Hartwig H. Hochmair ................................................................................. 339

Spatial Decision Support in the Pedagogical Area: Processing Travel Stories to Discover Itineraries Hidden Beneath the Surface
   Pierre Loustau, Thierry Nodenot, Mauro Gaio ............................................ 359

Spatial Information Infrastructures

Ownership Definition and Instances Integration in Highly Coupled Spatial Data Infrastructures
   Alberto Belussi, Federica Liguori, Jody Marca, Mauro Negri, Giuseppe Pelagatti .......................................................... 379

Spatial Data Integrability and Interoperability in the Context of SDI
   Hossein Mohammadi, Abbas Rajabifard, Ian Williamson .................................... 401

Information Services to Support Disaster and Risk Management in Alpine Areas
   Alexander Alme, Thomas Schnabel, Klaus Granica, Manuela Hirschmugl, Johann Raggam, Michael van Dahl .......................................................... 415

User Performance in Interaction with Web-GIS: A Semi-Automated Methodology Using Log-Files and Streaming-Tools
   Jens Ingensand, François Golay .................................................................. 433

Index .............................................................................................................. 445
Forest Stand Volume of Sitka Spruce Plantations in Britain: Can Existing Laser Scanning Methods Based on the Conventional One Provide Better Results, a Comparison of Two Approaches

Michal Petr$^{1,2}$, Genevieve Patenaude$^1$, Juan Suárez$^2$

$^1$ Institute of Geography, School of Geosciences, The University of Edinburgh, Edinburgh EH8 9XP, UK, M.Petr@sms.ed.ac.uk, genevieve.patenaude@ed.ac.uk
$^2$ Forest Research, Northern Research Station, Roslin, Midlothian EH25 9SY, UK, michal.petr@forestry.gsi.gov.uk, juan.suarez@forestry.gsi.gov.uk

Abstract. This paper looks at different datasets obtained from an airborne Light Detection And Ranging (LiDAR) system and compares the reliability of two contemporary analysis approaches. Estimates of different stand parameters, such as top tree height, were derived using regression analysis and a segmentation approach on data obtained from small-footprint laser scan were contrasted with the field measurements in 7 plots, specifically volume and basal area. Plots of 2,500 m$^2$ containing plantations of Sitka spruce (Picea sitchensis Bong. Carr.) were scanned with two different point densities in years 2003 and 2004. These plots were divided into training and test regions of 625 m$^2$ each. Regression analysis was performed using percentiles corresponding to the canopy tree height at different vertical levels and a segmentation method was used to delineate individual tree crowns where tree metrics can be determined. The bias of the estimated values for the stand volume and basal area ranged from 1.21 to 6.49 m$^3$ha$^{-1}$ (0.17 to 0.92 %) and -2.69 to 1.23 m$^2$ha$^{-1}$ (-3.9 to 1.7 %), respectively; and the bias calculated from the segmentation using 0.5 and 1m dataset ranged between -349.77 to -434.76 m$^3$ha$^{-1}$ (-49.7 to -61.8 %) for the stand volume and -33.36 to -42.24 m$^2$ha$^{-1}$ (-48.5 to -61.4 %) for the basal area. The results showed that the regression models estimated stand volume and basal more accurately compared with values calculated from the segmentation. Furthermore, it is shown that there was no significant difference in the estimates from the regression model when using different point densities.
Keywords: forest, stand volume, basal area, LiDAR, segmentation

1 Introduction

The forest is a dynamic and complex ecosystem that represents an integral part of world’s environment providing essential elements to different spheres of our society. In the United Kingdom, coniferous forest covers more than 57% of the total woodland (Forestry Statistics 2006), and the current growing stock is important for a sustainable forest management. The need for faster methods of forest inventory compared with the conventional field measurements has motivated the employment of remote sensing techniques that can cover large areas of woodland forest and with fewer constraints such as manpower. Except for the common techniques such as aerial photography, new systems such as Light Detection and Ranging and radar are currently being applied.

