

RAPPORT

8b • 2005

Proceedings of
ForestSat 2005
in Borås May 31 - June 3



Håkan Olsson (ed.)

© National Board of Forestry May 2005

Edited by

Håkan Olsson, SLU Umeå

Layout

Bigitta Nyberg

Papper

brilliant copy

Printed

JV, Jönköping

Edition

350 ex

ISSN 1100-0295

ORDER NR 1741

Skogsstyrelsens förlag
551 83 Jönköping

This report is also available in digital form on-line at www.svo.se/bookshop
for illustrations in colour, please see the digital version.

Some minor typographical errors have been corrected in the digital version of the report.

Session 3b

SANKT-PETERSBURG GREEN BELT LAND COVER MAP DEVELOPMENT USING LANDSAT 7 ETM+ IMAGES

A.S. Alekseev^a, B.V. Shilin^b, A. A. Tronin^b and R. F. Treifeld^c

^aDepartment of Forest Inventory, Management and GIS, Sankt-Petersburg State Forest Technical Academy, Sankt-Petersburg, Institutsky per.,5, Russia, email: Alekseev@AA2996.spb.edu

^bResearch Center for Ecological Security of Russian Academy of Sciences, Sankt-Petersburg, Russia

^cNorth-West State Forest Inventory and Management Enterprise, Sankt-Petersburg, Russia.

ABSTRACT

Land cover map development was done on the base of the Landsat 7 ETM+ images. As a background topographical map of the scale 1:100 000 was used. Satellite image interpretation was done automatically using coefficients of spectral brightness for image and etalon (standard) types of landscape surfaces and the rule of minimum distance. Total number of 14 types of lands for Sankt-Petersburg greenbelt area was identified: broadleaf forest, mix forest, Scots pine dominated forest, Norway spruce dominate forest, clear cuttings, bushes and low dense forest, meadows and hay fields, agricultural lands, open areas (badlands), dacha settlements and villages, cities areas (including roads), areas densely covered by houses, bogs, waters. The results of satellite image interpretation were checked by comparison with ground truth data on 37 sample plots, statistical reliability of classification appears to be equal 85%. As a result of satellite image interpretation the digital land cover map for the Sankt-Petersburg green belt area of the scale 1:100 000 was developed. Each of the 14 types of the land was represented as a separate layer in ArcView3.2 GIS in the format SHP (ArcView3.2 Shape file). It appears that all kind of the forests cover 56.0% of the total green belt area with serious prevailing of Scots pine dominated forests – 31.1%, agricultural lands – 22.0%, all kind of populated areas – 18.6%, bogs and waters – 2.9%. This data and spatial patterns of each type of land distribution were compared with the data and maps obtained from other sources by means of GIS technique.

Keywords: Satellite image, greenbelt area, land classification, land cover map, GIS analysis

1 INTRODUCTION

Sankt-Petersburg green belt area account in total as much as 693 394 hectares, destine first of all for social and recreational use and last years is affected by a number of different kinds of influences such as house and industrial construction, pipelines construction, mining operations for peat, granite stove, gravel and sand, forestry operations such as thinning, felling, clear cuttings (sometimes illegal) etc. That is why was made a decision to create for this area special multi level, multi layer and multi functional GIS to manage and control the possible developments and changes in green belt land cover. Main layers which were created represent information on economics and management, ecological, social, cultural and historical aspects of greenbelt area.

We consider the GIS as a universal mean for combining, integrating, visualization and analysis of the spatial data relevant for green areas inventory and management from different sources and fields of interest such as: geology and geomorphology; climate, soil and water sciences; forestry, botany, dendrology and zoology; landscape ecology and nature protection; demography, economy, sociology, logistics, transport etc.

In such a way GIS implementation provide a new view on many problems and also help in decision such actual tasks as: socio-economic estimation of the lands and development of cadastre; organization and providing advanced systems for continues green areas inventory, land use management and planning; development of the system of protected forest territories for nature conservation; integration of the forest lands use into general regional land use system

Land cover map developed on the base of satellite images was considered as an essential component of GIS because first, such a map originated from an independent information source different from official and traditional data and may be used for verification of last, second, land cover map may be used for detection of the future changes in land use within greenbelt area.

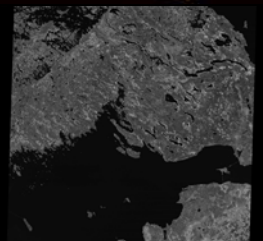
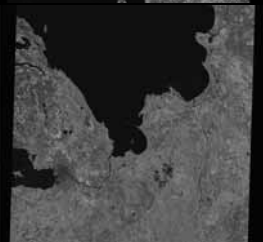
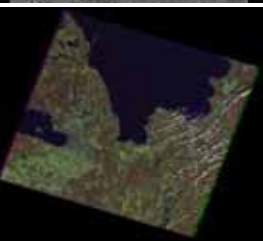
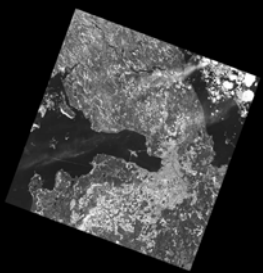
It was planned that land cover map should include such main categories of land as forests (with subdivision on coniferous, broadleaves, mixed and of low dense), clear cuttings, bushes, meadows, agricultural lands, open soils, bogs, waters, populated areas (with subdivision on villages and dacha settlements, cities and areas look like cities). In order to be joining with the whole GIS for greenbelt area and be compatible with used ArcView3.2 software land cover map should be presented as a number of separate layers, corresponding to each theme in SHP format.

Next demand on land cover map consist in it registration in WGS-84 coordinate system for further geo referenced calculations and different kinds of spatial analysis.

2 SATELLITE AND BACKGROUND DATA FOR MAP DEVELOPMENT

For the land cover map development was analyzed 12 of multi spectral digital Landsat scenes with spatial resolution of 30 meters 5 of which was chosen for following treatment. Chosen scenes are presented in the table 1.

Table 1. Satellite data used for land cover map development.

Date	Satellite	Scanner	Image
07.05.1993	Landsat 5	TM	
26.08.1996	Landsat 5	TM	
31.08.1997	Landsat 5	TM	
<u>25.04.2000</u>	<u>Landsat 7</u>	<u>ETM+</u>	
<u>22.04.2002</u>	<u>Landsat 7</u>	<u>ETM+</u>	

As a background topographical map of the scale 1:100 000 was used.

3 SATELLITE IMAGES TREATMENT AND MAP DEVELOPMENT

Two types of images corrections were done: geometric and radiometric. Geometric correction of the image was done, first, to compensate Earth rotation and curvature and second, to combine image and background map. Radiometric correction removes the noise and defective scanning lines. Also as a result of pretreatment were calculated coefficients of spectral brightness (figure 1) for image as a main source of physical information on scanned natural surface.

After satellite images geometric and radiometric corrections composite image (figure 2) on the area in interest was created and transformed into WGS-84 coordinate system. This synthesized image contain information needed for further thematic land cover classification (Jackson and Reginato, 1977; McGinnis and Tarpley, 1985; Chapursky, 1986; Spectral library for ASTER).