Many studies have used airborne laser scanning (ALS) for the estimation of various forest stand metrics, mainly mean tree height, stand basal area, and stand volume (Naesset 1997a; Holmgren 2004; Hyppa et al. 2001; Maltamo, Eerikainen et al. 2004). The most significant advantage of LiDAR is its ability to provide almost three-dimensional information about forest canopy structures (Naesset 2004). Data gathered by ALS systems can provide complete coverage as opposed to the sampling normally achieved with the field data collection. The width of study area will depend on the flight altitude and scanning angle (Hyppa et al. 2005).

LiDAR systems have been frequently employed in forest studies because of their ability to provide both horizontal and vertical information about trees, whereas most optical sensors are only capable of providing horizontal information (Lim, Flood et al. 2003; Donoghue and Watt 2006). Another significant advantage of LiDAR sensing is that it readily obtains specific information about trees, such as canopy height, very accurately (Naesset 2004). Even though the cost of obtaining LiDAR data is high (Tilley et al. 2004), it gives an essential and effective tool for the estimation of tree characteristics; and as this data can be also adopted to various other purposes, a reduction in the cost can be expected with wider application.

The use of airborne LiDAR systems in forestry focuses primarily on the determination of tree metrics in the stand level for inventory purposes (Naesset 2004; Holmgren 2004; Coops et al. 2007; Naesset 1997a) or at the tree level for measurement of single trees (Hyppa et al. 2005; Maltamo, Yu et al. 2004; Peuhkurinen et al. forthcoming). Airborne LiDAR
with small footprint and low point density (about 1 point/m$^2$) has proven to be an efficient tool at the stand level in the determination of tree metrics (Holmgren 2004). However, at the tree level a point density of at least 5 points/m$^2$ is required to achieve good results for tree segmentation and to provide sufficient information about individual trees (Maltamo et al. 2005).

In forest inventories, where information is related to the stand, several parameters such as mean tree height or basal area are necessary to determine. Previously airborne LiDAR has effectively and accurately determined mean tree height (Naesset 1997a; Holmgren, Nilsson, and Olsson 2003; Hyyppä et al. 2000), basal area (Hyyppä et al. 2000; Lim, Treitz et al. 2003), stand volume (Naesset 1997b; Means et al. 2000), and dominant tree height (Lim, Treitz et al. 2003). When airborne discrete LiDAR data was used for stand volume estimation it was shown that the accuracy varies for different tree species (Naesset 1997b). Reasons for this variation are mainly due to crown shape and the type of vegetative material (leaves or needles) of specific species. Furthermore, the scanning point density influences the capability of LiDAR to detect tree tops; however, in most cases it omitted the correct tree top (Suárez et al. 2005) and therefore underestimated tree height (Nilsson 1996; Naesset 1997b). Many studies also focused on the estimation of dominant tree height of young forests and proved that mature and even young trees can be accurately estimated by LiDAR (Naesset and Bjerknes 2001).

Another method to determine forest metrics from LiDAR data is the single-tree level approach. This approach is based on tree delineation algorithms that identify, locate, and relate tree characteristics to a single tree. Several studies have suggested and used high scanning point densities from 10 - 20 points/m$^2$ for the detection of single trees (Maltamo, Yu et al. 2004; Hyyppä et al. 2001). The canopy height model (CHM) which gives the canopy height as the difference between the Digital Elevation Model (DEM) and the Digital Surface Model (DSM) is derived from LiDAR data and has been used for tree segmentation. Segmentation algorithms have been primarily used for the detection of trees from the CHM where local maxima were given by tree tops in segments. The retrieved height using this method provided reliable results and a recent study has shown an underestimation of only 0.97 m (Maltamo, Yu et al. 2004). The stem volume of an individual tree has been calculated from volume equations that used predetermined variables, tree height and crown diameter (assumed to be a circular shape) from the CHM and predicted stem diameter as well (Naesset et al. 2004). The study of (Persson, Holmgren, and Söderman 2002) presented estimates of stem volume with RMSE of 0.21 m$^2$, corresponding to 22% of the field stem volume. The main advantage of the tree
segmentation compared with the stand approach is its capability of identifying spatial heterogeneity within a stand where mean variables are estimated. However, these algorithms are developing and still need improvement.