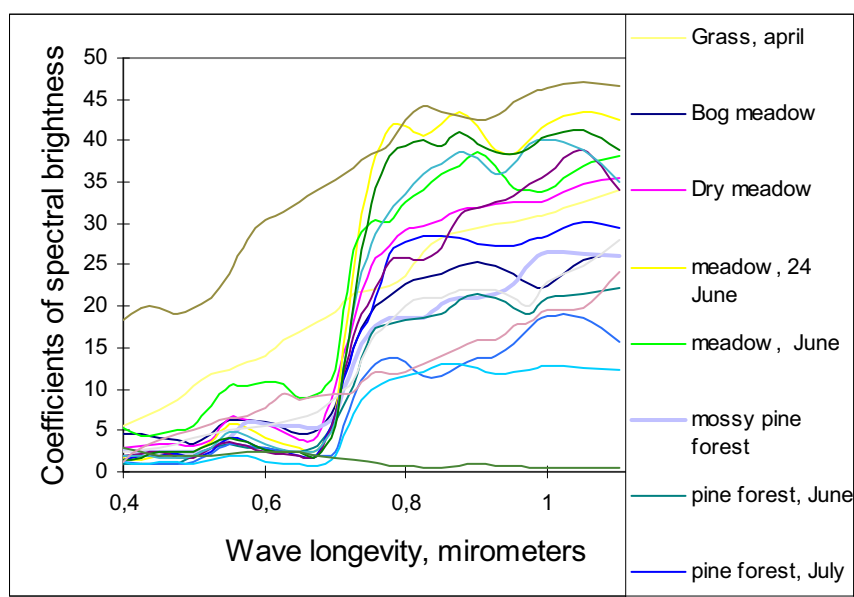


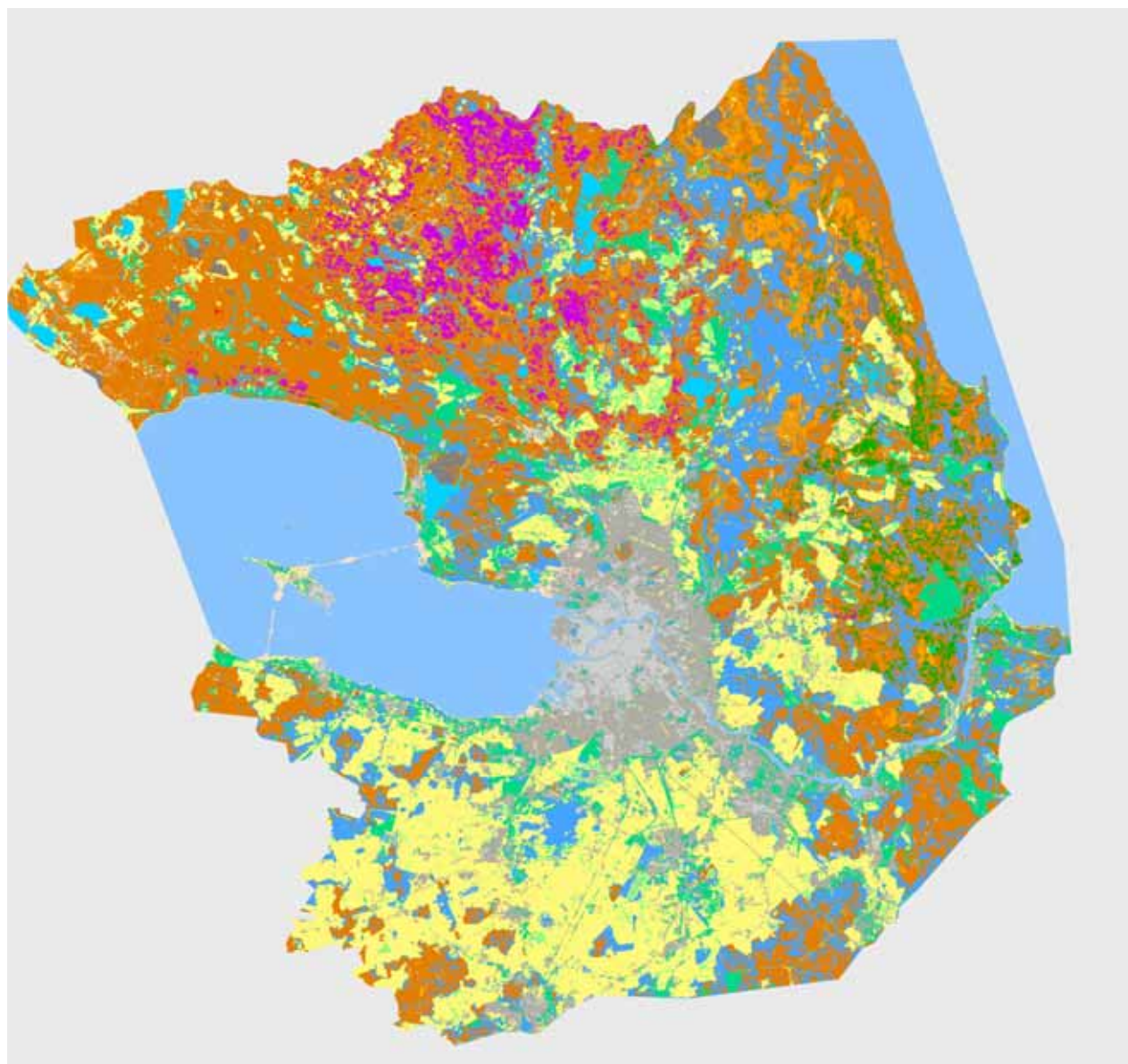
Figure 1. Coefficients of spectral brightness for main types of vegetation of Sankt-Petersburg greenbelt area.



Figure 2. Synthesized image prepared for thematic interpretation.

On figure 2. also delineated the borders of Sankt-Petersburg city and its greenbelt area.

Final land cover map for Sankt-Petersburg greenbelt area was obtained after last image automatic classification procedure using etalons (Curran and Hay, 1986; Homer *et.al.*, 1997; Cohen *et.al.*, 2001). Total number of 14 land categories was detected. The results of satellite image interpretation were checked by comparison with ground truth data on 37 sample plots, statistical reliability of classification appears to be equal 85%.



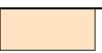
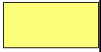
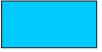
Legend					
	Coniferous, mainly Norway spruce		Bushes and low dense forests		Open soils
	Coniferous, mainly Scots pine		Bogs		City like areas
	Cuttings in coniferous forests		Meadows		Cities
	Mixed forests		Agricultural lands		Waters
	Broadleaves forests		Villages and settlements		Roads

Figure 3. Results of image interpretation and land classification using etalons.

Each of the 14 types of the land was represented as a separate layer in ArcView3.2 GIS in the format SHP (ArcView3.2 Shape file) and became an essential part of total GIS for the whole greenbelt. For each land category its area was calculated using GIS options. The results are presented in the table 2.

Table 2. Main types of lands of Sankt-Petersburg greenbelt area

Type of land	Area, hectares	Area, %
Broadleaves forests	110 659	16.0
Mixed forests	11 162	1.6
Coniferous, mainly Scots pine forests	215 729	31.1
Coniferous, mainly Norway spruce forests	20 150	2.9
Clear cuttings in coniferous forests	556	0.1
Bushes and low dense forests	30 721	4.4
Meadows	9 982	1.4
Bogs	8 104	1.2
Agricultural lands	142 866	20.6
Villages and dacha settlements	51 488	7.4
Cities	16 543	2.4
City like areas	60 770	8.8
Waters	11 929	1.7
Open soils	2 735	0.4
Total	693 394	100

From table 2 follows that all kind of the forests (including bushes and low dense forest) cover 56.0% of the total green belt area with serious prevailing of Scots pine dominated forests – 31.1%, agricultural lands (together with meadows usually used for hay collection) – 22.0%, all kind of populated (covered by houses or industrial objects) areas – 18.6%, bogs and waters – 2.9%.

ACKNOWLEDGMENTS

This report was prepared using support of Danish Center for Forest, Landscape and Planning of Copenhagen University within joint project “Planning and Management System for the Sankt-Petersburg Forest Greenbelt” granted by Danish Ministry of Environment and Energy. Special thanks to our colleagues Jasper Schipperijn and Ole Kaspersen.

REFERENCES

- Chapursky, L. 1986. Reflective features of natural objects in the diapason of 400-2500 nanometers. Published by Ministry of Defence. Moscow. 160 p.p.
- Cohen, W. Maersperger, T. Spits, T. and Oetter, D. 2001. Modeling forest cover attributes as continuous variables at a regional context with Thematic Mapper data. *International Journal of Remote Sensing*. 6: 124-135.
- Curran, P. and Hay, A. 1986. The importance of measurement error for certain procedures in remote sensing at optical wavelengths. *Photogrammetric Engineering and Remote Sensing*. 52(2): 229-241.
- Homer, C. Ramsey, R. Edwards, J. and Falkoner, A. 1997. Landscape cover-type modeling using a multi-scene Thematic Mapper mosaic. *Photogrammetric Engineering and Remote Sensing*. 63(1): 59-67.
- Idso, S. Jackson, R. and Reginato, R. 1977. Remote sensing of crop yields. *Science*. 196: 19-25.
- McGinnis, D. and Tarpley, J. 1985. Vegetation cover mapping from NOAA/AVHRR. *Advanced space research*. 5 (6): 359-369.
- Spectral library for ASTER. <http://speclib.jpl.nasa.gov/>

A METHOD FOR CALIBRATED MAXIMUM LIKELIHOOD CLASSIFICATION OF FOREST TYPES

Olle Hagner and Heather Reese

Remote Sensing Laboratory, Department of Forest Resource Management and Geomatics, Swedish University of Agricultural Sciences (SLU); Olle.Hagner@resgeom.slu.se

ABSTRACT

A new method for calibrated maximum likelihood (ML) classification of forest types has been developed for the Swedish CORINE land cover mapping project. The method corrects for the tendency of the ML-algorithm to over-represent dominant classes and to under-represent less frequent ones. The correction procedure iteratively adjusts the *prior* weights until class frequency in the output corresponds to objective (field-inventoried) estimates of the frequency for each class. National forest inventory data measured from a five-year period is used both for the estimation of relative frequency and to derive spectral signatures for each class. The method was implemented operationally in an automated production line which enabled rapid production of a country-wide forest type map from Landsat TM satellite data. This paper describes the details of the method and demonstrates results from operational use.

Keywords: *a priori* probabilities, Maximum likelihood, iterative, Landsat, proportions, forest inventory.

1 INTRODUCTION

The EU-CORINE Land Cover project has as an objective to map land cover from satellite imagery with a common classification scheme for all participating EU countries. The project started with the target year of 1990 and is repeated on a 10 year cycle; Sweden joined the CORINE mapping effort for the target year 2000. Simultaneous to producing CORINE, a spatially and thematically more detailed “Swedish Land Cover” map was made from which CORINE was generalized. To produce a cloud-free coverage, 50 Landsat ETM+ images were classified within two years. To reach this goal, an automated system was developed to perform the classification with minimal manual intervention. This paper describes the classification method used for the forest classes, referred to as a “calibrated Maximum Likelihood classification.” The procedure iteratively adjusts the *prior* weights until expected proportions of forest classes are achieved. Brief descriptions of the steps leading up to the classification, such as image pre-processing and updating of forest inventory data, are also given.

The classification method chosen for CORINE needed to be repeatable and robust since mapping should be repeated every 10 years. For this purpose, it is desirable when classifying large areas with multiple satellite images and a large number of classes to have access to a current and large field data set. Sweden has access to such field data for forestland through the annual National Forest Inventory (NFI), which is carried out by the Swedish University of Agricultural Sciences (SLU). The NFI data, through their systematic random sample over the country, provide statistically sound information about the presence and proportion of different forest types. The forest classification, also carried out by SLU, used NFI plot data to a large extent, for training data. In this case, the roll of the *prior weights* as used in the calibrated ML classification is to obtain an accurate spatial mapping of each class based on the satellite image and area estimates derived from the forest inventory.

The Maximum Likelihood (ML) algorithm is a classifier often used in land cover mapping. A documented effect when assuming actual prior probability for all classes is that dominant classes tend to be over-classified and more rare classes are under-represented in order to achieve the goal of minimising the number of misclassified pixels (Lillesand and Kiefer 2000). The problem can be corrected by carefully modifying the *prior* weights for each class until it gets the proper representation in the output. Several early studies such as Strahler (1980), and Skidmore and Turner (1988) pointed out the utility of including *a priori* probabilities within the ML classifier. Other studies have applied these principals to land cover mapping (Maselli et al., 1992; Gorte and Stein, 1998; Pedroni, 2003; Schuck et al, 2003), however it has been done on smaller scales and often without the use of extensive field data. The method described here was first introduced in 2001 (Hagner, 2001), within the scope of the CORINE project.

2 MATERIALS AND STUDY AREA

The materials used for the project were Landsat satellite data, NFI plots, and map data. In all, 50 Landsat scenes were processed to obtain a cloud-free classification. Primarily Landsat 7 ETM+ data acquired from 1999 to 2002 were used, with a geometric accuracy of $RMSE \pm 1$ pixel, and local errors of three pixels allowed (Bossard et al., 2000). For the classification, TM bands 3, 4, 5 and 7 were used. The forest inventory data were taken from the NFI which annually collects forest information on 13,000 plots over the entire country. The plot inventory is taken on over 200 variables on a 7, 10 or 20 m radius plots (Anon 2004). For the classification, all NFI plots from a minimum of 5 years to a maximum 10-year time period (if needed for a sufficient number of plots) were extracted for a single satellite scene. This gave approximately 2,000 plots per scene on average, with an absolute minimum of 600 plots and a maximum of 7,000. The digital map data used were land cover categories from 1:50 000 scale (available only for parts of Sweden) and 1:100 000 maps, and a 50m resolution DEM.

The project area covered the country of Sweden, which is approximately 44 million ha in size (National Atlas of Sweden, 1990). Forest cover types comprise 65% of the land area, with a mix of deciduous and coniferous species which varies widely from north to south within the country. Coniferous species of Norway spruce and Scots pine dominate in the northern two-thirds of the country with birch species also present. In the southern third of Sweden, more deciduous species occur (e.g., Oak, Beech and Aspen), along with Norway spruce, Scots pine and birch.

3 METHODS

The objective of the forest classification work presented here was to derive an accurate mapping of nine forest classes from satellite data using a robust and repeatable method to be automated into a production line. The main forest classes were clear-felled areas; regenerating areas; young coniferous forests; young deciduous forests; deciduous forests; mixed forests; coniferous forests of less than 15 m height; coniferous forests taller than 15 m; and, lichen dominated coniferous forests. For the classification, spectrally homogenous sub-classes to the nine target classes were defined, resulting in approximately 30 sub-classes (Table 1). In order to derive the final classes in the Swedish Land Cover and CORINE, GIS operations with map data (to add wetland and bedrock information) were performed. The forest class definitions were determined by previous pilot studies (Swedish Space Corporation, 1999) and later easily aggregated into the four CORINE classes of young forest, deciduous forest, mixed forest, and coniferous forest.

National Forest Inventory plots from a five-year time span covering a satellite scene's extent were assembled in a database. Variables included site index, percent forest cover within plot, tree species percentage, height, age, mean diameter, wood volume per tree species, field layer vegetation, soil moisture, age of clear-felling, among several others. Some variables necessary for the classification, such as crown closure and lichen percentage, were not collected by the NFI but could be calculated based on other variables. Since the plot information was collected at dates other than the scene acquisition date, variables for volume, age and height required updating which was done using growth functions (Söderberg 1986). Pre-processing steps for the satellite data included reduction of within-scene haze differences and illumination correction. (Hagner and Olsson, 2004). The processing involved detailed radiometric modeling with the spectral signature associated with each plot as the response variable and the following explanatory variables: land use category, wood volume, tree species proportions, age, soil moisture, site index and vegetation type, local atmospheric optical thickness and illumination effects due to sloping terrain.

Table 1. The main classes and sub-classes for the Swedish Land Cover classification.

Main class	Height (m)	Canopy closure	Definition Species Proportions	Sub classes used In the classification
Clear-felled	0-.5m	< 20 %	None	Bedrock; New clear-cut
Regenerating	.5-2m	< 20%	None	Open wetland; Fertile land
Young Forest	2-5m	> 20%	None	< 70% of each species
Coniferous 5-15m	5-15m	> 20%	>70% Coniferous	Coniferous on wetland; Coniferous on bedrock; >70% Pine; >70% Spruce; >70% Lodgepole
Deciduous	> 5m	> 20%	>70% Deciduous	Deciduous on wetland; Deciduous on bedrock; >70% Birch; >70% Oak; >70% Beech
Mixed Forest	> 5m	> 20%	30-70% Deciduous	Mixed on wetland; Mixed on bedrock; > 50% Pine; > 50% Spruce; > 50% Deciduous
Coniferous > 15m	> 15m	>20%	>70% Coniferous	Coniferous on wetland; Coniferous on bedrock; >70% Pine; >70% Spruce; >70% Lodgepole
Coniferous, lichen dominated	> 5m	> 20%	Lichen>12% Pine>50% Decid<30%	Volume > median for lichen dominated forest; Pine forest on bedrock

Each NFI plot had geographic coordinate information, primarily with differentially corrected GPS coordinates (used since 1996). However, due to local geometric errors in the image data and the lower accuracy of the pre-GPS NFI plot coordinates, a routine for matching plot and image data was implemented (Hagner et al., 2004).

The NFI plots to be used for classification were then assigned one of the nine primary forest classes and a sub-class. The class assignment was based on tree height, tree species, percent crown closure, and the land cover type according to the map data. Because forest also occurs on wetland and bedrock, and since their signatures can be confused with others, a stratification of the satellite data and NFI plots into forest, wetland (both forested and non-forested), and bedrock was made based on map data. Spectral signatures with highly atypical values for their assigned forest type were identified and replaced with a modeled signature appropriate for the forest conditions (Hagner and Tingelöf, 2002). However, to preserve the proper proportions of classes, only plots covered by dense clouds and shadows were removed in this step. At this point, the proportions of each sub-class were calculated from the NFI data.

The resulting database of NFI plots, a priori probabilities, sub-classes, and spectral signatures was assessed with quadratic discriminant analysis. From the result, the proportion of pixels per class was compared to class proportions from the NFI. If a class was under-represented, the *prior* weight was increased slightly and vice versa. The procedure was re-iterated until output class proportions converged to those estimated by the NFI.