The aim of this study is to estimate the stand volume of coniferous trees with the use of airborne laser scanning data and to compare the accuracy of this estimate with ground measurements. The main objectives pursued in this study are: 1) to determine whether a relationship between the stand volume and canopy heights obtained from LiDAR data exists for our study site; 2) to determine how precisely it is possible to estimate the stand volume of coniferous trees from LiDAR data in the study area; 3) to evaluate the effect of different hit densities on the estimation of stand volume; and 4) to determine the differences of volume estimates between a regression based method and a segmentation method at the stand.

This paper is divided into 5 sections, each focusing on a different part of the research. The following section describes available field and LiDAR datasets, and also deals with an applied segmentation algorithm. Section 3 then suggests an appropriate methodology which should be used whereas section 4 will presents the final results of the study. Finally, section 5 will provide discussion of the results and overall conclusions.

2 Materials

2.1 Study Site

The study area of this project is the Kielder Forest District located in northern England (Fig. 1) (55° 14' N 2° 35' W), and is managed and owned by the UK Forestry Commission (FC). The area is characterised mainly by low hills with altitudes between 30 m and 600 m, and a mean slope angle of 6°. Sitka spruce (*Picea sitchensis* Bong. Carr.) is the dominant species in the area followed by other species like Norway spruce (*Picea abies* (L.) H. Karst) and Lodgepole pine (*Pinus contorta* Douglas). This study focuses on the Sitka spruce plantations located within the study plots.
2.2 Field Data

The field work in the study area was carried out by the Forestry Commission between February and April 2003. Field data was obtained from 7 study plots with dimensions 50 m x 50 m. These plots were located separately in three different sites with similar climate conditions. In the first location the trees were 64 years old, in the second location trees of 33 prevailed and in the third location the dominant tree age was 36 (Table 1). Each study plot and an individual tree within each plot were accurately located with the differential GPS and a Total station. Tree parameters collected consisted of tree height, diameter at breast height (DBH) and dominance type. Furthermore, for selected trees, additional information about tree crown dimensions in N-S and E-W directions and height to the first live whorl was registered.

Then, the reference data, which is presented in Table 1, was calculated. First, top height ($t_h$) was calculated as the average of the 100 largest trees
with the largest DBH per hectare as in (Philip 1994); however, in this study only 25 trees per plot were used. Second, the mean diameter at breast height \( (m_{\text{DBH}}) \) was computed as an average of all living trees. Next, the basal area of a plot \( (G) \) was computed as the sum of all individual basal areas. The Crop form coefficients were derived from the top height of the study plots. Finally, the stand volume \( (V) \) was computed as the sum of stem volumes of all living trees located within a study plot. The stem volume for each tree was calculated as basal area multiplied by a crop form coefficient for Sitka spruce from (Hamilton 1975).

**Table 1** Summary of the field plot reference data (50 m x 50 m)

<table>
<thead>
<tr>
<th>Plot</th>
<th>( m_{\text{DBH}} )</th>
<th>( m_{\text{DBH}} )</th>
<th>( m_{\text{DBH}} )</th>
<th>( m_{\text{DBH}} )</th>
<th>( m_{\text{DBH}} )</th>
<th>( m_{\text{DBH}} )</th>
<th>( m_{\text{DBH}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plot1</td>
<td>32.5</td>
<td>948</td>
<td>32.5</td>
<td>82.6</td>
<td>1209.6</td>
<td>64</td>
<td></td>
</tr>
<tr>
<td>Plot2</td>
<td>27.5</td>
<td>1168</td>
<td>25.0</td>
<td>61.7</td>
<td>748.0</td>
<td>64</td>
<td></td>
</tr>
<tr>
<td>Plot3</td>
<td>19.8</td>
<td>2024</td>
<td>18.2</td>
<td>62.8</td>
<td>509.3</td>
<td>38</td>
<td></td>
</tr>
<tr>
<td>Plot4</td>
<td>19.9</td>
<td>2040</td>
<td>18.1</td>
<td>62.0</td>
<td>521.6</td>
<td>38</td>
<td></td>
</tr>
<tr>
<td>Plot5</td>
<td>17.3</td>
<td>1980</td>
<td>19.9</td>
<td>59.4</td>
<td>476.2</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td>Plot6</td>
<td>20.6</td>
<td>2792</td>
<td>17.3</td>
<td>72.6</td>
<td>611.9</td>
<td>36</td>
<td></td>
</tr>
<tr>
<td>Plot7</td>
<td>21.0</td>
<td>2196</td>
<td>20.0</td>
<td>72.4</td>
<td>651.6</td>
<td>33</td>
<td></td>
</tr>
</tbody>
</table>