In a second step, weights for the map masks of forest, wetland, and bedrock, were adjusted based on the relative occurrence of a class within each map mask, based on Equation1:

$$v_{k,m} = v_k * (p_{k,m} + e) / (p_k + e) \quad (1)$$

- $v_{k,m}$ = weight for class k under map mask m
- v_k = weight for class k in the whole image (disregarding map mask) calculated in step 1
- $p_{k,m}$ = proportion for class k under map mask m
- p_k = proportion for class k for all map masks together
- e = an error term for avoiding over-adjustment when the statistical basis for a class is low

When classifying wetland areas according to the map mask, bedrock areas had zero weight and when classifying non-wetland areas, wetland classes had zero weight.

4 RESULTS AND DISCUSSION

The output from the production line was a pixel-wise classification of the nine forest classes over the entire satellite scene. In the initial stages of the project (i.e., but not throughout the project), the result from the ML classification with equal prior probabilities was compared to the method using the calibrated-ML classification. To do this, a per-pixel cross validation was performed with a total of 2443 NFI plots used for a Landsat scene in southern Sweden (Path 193, Row 021). The use of a calibrated ML classification resulted in an overall increase of correctly assigned classes by nearly 9% (from 58.6% to 67.2%), with improvement in every class, except Mixed Forest (Table 2). As expected the relative frequency of each class was much more realistic with the calibrated ML. Classes which had a low representation in the training data, such as coniferous forest on lichen and clear-felled areas, were classified more accurately using the calibrated ML.

Table 2. Result from per-pixel cross-validation of correctly placed plots using ML versus calibrated-ML classification.

Class	ML-classification	Calibrated-ML
Clear-felled	47/125 = 37.5 %	95/125 = 76.0 %
Regenerating	140/240 = 58.3 %	150/240 = 62.5 %
Young forest	61/156 = 38.8 %	68/156 = 43.6 %
Conif. 5-15m	101/309 = 32.7 %	135/309 = 43.7 %
Mixed forest	83/228 = 36.3 %	77/228 = 33.8 %
Decid. Forest	269/469 = 57.4 %	362/469 = 77.2 %
Conif. > 15m	731/916 = 79.8 %	749/916 = 81.8 %
Conif. on lichen	0/7 = 0.0 %	5/7 = 71.4 %
Overall	1432/2443 = 58.6%	1641/2443 = 67.2%

This method has an advantage over other studies which initially do not have accurate data on class frequencies thus needing to estimate it from the available data. Also due to the detailed description of each field plot it is possible to define spectrally *and* thematically homogenous subclasses which in turn enable robust outlier detection. This makes the calibrated ML classification presented here a robust starting point for repeatable and accurate forest/land cover classifications.

For the Swedish Land Cover classification, 50 Landsat scenes were processed at a rate of approximately one scene per week. Without the automated production line, such a production rate would not have been possible. For the forest classification, accuracy assessment was carried out in two ways: photo-interpretation and cross-validation. An accuracy assessment using photo-interpretation was limited to three areas in the country due to expense. A compilation of the accuracy over the three areas for the forest classes alone gave approximately 75% accuracy for forest types (Hagner and Reese, 2004).

5 CONCLUSION

A robust and repeatable classification method for forest classes was needed as part of the Swedish CORINE project. A calibrated ML classifier was developed using *prior weights* derived from forest inventory data which were then modified and run iteratively for three different map strata. The use of a calibrated ML classification resulted in an overall increase of correctly assigned classes by nearly 9% (from 58.6% to 67.2%), with improvement in every class, except Mixed Forest and class frequencies closely corresponding to those predicted from the NFI. The algorithm was used operationally for over 50 Landsat images.

REFERENCES

- Anonymous, 2004. *Instruktion för Fältarbetet vid Risksskogstaxeringen*. Swedish University of Agricultural Sciences, Department of forest resource management and geomatics, Umeå, Sweden (In English: Instructions for Fieldwork of the National Forest Inventory).
- Bossard, M., Feranec, J., and Otahel, J., 2000. CORINE land cover technical guide – Addendum 2000. Technical report No. 40. European Environment Agency. 105 p.
- Gorte, B. and Stein, A., 1998. Bayesian Classification and Class Area Estimation of Satellite Images Using Stratification. *IEEE Transactions on Geoscience and Remote Sensing*, 36(3), 803-812.
- Hagner, O., 2001. Frequency calibrated Maximum Likelihood classification. Abstract presented at the 2001 RESE Annual meeting, 14-16 November, 2001, Holmsund, Sweden.
- Hagner O., and Olsson, H., 2004. Normalisation of within-scene optical depth levels in multispectral satellite imagery using national forest inventory plot data. In *Proceedings from the 24th EARSeL Symposium, Workshop on "Remote sensing of land use and land cover"*, Dubrovnik, Croatia, May 28-29, 2004.
- Hagner, O. Nilsson, M., Reese, H., Egberth, M., Olsson, H., 2004. Procedure for Classification of Forest for CORINE Land Cover in Sweden. In *Proceedings from the 24th EARSeL Symposium, Workshop on "Remote sensing of land use and land cover"*, Dubrovnik, Croatia, May 28-29, 2004.
- Hagner, O., and Tingelöf, U., 2002. Metod för pixelvis klassning av skog i Svenska Marktäckedata (SMD) (In English: The method for pixel-wise classification of forest for the Swedish Land Cover Data), manuscript.
- Lillesand, T. and Kiefer, R., 2000. *Remote Sensing and Image Interpretation*. Fourth edition, John Wiley and Sons, 724 p.
- Maselli, F., Conese, C., Petkov, L., and Resti, R., 1992. Inclusion of Prior Probabilities Derived from a Nonparametric Process into the Maximum-Likelihood Classifier. *Photogrammetric Engineering and Remote Sensing*, 58(2): 201-207.
- National Atlas of Sweden, 1990. *The Forests*. Editor N.E. Nilsson. SNA Publishing, Italy. 144 p.
- Pedroni, L. 2003. Improved classification of Landsat Thematic Mapper data using modified prior probabilities in large and complex landscapes. *International Journal of Remote Sensing*, 24(1):91-113.
- Schuck, A., Päivinen, R., Häme, T., Van Brusselen, J., Kennedy, P., and Folving, S., 2003. Compilation of a European forest map from Portugal to the Ural mountains based on earth observation data and forest statistics. *Forest Policy and Economics* 5(2003):187-202.
- Skidmore, A.K., and Turner, B.J., 1988. Forest mapping accuracies are improved using supervised nonparametric classifier with SPOT data. *Photogrammetric Engineering and Remote Sensing*, 54, 1415-1421.
- Söderberg, U. 1986. Functions for forecasting of timber yields – Increment and form height for individual trees of native species in Sweden. Report No. 14. Swedish University of Agricultural Sciences. Department of Forest Mensuration and Management.
- Strahler, A.H. 1980. The Use of Prior Probabilities in Maximum Likelihood Classification of Remotely Sensed Data. *Remote Sensing of Environment*, 10:135-163.
- Swedish Space Corporation, 1999. Pilotproduktion av CORINE Landtäckedata – beskrivning av produktionsprocess. Rapport för Lantmäteriverket, Miljödatacentrum, Naturvårdsverket, and Försvarsmakten. Rymdbolaget.
- Teillet, P.M., Guindon B., and Goodenough D.G., 1982. On the slope-aspect correction of multispectral scanner data, *Canadian Journal of Remote Sensing* 8:84-106.

GA-DRIVEN FEATURE SELECTION IN OBJECT-BASED CLASSIFICATION FOR FOREST MAPPING WITH IKONOS IMAGERY IN FLANDERS, BELGIUM

Frieke M.B. Van Coillie, Lieven P.C. Verbeke and Robert R. De Wulf

Ghent University, Laboratory of Forest Management and Spatial Information Techniques,
Coupure Links 653, 9000 Gent, Belgium, email: frieke.vancoillie@ugent.be

ABSTRACT

Obtaining detailed information about the amount of forest cover is an important issue for governmental policy and forest management. This paper presents a new approach to update the Flemish Forest Map (FFM) using IKONOS imagery. The proposed method is a three-step object-oriented classification routine that involves the integration of 1) image segmentation, 2) feature selection by Genetic Algorithms (GAs) and 3) joint Neural Network (NN) based object-classification. Results showed that, with GA-feature selection, the mean classification accuracy (in terms of Kappa Index of Agreement (KIA)) was significantly higher ($p < 0.01$) than without feature selection. On average, the summed output of 50 networks provided a significantly higher ($p < 0.01$) classification accuracy than the mean output of 50 individual networks. Finally, the proposed classification routine led to a significantly higher ($p < 0.01$) classification accuracy as compared with a traditional working strategy, i.e. without feature selection and joint network output. In addition, the proposed method showed its potential when few training data were available.

Keywords: Segmentation, Genetic Algorithms, Feature Selection, Neural Networks, Classification, Very High Resolution Imagery, Forest Mapping.

1 INTRODUCTION

Because of a typically long planning horizon, forest management requires detailed information on the amount of forest cover, not only to support current operations, but also to provide a record of past activities and to predict the possible outcomes of management decisions (Weir 2002). In Flanders, the current FFM (OC GIS 2001) was produced through visual aerial photo interpretation and manual stand delineation: an expensive, labour-intensive and sometimes subjective assignment. Semi-automated and computer assisted interpretation of digital imagery is assumed to offer an alternative to acquire forest information, reducing time and costs, and increasing consistency.

This paper presents a new approach to update the FFM according to the forest definitions specified by the Flemish Forest Decree. The use of IKONOS VHR imagery sets new demands for image analysis methods that usually operate on a pixel-by-pixel basis and do not utilize the spatial information present in the image. One way to exploit this information is by using image analysis units larger than a single pixel created through segmentation. The proposed method is a three-step object-oriented classification routine that involves the integration of 1) image segmentation, 2) feature selection by GAs and 3) joint NN based object-classification.

2 DATA

The current FFM is a digital vector data layer divided into 280 map sheets of 8 by 10 km. The proposed method should be applicable on all possible forest landscapes, from low to high forest cover. Therefore representative study areas were selected based on a *stratified random sampling/clustering* strategy. We concentrated on one study region: Beringen, map sheet 25-3, sampled from cluster 3 and situated in the province of Limburg (Figure 1).

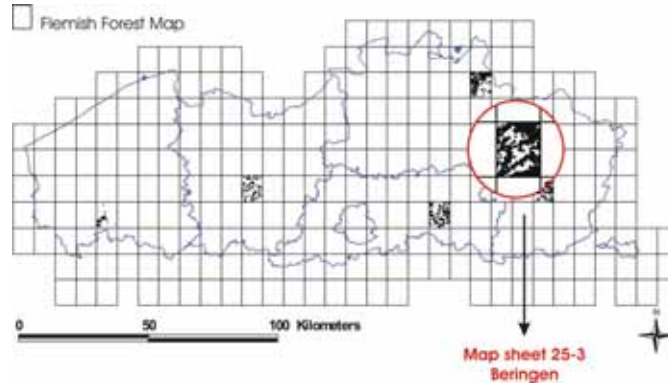


Figure 1: Study area of Beringen situated in Flanders

For Beringen, IKONOS imagery was available from 3 different scenes. Prior to analysis, the data were geo-radiometrically corrected and mosaiced (including histogram matching) by the image supplier OC GIS Vlaanderen. Next to the 1m panchromatic and the four 4m multispectral bands, a Normalized Difference Vegetation Index (NDVI) served as input data.

3 METHODOLOGY

3.1 IMAGE SEGMENTATION

Image segmentation is a commonly applied technique in the fields of machine vision and pattern recognition. Segmentation methods can be divided into point-based (e.g. grey-level thresholding), edge-based (e.g. edge detection techniques) and region-based (e.g. split and merge) methods. For reviews of different segmentation techniques, see for example Haralick and Shapiro (1983) and Pal and Pal (1993).

In the first step of the methodology, image objects are created by means of the segmentation algorithm introduced by Baatz and Schäpe (2000), which is implemented in eCognition (eCognition 2000). This is a bottom-up region merging technique and is therefore regarded as a region-based algorithm. It starts by considering each pixel as a separate object. Subsequently, pairs of objects are merged to form larger segments. The merging decision is based on a local homogeneity criterion, describing the similarity between adjacent image objects. The pair of objects with the smallest increase in the defined criterion is merged. The process terminates when the smallest increase of homogeneity exceeds a user-defined threshold (the so-called scale parameter, here set to 25, based on trial and error). For this application, the segmentation algorithm utilizes spectral information to extract spatially continuous, independent and homogenous regions or image objects (264981 in total) (Figure 2).

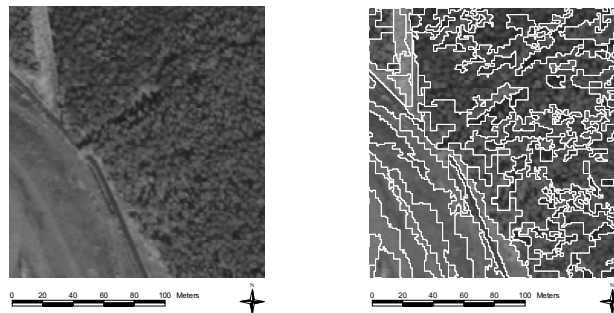


Figure 2: Detail of panchromatic scene before (left) and after (right) segmentation

In a second step the image objects are characterized by calculating spectral, shape and textural features. The *spectral* features are channel means, standard deviations, brightness and maximum difference. The object *shape* is represented by calculating object Length/Width proportion and a Shape Index. Concerning *texture*, first and second order texture measures are derived from either the Grey-Level Co-occurrence Matrix (GLCM) or the Grey-Level Divergence Vector (GLDV), such as Homogeneity, Contrast, Dissimilarity, Entropy, Angular Second Moment, Mean, Standard Deviation and Correlation. A total of 89 different object features are extracted. For classifying this high dimensional dataset, a large NN (with many inputs and a large number of hidden neurons) can be used. However, if NNs are large, many parameters

(weights and biases) need to be estimated using a finite number of training samples. In that case, overfitting might occur: the networks may not generalize well for unseen data although high classification accuracy can be achieved for the training data (Benediktsson and Sveinsson 1997). Hence, it is necessary to reduce the input dimensionality in order to obtain smaller networks performing well both in terms of training and test classification accuracies. This leads to the importance of feature extraction for NNs. The second step of the proposed method uses GAs to select appropriate object features according to the underlying classification problem.

3.2 GA-DRIVEN FEATURE SELECTION

GAs are search and optimization algorithms based on natural selection and natural genetics (Goldberg 1989). The basic data structure is a fixed-length binary encoded string and is referred to as a chromosome. Just like its biological counterpart, a chromosome consists of variables or genes that can have different values or alleles. Each individual can be evaluated using an objective value, also called fitness value that can be calculated from the information contained in the chromosome. A set of individuals together form a population. Each member of the population receives a fitness value based on the coding of its genes. To continue the biological analogy, fitter individuals will be enriched in number towards future generations. This occurs in a process called *selection* where fitter individuals are preferred over others. If only this process of recombination was at work, diversity would not be guaranteed. That is why *mutation* is essential. Mutation causes *mistakes* in the reproductive cycle (Booker *et al.* 1997). The main loop of a GA is described as follows. Initially, a random parent population is established. The population size is set, determining the amount of individuals inside the population. From the initial population a selected group of fitter individuals is chosen as members for the mating pool. On these individuals a binary crossover operator and a unary mutation operator are applied with probabilities respectively called the crossover and mutation probability. The result of this recombination and mutation yields the next generation of individuals. This cycle is repeated until a stopping criterion is met (Goldberg 1989).

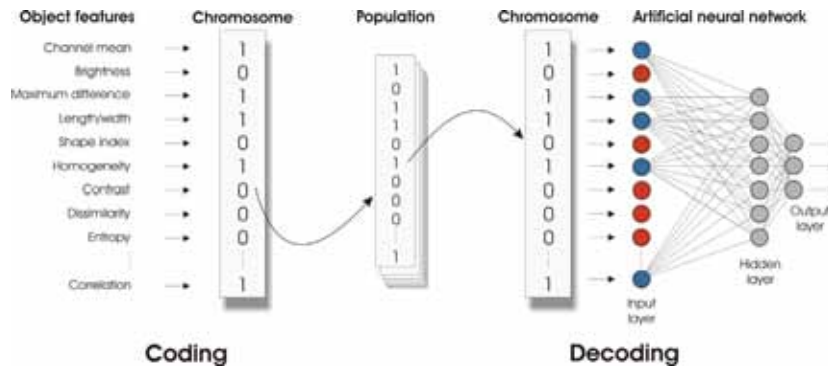


Figure 3: Coding and decoding in GA-driven feature selection

In this study GAs are used to address the problem of object feature selection. The problem is *coded* into a binary string: 89 bits long (one for each object feature), where the genes can have the value 0 or 1 (Figure 3). Starting with a mating pool of 100 individuals having on average 5 non-zero values, the fitter ones are selected based on their fitness value. An objective function is applied calculating fitness, i.e. classification accuracy in terms of KIA. The KIA index serves as indicator of the extent to which the percentage correct values of an error matrix are due to true versus chance agreement (Cohen 1960). Therefore, the individuals have to be *decoded* to serve as inputs for NN classification (Back-propagation training with learning rate=0.0001 and momentum=0.001). Figure 3 also shows the decoding process: all object features valued by 1 are represented by fully-connected input neurons (blue), all other neurons (valued by 0 and represented in red) are disconnected. The GA replacement strategy designed for this application uses tournament selection, elitism and uniform crossover. For more detailed information about these operators we refer to Goldberg (1989), Davis (1991) and Booker *et al.* (1997). The crossover probability is set to 1.0, while the mutation probability equals 0.7/number of features, here 89. Reproduction stops after 20 generations. All offspring individuals represent spectral-shape-texture combinations that result in high classification accuracy. From the offspring, the 50 best performing low dimensional NN architectures are selected for subsequent classification.