\( t_h \) top height, \( N \) stem number, \( m_{\text{DBH}} \) mean diameter at breast height, \( G \) stand basal area, \( V \) stand volume.

Due to the small number of reference plots available for this study, each 50 m x 50 m study plot was split into four sub-plots of 25 m x 25 m (625 \( m^2 \)). These 28 sub-plots were randomly divided into two same size groups of training and test plots, where each group always contained two sub-plots from the original plot. The stand volume and the basal area were the only forest metrics computed in this study with the same methods as above. For the calculation, just the trees registered within a sub-plot were used. A summary of training and test sub-plots for the basal area and the stand volume is presented in Table 2.
Table 2 Summary of the sub-plots reference data (25 m x 25 m)

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Range</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Training plots (n = 14)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G [m² ha⁻¹]</td>
<td>51.64 - 85.63</td>
<td>66.46</td>
</tr>
<tr>
<td>V [m³ ha⁻¹]</td>
<td>465.02 - 1210.81</td>
<td>678.21</td>
</tr>
<tr>
<td><strong>Test plots (n = 14)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G [m² ha⁻¹]</td>
<td>54.85 - 92.76</td>
<td>68.79</td>
</tr>
<tr>
<td>V [m³ ha⁻¹]</td>
<td>466.72 - 1311.65</td>
<td>703.79</td>
</tr>
</tbody>
</table>

G stand basal area, V stand volume.

2.3 Laser Scanner Data

LiDAR data for the study area was acquired in two years, 2003 (26th March) and 2004 (July). The contractor for data from year 2003 was UK Environmental Agency using an Optech ALTM 2033 LiDAR system which was the same system for 2004. Operating at 1064 nm the Optech 2033 is a discrete return system, recording only the first and the last laser returns. Detailed information about ALTM 2033 is presented in Table 3. The LiDAR data from 2003 was acquired with a 305 m wide swath and 70% side overlap, and the flight altitude was 905 m. The scanning angle at nadir was 20°, and produced distances of 1.8 m between scanning lines and 0.3 m between points. The provided LiDAR data was in ASCII format with the X, Y, Z coordinates in the British National Grid (BNG); the intensity of the laser pulse was included in this information. Height was stored in Z coordinate and referenced to the Ordnance Survey of Great Britain OSGB 1936 Datum (Donoghue and Watt 2006).

Table 3 Summary of the laser scanner

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensor</td>
<td>ALTM Optech 2033</td>
</tr>
<tr>
<td>Laser pulse frequency</td>
<td>33 kHz</td>
</tr>
<tr>
<td>Field of view</td>
<td>10° (20°)²</td>
</tr>
<tr>
<td>Beam divergence</td>
<td>0.3 Mrad</td>
</tr>
<tr>
<td>Horizontal accuracy</td>
<td>± 0.15 m</td>
</tr>
<tr>
<td>Vertical accuracy</td>
<td>0.60 m</td>
</tr>
<tr>
<td>Laser classification</td>
<td>Class IV laser product (FDA CFR 21)</td>
</tr>
</tbody>
</table>

² for LiDAR data acquired in 2003.
The LiDAR data from 2003 and 2004 have different point densities. The density of LiDAR data from year 2003 is lower (above 2 points per m²) when compared with the LiDAR 2004 data with point density of about 7 points per m² (Table 4). Nevertheless, there were almost no differences in point densities between training and test plots.

Table 4 Sampling density of laser scanner data

<table>
<thead>
<tr>
<th></th>
<th>No. of observations</th>
<th>No. of transmitted pulses&lt;sup&gt;a&lt;/sup&gt; (ha&lt;sup&gt;-1&lt;/sup&gt;)</th>
<th>Mean point density (points/m²)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>LiDAR 2003</strong></td>
<td>Training plots</td>
<td>14 12192 – 35584</td>
<td>2.29</td>
</tr>
<tr>
<td></td>
<td>Test plots</td>
<td>14 11680 – 43808</td>
<td>2.38</td>
</tr>
<tr>
<td><strong>LiDAR 2004</strong></td>
<td>Training plots</td>
<td>14 41440 – 96960</td>
<td>7.01</td>
</tr>
<tr>
<td></td>
<td>Test plots</td>
<td>14 39456 – 98976</td>
<td>6.89</td>
</tr>
</tbody>
</table>

<sup>a</sup> refers to first pulse data

The last laser points of the LiDAR 2004 data were filtered and classified using Terrascan software (Terrasolid, Finland) to determine and separate the ground points from the vegetation hits. Only the last laser points from 2004 data assumed to represent ground were used to generate two grids with 0.5 m and 1 m spacing using the Surfer software (Goldensoftware, Golden, Colorado, USA). Then, these grids were interpolated by method of Delaunay triangulation with linear interpolation. The final DEM was created from previous grids in ArcGIS using the Nearest Neighbor interpolation algorithm with predefined 0.5 m and 1.0 m pixel sizes. However, the DEM from 2003 was delivered already complete using Treesvis software (Weinacker, 2004) with 1.0 m pixel size.

From the original 2003 and 2004 LiDAR datasets using only the first laser returns, new data was derived for further analysis. All LiDAR points in the study area were registered to a specific DEM based on their coordinates in the OS BNG. The relative height of each registered LiDAR point was calculated as a difference between the DEM and elevation value (Z). The same approach for the determination of relative height has been employed in many studies (Naesset 2004; Naesset and Okland 2002). In some cases the relative height of several LiDAR points was low caused mainly by ground vegetation, such as grasses and shrubs. Hence, the values of relative height below 2 m, which represented low shrubs or stones were excluded from further analysis (Naesset 1997a; Nilsson 1996).
2.4 Single-Tree Segmentation

The purpose of the segmentation is to delineate individual tree crowns and to derive information about tree height and other important parameters. The input data necessary for the segmentation is the CHM consisting of the DEM and the DSM. The DSM was assumed to represent the vegetation cover which was created only from the first laser returns (the first point reflected from the top of objects) of the LiDAR 2004 data. Similar steps were followed as in the creation of DEM for the DSM, with emphasis on vegetation using Surfer software operating with pixel sizes of 0.5 m and 1.0 m. The surface layers, corresponding to the ground as the DEM and the vegetation cover as the DSM, were afterwards employed for the calculation of the CHM. This model giving the relative height of vegetation was calculated on a per pixel basis where each pixel from the DEM was subtracted from the DSM (Fig 2.).

The segmentation algorithm that was applied to the high-resolution CHM was previously introduced by (Pitkänen et al. 2004; Pitkänen 2005) for the segmentation of coniferous trees. The high point density LiDAR 2004 dataset (about 7 points/m$^2$) was exclusively used as it had shown success in recent studies (Hyyppä et al. 2001; Peuhkurinen et al. forthcoming).

The first step in the segmentation process was the filtering of the CHM which utilized a height-based filtering method dependent on the degree of low-pass Gaussian filter with used input value of 0.8 to reduce the effects of noise caused by branches (Peuhkurinen et al. forthcoming). Afterwards, the local maxima were detected and assumed to be the tree tops in the filtered image. The individual tree crowns were then discriminated using a watershed algorithm with a drainage direction (Pitkänen 2005); segments with local maxima (i.e. tree height) below 2 m were discarded. Each segment or individual tree contains information, for example, tree height (local maxima), coordinates of the local maxima and centroid, the crown area, and the maximum crown diameter.
Methods

3.1 Estimation of the Stand Volume and the Basal Area with Percentiles

The relative height values of the vegetation, calculated from derived LiDAR data and the DEM, was spatially registered to 14 training and 14 test plots (25 m x 25 m); data outside these plots was excluded from further analysis.