3.3 JOINT NN IMAGE CLASSIFICATION

Instead of interpreting all classification results individually, all predictions are combined by summing network outputs: the networks thus jointly provide a classification result. It has been reported that the combination of NNs can outperform the accuracy of the best single network (Bishop 1992). Giancinto and Roli (2001) stated that classifier combination is effective only when the individual classifiers are *accurate* and *diverse*, that is, if they display low error rates and commit different errors. Since the NNs here are the result of a GA selection procedure and thus represent the fittest and best performing individuals, their error rates are low. Moreover they exhibit different architectures and thus make different errors.

3.4 EVALUATION

The three-step classification routine is evaluated against a traditional NN classification strategy, i.e. without feature selection and joint network output. First, the effect of GA-driven feature selection is evaluated with a t-test. Secondly, the added value of joint network outputs is considered using a paired t-test on fully-connected networks (i.e. network with 89 fully-connected input neurons). Next, the accuracy levels of both classification methods are compared with a z-test. Finally, the performance of the proposed classification routine is evaluated in function of the training set size, considered important for real life situations.

4 RESULTS AND DISCUSSION

From visual interpretation it appears that the proposed classification routine outperforms the current FFM: all forest fragments are accurately delineated, ranging from individual trees, over tree rows to large forest patches. Mistakes which were inherent in the FFM (due to subjective interpretations) are eliminated (Figure 4).

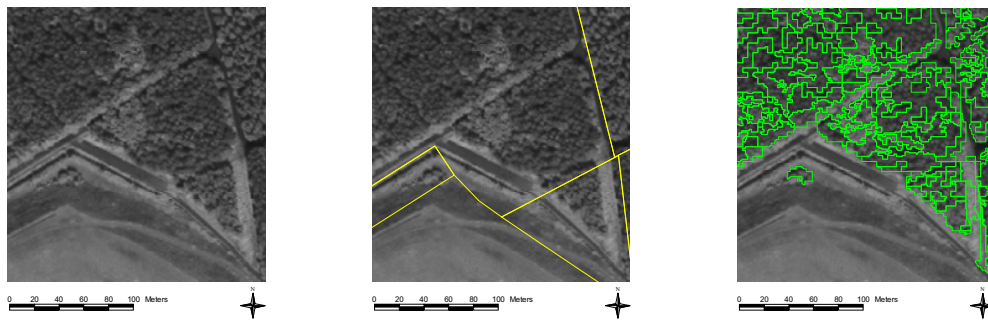


Figure 4: Detail of panchromatic scene (left), forest cover according to the current FFM (middle) and according to the proposed method (right)

In terms of classification accuracy, the proposed method outperforms a traditional strategy at three different levels. First, a t-test answers the question whether in a network population of 100 individuals, fully-connected networks reach the same mean KIA as partially-connected networks resulting from feature selection. Table 1 shows that, with GA-feature selection, the mean classification accuracy is significantly higher than without feature selection.

Secondly, in order to test whether the combination of NNs yields accuracy improvement, 1000 networks were trained. Since repeating the GA-feature selection procedure 1000 times is computationally too demanding and extremely time-consuming, this experiment is performed with fully-connected networks. Hereby we assume that, if network combination is successful with fully-connected networks, all of which are *accurate*, it will surely work with partially-connected networks, which are *accurate* and *diverse*. For each group of 50 networks, mean and joint accuracies were calculated and compared using a paired t-test. From Table 1, we conclude that the combination of NNs significantly outperforms the use of individual networks in classification accuracy justifying the implementation of joint NN outputs as third step in the classification routine.

Thirdly, the proposed classification routine is evaluated by comparing its accuracy level with the mean accuracy of 50 individually trained NNs. A significant KIA improvement of 1.9% is observed¹.

Table 1. Results from t- and z-tests

<i>Classification routine</i>	<i>Mean KIA</i>	<i>N</i>	<i>t/z-statistic</i>	<i>p</i>
Feature selection	0.9539	50	t=16.748	< 0.01
No feature selection	0.9465	50		
Joint output	0.9536	20	t=15.401	< 0.01
Individual output	0.9483	20		
¹ Feature selection + joint output	0.9644	1	z=5.838	< 0.01
No feature selection + individual output	0.9465	50		
² Feature selection + joint output	0.8449	1	z=16.789	< 0.01
No feature selection + individual output	0.7788	50		

Considered the fact that, with a finite number of training data, overtraining might occur with high dimensional NNs, we assume that the relevance of feature selection is inversely proportional to the number of training samples. In a fourth and last evaluation experiment, the number of training samples is reduced from ± 5000 objects-per-class¹ to $\pm 100^2$. With this small training set, 50 individually trained networks reach a mean accuracy of 0.7788 while a KIA of 0.8449 is obtained using the proposed method. This yields a significant accuracy improvement of 8.5%.

5 CONCLUSIONS AND FUTURE RESEARCH

The ultimate goal of this study is to provide the Flemish government with an efficacious method to update the FFM based on IKONOS imagery. Comparing the classification result with the current FFM, the proposed strategy presents a valuable alternative to visual photo interpretation. Feature selection and network combination assisted in building a high performance image classifier. Preliminary examination showed that the method's added value increases as training set size decreases.

Future research will concentrate on the effect of training set size on the classifier's performance. The proposed method will also be validated on its applicability for other study regions. The GA offspring of each study area will be compared in order to design a working strategy for all other map sheets within each cluster, resulting in an applicable method for the whole region of Flanders.

ACKNOWLEDGEMENTS

The authors are grateful to the OC GIS Vlaanderen for providing and correcting the IKONOS imagery covering the study site. Bos & Groen, Ministry of the Flemish Government is acknowledged for supplying detailed information on the current Flemish Forest Map and the Flemish Forest Decree.

REFERENCES

- Baatz, M. and Schäpe, A. 2000. Multiresolution segmentation - an optimization approach for high quality multi-scale image segmentation. In *Angewandte Geographische Informationsverarbeitung*. Wichmann-Verlag, Heidelberg, pp.12-23.
- Benediktsson, J. A. and Sveinsson, J. R. 1997. Feature extraction for neural network classifiers. In *Neuro-computation in Remote Sensing Data Analysis*. Springer-Verlag, Berlin, pp.97-104
- Bishop, C. M. 1997. *Neural Networks for Pattern Recognition*. Clarendon Press, Oxford.
- Booker, L. B., Fogel, D. B., Whitley, D. and Angeline, P. J. 1997. Recombination. In *Handbook of Evolutionary Computation*. IOP Publishing Ltd and Oxford University Press, Oxford, pp. C3.3:1– C3.3:27.
- Cohen, J. 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20:37–46.
- Davis, L. 1991. *Handbook of Genetic Algorithms*. Van Nostrand Reinhold, New York.
- eCognition 2000, Definiens Imaging GmbH, München, Germany.
- Giacinto, G. and Roli, F. 2001. An approach to the automated design of multiple classifier systems. *Pattern Recognition Letters*, 22:23-25.
- Goldberg, D. E. 1989. *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley, USA.
- Haralick, R. and Shapiro, L. 1985. Survey: image segmentation techniques. *Computer Vision, Graphics and Image Processing*, 29:100-132.
- OC GIS 2001. *Digital version of the Flemish Forest Map*. AMINAL, Department Bos & Groen, OC GIS-Vlaanderen.
- Pal, N. R. and Pal, S. K. 1997. A review on image segmentation techniques. *Pattern Recognition*, 26:1277-1294.
- Weir, M. 2002. Monte Carlo simulation of long-term spatial error propagation in forestry databases. In *Spatial Data Quality*. Taylor and Francis, London, pp. 294-303.