A multiple regression analysis was used to determine the relationship between field and laser derived data, with a final analysis applied for the estimation of the stand volume and basal area. First, the creation of the regression models employed independent variables represented by the mean height value, percentiles of 10, 20, ..., 100% and also two additional of 95 and 99% of the relative height values. Similar statistics for the laser derived relative height data were introduced in many previous studies (Naesset 2004; Holmgren 2004; Nelson, Krabill, and Tonelli 1988; Naesset 1997a; Naesset 2002). An additional independent variable for the age of trees, as it varies among plots, was also tested. So in total, this study employed and evaluated 13 independent candidates to find the best predic-
tor variable or variables using multiple linear regression with the first model, Model (1), formulated as:

\[ Y = \beta_0 + \beta_1 h_{10} + \beta_2 h_{20} + \ldots + \beta_{10} h_{100} + \beta_{11} h_{95} + \beta_{12} h_{99} + \beta_{13} h_{\text{mean}} \] (1)

Including the additional variable for age gives Model (2) as:

\[ Y = \beta_0 + \beta_1 h_{10} + \beta_2 h_{20} + \ldots + \beta_{10} h_{100} + \beta_{11} h_{95} + \beta_{12} h_{99} + \beta_{13} h_{\text{mean}} + \beta_{14} v_{\text{age}} \] (2)

where \( Y \) represents stand volume \( V \) (m\(^3\)ha\(^{-1}\)) or stand basal area \( G \) (m\(^2\)ha\(^{-1}\)); \( h_{10}, h_{20}, \ldots, h_{100}, h_{95}, h_{99} \) represent percentiles of relative height for 10%, 20%,...,100%, 95% and 99%; \( h_{\text{mean}} \) represents mean value for relative height; and \( v_{\text{age}} \) represents age of trees in the stand.

Selection of the most suitable independent variables for (1) and (2) regression models was based on the stepwise multiple selection similar to that used by (Naesset 2004). To determine whether they should be retained in a model, these variables were tested with a partial F statistic at significance level greater than 0.05; only variables below this significance level were included in the final prediction. In equation (2), age was excluded from testing with F statistics and entered as a second independent variable into a regression model.

Residuals of both models were studied which assisted to decide and improve the final regression models. Additionally, an analysis of variance of these models was considered with the emphasis on the principle of the least squares method using a minimum distance between predicted and observed values. Then the validation of the 2 previously created models (1) and (2) was done on 14 test plots. The limitation of these 2 models was that the testing plots were not spatially independent from the training plots. However, in this study it was the best available scenario which proved the usefulness of the studied models. Furthermore, these two models were used to estimate the stand volume and basal area, with the results compared against the field data.

In the end, the differences between estimated and observed values for the stand volume and basal area in the test plots were evaluated with a t-test for a significance. A paired two-tailed t-test was employed to compare if there was a significant difference between calculated means of estimated and observed values. Identical testing of this significance was previously carried out by (Naesset 2004).
3.2 Estimation of the Stand Volume and the Basal Area with Single Tree Segmentation

3.2.1 Individual Tree Parameters

The segmented high-resolution LiDAR images from 2004 with 0.5 m and 1.0 m spatial resolution were used for the determination of tree parameters. First, the created segments were spatially registered to the test plots based on the location of tree tops (local maxima), and only those segments within these plots were recorded for later analysis. The tree height for each segment was obtained as a maximum pixel value corresponding to a tree top. The crown diameter (CD) was calculated from every segment’s area (A) corresponding to the area of tree crown, using the following formula:

$$CD = \sqrt{\frac{4A}{\pi}}$$  \hspace{1cm} (3)

Another important parameter for the estimation of stem volume is the basal area and the DBH that was calculated using Forest Research Environment Database (FRED), which is a model based on the linear relationship between the DBH, tree height and crown diameter. Then, the calculated DBH was employed to compute basal area (BA) for each tree using the formula:

$$BA = \frac{\pi \cdot DBH^2}{40,000}$$  \hspace{1cm} (4)

where the value of 40,000 was used to relate the calculated BA to square meters. Finally, the stem volume for single tree was determined by the basal area and the crop form coefficient for Sitka spruce (Hamilton 1975).

Additionally, a linking process was made in ArcGIS (ESRI, Redlands, USA) for those trees detected by the segmentation algorithm. They were compared based on the location of their tree tops with the dominant and the co-dominant trees in each plot and then matched only with the closest one in the field. The detected trees with the distance less than 1.5 m to the field trees were included in further analysis. Metrics derived from the detected trees were compared against those in the field.

3.2.2 Stand Parameters

The estimates of stand parameters were determined by individual segmented trees located within the test plots. The stand volume (m$^3$ha$^{-1}$) and basal area (m$^2$ha$^{-1}$) from segmented trees were calculated as the sum of all
stem volumes and the sum of all basal areas for detected trees, respectively.

4 Results

4.1 Stand Level

The results of the regression analysis showed that only one independent variable derived from LiDAR datasets was suitable for both estimation of the stand volume and also of the stand basal area; therefore linear regression models were employed. The final regression models for the estimation of the stand volume were based on the 30th percentile of relative height \(h_{30}\) for both LiDAR datasets. However, when additional information about tree age was available in a training plot, the regression models containing age as a predictive variable \(v_{age}\) were used. The coefficient of determination \(R^2\) for Model (1) was 0.94 for the LiDAR 2003 dataset and 0.91 for 2004 (Table 5). When age was used as an additional predictive variable in Model (2), the coefficient of determination \(R^2\) increased for both datasets to 0.95 (LiDAR 2003) and 0.93 (LiDAR 2004). The accuracy of the predictive models stated by root mean squared error (RMSE) for the stand volume showed for Model (1) results of 75.84 and 78.40 in m\(^3\)ha\(^{-1}\) for LiDAR 2003 and LiDAR 2004 respectively, and for Model (2) showed higher accuracy with results of 73.44 and 76.48 respectively (Table 5).

The regression models were also employed to determine the stand basal area using the stepwise selection to find the best predictor. The 10% percentile of the relative height \(h_{10}\) was used as the most suitable independent predictor in all final models. A similar procedure for age as an additional variable was used as in the estimation of the stand volume. The range of the coefficients of determination \(R^2\) for both datasets and regression models varied from 0.64 to 0.82 (Table 5). The RMSE for the stand basal area in (m\(^2\)ha\(^{-1}\)) for Model (1) was 5.60 (LiDAR 2003) and 6.08 (LiDAR 2004); and for Model (2) was 4.48 and 5.44 respectively (Table 5).
Table 5 Representing relationship between ground values from training plots (25x25 m) and laser derived metrics from stepwise multiple regression analysis

<table>
<thead>
<tr>
<th>Dependent variable&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Predictive model&lt;sup&gt;b&lt;/sup&gt;</th>
<th>R&lt;sup&gt;2&lt;/sup&gt;</th>
<th>RMSE&lt;sup&gt;e&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>LiDAR 2003</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V (1)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>-11.079 + 3.459 * h&lt;sub&gt;30&lt;/sub&gt;</td>
<td>0.94</td>
<td>75.84</td>
</tr>
<tr>
<td>V (2)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>-10.517 + 3.883 * h&lt;sub&gt;30&lt;/sub&gt; + 0.163 * v&lt;sub&gt;age&lt;/sub&gt;</td>
<td>0.95</td>
<td>73.44</td>
</tr>
<tr>
<td>G (1)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>2.546 + 0.111 * h&lt;sub&gt;10&lt;/sub&gt;</td>
<td>0.64</td>
<td>5.60</td>
</tr>
<tr>
<td>G (2)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>2.763 + 0.166 * h&lt;sub&gt;10&lt;/sub&gt; -0.023 * v&lt;sub&gt;age&lt;/sub&gt;</td>
<td>0.80</td>
<td>4.48</td>
</tr>
<tr>
<td><strong>LiDAR 2004</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V (1)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>-12.927 + 3.415 * h&lt;sub&gt;30&lt;/sub&gt;</td>
<td>0.91</td>
<td>78.40</td>
</tr>
<tr>
<td>V (2)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>-12.6 + 3.726 * h&lt;sub&gt;30&lt;/sub&gt; -0.123 * v&lt;sub&gt;age&lt;/sub&gt;</td>
<td>0.93</td>
<td>76.48</td>
</tr>
<tr>
<td>G (1)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>2.595 + 0.113 * h&lt;sub&gt;10&lt;/sub&gt;</td>
<td>0.72</td>
<td>6.08</td>
</tr>
<tr>
<td>G (2)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>2.842 + 0.174 * h&lt;sub&gt;10&lt;/sub&gt; -0.025 * v&lt;sub&gt;age&lt;/sub&gt;</td>
<td>0.82</td>
<td>5.44</td>
</tr>
</tbody>
</table>