OBJECT-ORIENTED CLASSIFICATION OF REMOTE SENSING DATA FOR THE IDENTIFICATION OF TREE SPECIES COMPOSITION

Filip Hájek

Department of Forestry Management, Faculty of Forestry and Environment CUA Prague
Kamýcká 1176, 165 21 Praha 6

hajek@fle.czu.cz

ABSTRACT

This paper deals with the classification of tree species composition from Ikonos imagery (4m resolution) based on the object-oriented image analysis in eCognition software. The image was acquired over a man-planted forest area with proportion of various forest types (conifers, broadleaved, mixed) in the Krušné Hory Mts., Czech Republic. In order to enlarge class feature space, additional channels were produced by low-pass filtering, IHS transformation and influence of various Haralick texture measures on classification was also examined. The principal component calculated from original Ikonos bands was applied in the pre-processing phase. The segmentation and classification were conducted on three levels to be incorporated into the hierarchical image object network. The higher level separated image into smaller parts regarding the stand maturity and structure, the lower (detailed) level assigned individual tree clusters into classes for the main forest species. The third level was created to distinguish forest/non-forest boundaries. Classification accuracy was assessed by comparing the automated technique with the field inventory data using Kappa coefficient. Apart from transferable rule base creation, the study aimed to determine appropriate scale for species composition estimation using common image data. Therefore the methodology will be tested on colour aerial photos in the further research.

Keywords: Object-oriented image analysis, median filters, texture, tree species composition, forestry management

1 INTRODUCTION

Remote sensing and image interpretation have been utilized in forestry management for many years. These methods can be applied in various tasks ranging from forest thematic mapping to the detailed tree or stand characteristics survey. Besides the advancement in digital aerial methods, high-resolution satellite sensors (e.g. Ikonos, QuickBird) are now available for operational use. However, the automated classification of such data is still problematic due to greater spectral variation within one class (Halounová, 2003).

Previous studies on high-resolution data (Gougeon, 1995b) proved that traditional spectral-based methods result in rather poor or incorrect classification. Much information is contained in spatial relations of pixels and a few studies already showed the object oriented approach promising when classifying VHR data (Batz & Schäpe, 1999, Leckie et al., 2003). The contribution of textural and structural information was also examined (Haralick & Shapiro 1992, Brantberg, 1999) and various algorithms, such as co-occurrence matrix were applied to extract texture characteristics of trees (Zhang, 2003). Neural networks (Gopal & Woodcock, 1996) and fuzzy classification improved modeling of real-world dependencies (Benz et al., 2004). Furthermore, increased use of a priori knowledge and information extraction become important with the rapid development of GIS.

This paper explores and demonstrates capability of object oriented image analysis software eCognition (Definiens Imaging, Germany) for the tree species classification from Ikonos imagery. Combination of complex object description, hierarchical image object network and fuzzy system makes eCognition a challenge to knowledge-based image interpretation in a range of forestry management applications. Next project objective is to determine appropriate scale and accuracy of species composition estimation using common image data. The prospect of knowledge base creation for the high level automation in operational forestry is also discussed.

2 SITE AND FIELD DATA COLLECTION

The research was conducted in man-planted forests nearby the town Hrob (50°40'N, 13°43'E) in the Krušné Hory Mts., Czech Republic. This submontane area consists of patches of mature Spruce (*Picea Abies* L.) and Beech (*Fagus Silvatica* L.) forest, with the substantial proportion of Larch (*Larix deciduas* Mill.) and also young plantations of Beech, Birch (*Betula pendula* L.) and *Picea pungens* often mixed with *Larix* and *Betula Pubescens*. The planted mature stands are mostly of the same age, but very heterogeneous in species composition, stocking density and canopy structure. The natural regeneration in addition to the planted trees sometimes occurs. Silviculture practices range from clear cutting to seed felling with heavy thinning on some spots.

Based on the previous information from LHPO forest inventory, twenty 400m² plots covering areas with 100% species composition were located as a reference data. Sample plot selection put emphasis on size and class purity to provide representative basis for accuracy assessment. The boundaries of each plot were determined with differential GPS SX BlueTM and PDA with ESRI ArcPadTM mobile GIS.

3 IMAGERY PREPROCESSING

The Ikonos (Space Imaging, USA) image was acquired on 17th September 2003. Except for some hardwood species, most vegetation was still green and fully foliated. Data were delivered in a geo-registered UTM projection (zone N33) with 11-bit radiometric resolution. The image contained significant amount of clouds and atmospheric haze, so a 3x3 km subset of forested area with clear sky conditions was chosen for the analysis. There is also important amount of shadow fraction throughout the scene associated with solar and observation angles (Table 1)

Table 1. Ikonos viewing and illumination geometry

View Azimuth	View Elevation	Sun Angle Azimuth	Sun Angle Elevation
330.30 degrees	71.78 degrees	170.58 degrees	41.51 degrees

In the next step, class feature space was enlarged by the calculation of additional channels. Foremost the single principal component from original bands was derived and then Median filter (kernel sizes 3x3 and 5x5) was applied in order to suppress spatial frequency. Several Haralick (GLCM) texture measures were calculated and the contribution to class separability was tested. Measures Mean, Variance and Homogeneity with window sizes of 3x3 and 5x5 were chosen. Further, layers calculated by IHS transformation and edge detection (Sobel operator) were also applied in the classification.

4 OBJECT ORIENTED ANALYSIS

After the feature space enlargement, image segmentation was performed to further handle high spectral variation and overlapping values of classes. In this phase, image was split into smaller regions (object primitives) to simplify thematically complex data content. The classification was then performed using segments instead of single pixels.

4.1 MULTIREOLUTION SEGMENTATION

In the segmentation process, size and shape of desired objects is defined by the calculation of heterogeneity between adjacent pixels, where Scale is the main input parameter. Shape factor (colour/shape ratio) and spatial properties (smoothness/compactness ratio) are other variables to define homogeneity of object primitives.

Segmentation was conducted stepwise on several levels using different scales to construct the hierarchical image object network. The primary level was created using Scale parameter of 15. After preliminary classification, objects were merged by classification-based segmentation and the result (basic landuse classification) was re-imported into eCognition as a thematic layer. The sublevel was then segmented only within the area of interest (class Forest) using Scale parameter equal to 5. The finest objects with the scale value of 3 were calculated at the third level applying the same approach.

The Shape factor was set to higher value for the coarse segmentation and lower value at finer scale (higher influence of spatial properties). The two lower levels were processed using four Ikonos bands and median filtered channel of kernel size 3x3, the coarse landuse segments were made based only on thematic layer. Layer weights were set in relation to their standard deviations.

4.2 CLASS DEFINITION

The three level hierarchical image object network was used to delimit classes. Level 3 comprised basic “Landuse” types - Urban, Fields and Forest. This served to mask all non-forest areas. The lower level 2 “Forest” aimed to separate forest regions into Dense (young and mature stands with more less closed canopies), Sparse areas and Clearcuts. Sparse forests mostly consisted of low stocking mature beech trees with presence of visible ground. The detailed level 1 “Stand” was set to distinguish four main forest species in the area - Fagus, Picea, Larix and Betula. Further, structures of shadows and bare ground were classified on this level.

All classes of “Forest” level were also recognised at the lower “Stand” level for purpose of post classification improvement.

4.3 CLASSIFICATION

In order to create distinct and fully transferable rule base, fuzzy logic membership functions were used to define object features. Fuzzy description enables classes to be assigned according to membership degree rather than crisp threshold values. Following features were applied:

- Object features: mean layer values (blue, red, NIR, brightness, GLCM mean 3x3, IHS, Sobel NIR), ratio layer values (blue, red), area generic shape feature,
- Class-related features: relative border to neighbour objects, relative area of sub-objects, existence of sub-objects (super-objects)
- Customised features: NDVI, (red-green) vegetation index, IHS/ brightness index,

Besides MF classification, preliminary nearest neighbour classification was done on the lowest level and features suitable to separate tree species were evaluated using sample editor (histogram comparison). The masking technique of determining foremost the easy classes (ground, beech trees) and moving on to more difficult ones was often applied. Then the class boundaries were improved using class-related features and finally corrected by means of classification-based segmentation.

5 RESULTS

The classification accuracy was evaluated using field reference data. Sample areas were imported into project by means of TTA mask (Definiens Inc. 2003) and the corresponding classes were linked to form confusion matrix (Table 2). Several measurements such as Producer’s, User’s, Overall accuracy and Kappa index of agreement were derived for each class. Besides, classification reliability (Best classification result) and stability within fuzzy concept were assessed.