<sup>a</sup> V stand volume [m<sup>3</sup>ha<sup>-1</sup>], G stand basal area [m<sup>2</sup>ha<sup>-1</sup>].

<sup>b</sup> h<sub>10</sub>,h<sub>30</sub> percentiles of relative height for 10, 30 % (m), v<sub>age</sub> representing age of trees.

<sup>c</sup> Model (1) without age variable.

<sup>d</sup> Model (2) with age variable.

<sup>e</sup> RMSE for V [m<sup>3</sup>ha<sup>-1</sup>] and for G [m<sup>2</sup>ha<sup>-1</sup>].

The created regression models of (Table 5) were used afterwards for the estimation of the stand volume and the basal area on 14 separate test plots. The estimated values from the test plots were then compared with the field data. For the comparison of the difference between the estimated and field data, the mean difference (Mean) representing bias was used (Table 6). The stand volume derived from both LiDAR datasets overestimated the field values in Model (1) by 1.21 and 4.00 (m<sup>3</sup>ha<sup>-1</sup>), and in Model (2) by 4.35 and 6.49 (m<sup>3</sup>ha<sup>-1</sup>) (Table 6). However, values for the stand basal area also underestimated the field values in Model (1) from LiDAR 2003 dataset by 2.66 (m<sup>2</sup>ha<sup>-1</sup>) but overestimated them by LiDAR 2004 dataset by 0.03 (m<sup>2</sup>ha<sup>-1</sup>). Similarly, the results for G in Model (2) underestimated field values from LiDAR 2003 dataset by 2.69 (m<sup>2</sup>ha<sup>-1</sup>) but overestimated in the LiDAR 2004 dataset by 1.23 (m<sup>2</sup>ha<sup>-1</sup>). Overall, the estimated mean values of V and G were not significant on the predefined level of significance (p=0.05). The results of the estimation for both stand parameters are presented in Fig. 3 and Fig. 4.
Table 6  Difference (D) between estimated values from LiDAR derived metrics and ground measurement values for test plots (25x25 m)

<table>
<thead>
<tr>
<th>Variable a</th>
<th>Observed mean</th>
<th>Range</th>
<th>Mean</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>LiDAR 2003</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>703.79</td>
<td>-236.8 to 97.6</td>
<td>- 241.6 to 105.6</td>
<td>1.21 ns</td>
</tr>
<tr>
<td>G</td>
<td>68.79</td>
<td>-15.20 to 8.32</td>
<td>- 16.96 to 4.16</td>
<td>2.66 ns</td>
</tr>
<tr>
<td>LiDAR 2004</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>703.79</td>
<td>-219.2 to 116.8</td>
<td>- 220.8 to 123.2</td>
<td>4.00 ns</td>
</tr>
<tr>
<td>G</td>
<td>68.79</td>
<td>-11.84 to 11.68</td>
<td>- 12.32 to 8.64</td>
<td>0.03 ns</td>
</tr>
</tbody>
</table>

(1) Model (1) without age variable.
(2) Model (2) with age variable.
ns not significant (p> 0.05).
a V stand volume [m$^3$ha$^{-1}$], G stand basal area [m$^2$ha$^{-1}$].

Fig. 3: Field observed volume is plotted against stand volume estimated from LiDAR derived data: (A) representing Model (1) without additional independent variable of tree age; (B) representing Model (2) with additional independent variable of tree age