Table 2. Error matrix of classification accuracy assessment. The Overall Accuracy is 0.945 with the Kappa index of agreement equal to 0.914

User \ Referer. Class	Larix	Betula	Fagus	Picea	sparse	ground	shadows	fields	urban	Sum
Larix	211	104	2	5	0	0	0	0	0	322
Betula	25	322	5	0	35	0	0	0	0	387
Fagus	93	14	685	0	2	0	0	0	0	794
Picea	34	0	0	444	0	0	107	0	0	585
sparse	10	0	0	0	317	0	0	0	0	327
ground	0	0	0	2	17	207	0	0	0	226
shadows	0	0	0	37	0	0	216	0	0	253
fields	0	0	0	0	0	0	0	5849	0	5849
urban	0	0	0	0	0	0	0	0	1334	1334
unclassified	17	17	5	3	1	0	17	0	0	60
Sum	390	457	697	491	372	207	340	5849	1334	
Producer’s	1	0.541	0.705	0.983	0.904	0.852	0.635	1	1	
User’s	0.916	0.655	0.832	0.863	0.759	0.969	0.854	1	1	
KIA Per Class	1	0.526	0.693	0.981	0.898	0.847	0.626	1	1	

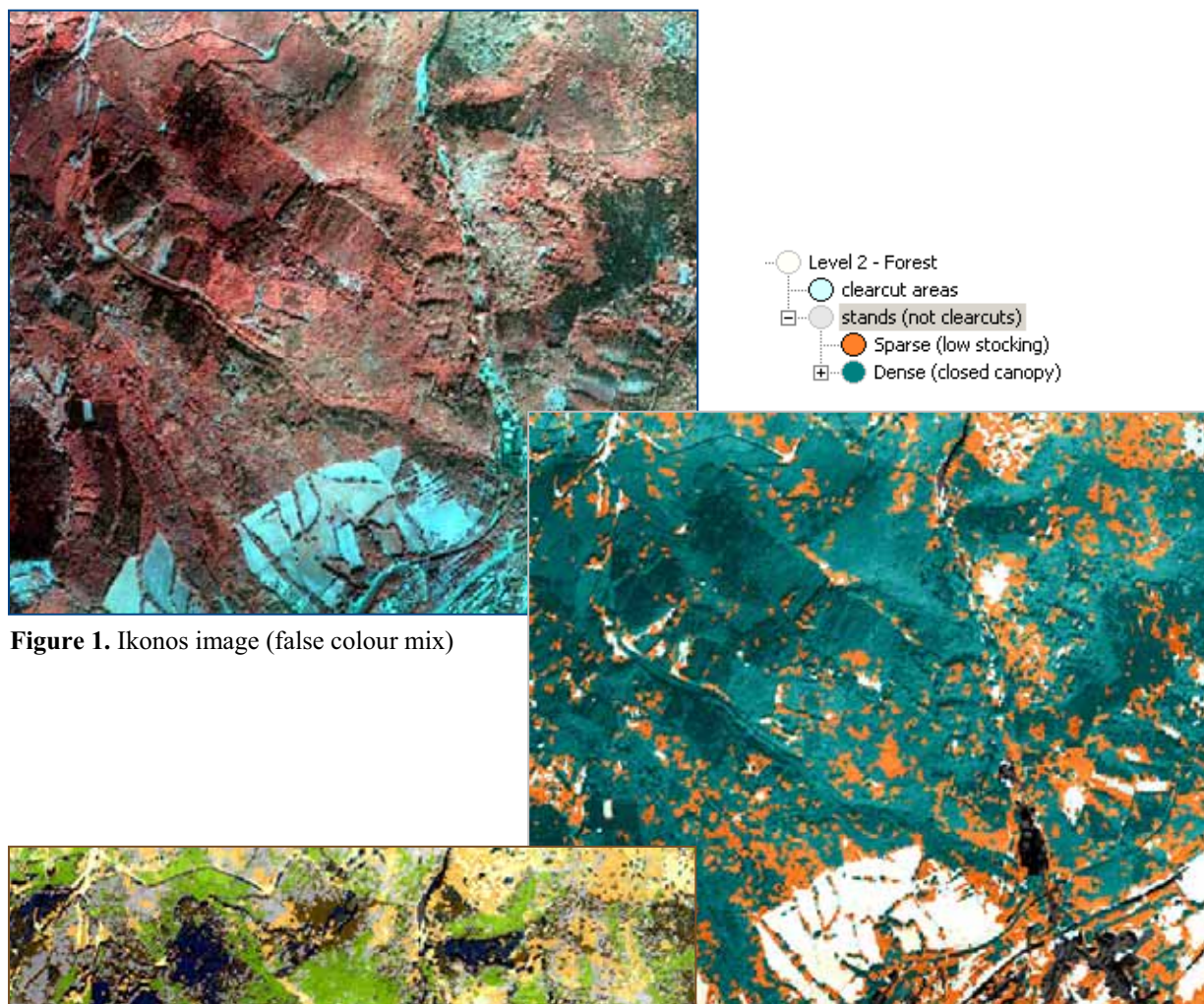


Fig. 1 shows the original image, Fig. 2 shows classification on higher “Forest” level. The result of classification at lower “Stand” level is on Fig. 3.

As indicated in the confusion matrix, proposed method offered very good overall results. Both Picea and Larix conifer species were classified with accuracy over 90%. Fagus achieved accuracy about 70%, which was caused by confusion with class Sparse (high proportion of beech trees). The most problematic tree class was Betula, not only by means of error matrix, but also classification reliability and stability. This tool estimating differences in membership degrees between the best and second best class assignment shows that class Betula and Larix often act as ambiguous. The two species have similar spectral and textural characteristics, especially at young age. Besides, shadows are frequently being confused or mixed with Picea pixels. Best results were obtained for all non-forest classes.

6 DISCUSSION

The results showed that classification of 4-m Ikonos data can be performed with relatively high accuracy. The image enables to estimate tree species composition at the sufficient scale. To get satisfying outcome, additional channels must be calculated in the pre-processing phase and then included in segmentation and subsequent object-oriented classification. The object shape attributes are influenced by the kernel size and layer weight in the process of segmentation, the contribution of additional layers to classification was high for Sobel Edge detector, IHS transformation and low-pass filters. Other textures measures had lower impact using on such spatial resolution data. The classification rules based on fuzzy membership functions are highly convertible, eCognition protocols developed in this project can be transferred and applied (with some threshold modification) to other datasets. To normalize imagery band ratios can be employed, yet data acquired under fixed viewing and solar geometry are recommended to use for the automated analysis.

Previous studies indicated that variations in image acquisition (different projection centres) become more problematic when analysing multitemporal aerial photos. Lower spectral, radiometric and temporal resolutions are also a drawback comparing to VHR satellite data. However, the current lower price at higher spatial resolution still account for aerial photography when developing knowledge base for the method utilization in forestry management. The necessary conditions to obtain good results are standardised screening plan and introduction of photography on IR material. The object-oriented analysis of aerial photos will be examined in the further research.

ACKNOWLEDGMENTS

I would like to thank my supervisor Lena Halounová from Remote Sensing Lab, Czech Technical University Prague, for her endless ideas and constructive comments. Also, I am grateful to Iris Lingenfelder from Definiens Imaging for showing me the right way of eCognition thinking.

REFERENCES

- Halounová, L. 2003. Textural classification of B&W aerial photos for the forest classification. *EARSeL conference*, Gent, Belgium
- Leckie, D.G., Gougeon, F.A., Walsworth, N., Paradine, D. 2003. Stand delineation and composition estimation using semi-automated individual tree crown analysis. *Remote Sensing of Environment* 85, 355–369
- Baatz, M. and Schäpe A., 1999. Object –oriented and multi-scale image analysis in semantic network *Proc. Of the 2nd International symposium on operationalization of remote sensing*, Enschede ITC.
- Haralick, R. and Shapiro, L. 1992. *Computer and Robot Vision vol. I*. Chap. 9. Texture. Addison-Wesley, Reading, USA, pp. 453-494
- Brandtberg, T. 1999. Automatic individual tree-based analysis of high spatial resolution remotely sensed data. *Acta universitatis agriculturae Sueciae*, Silvestria 118,
- Gopal, S. Woodcock, C. 1996. Remote sensing of forest change using artificial neural networks. *IEEE Control Systems Magazine* 8 (3) 42 – 48
- Benz, U. C., Hoffmann, P., Willhauck, G., Lingenfelder, I., & Heynen, M. 2004. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of Photogrammetry and Remote Sensing*, 58, 239– 258
- Culvenor, D.S., Coops, N., Preston, R. Tolhurst, K.G. 1998. A spatial clustering approach to automated tree crown delineation, *Proceedings of automatic interpretation of high spatial resolution imagery for forestry*
- Definiens Inc. 2003. *eCognition object oriented image analysis user guide*. Definiens Inc., Munchen, Germany,

MONITORING NATURA 2000 FOREST HABITATS IN BAVARIA BY THE USE OF ASTER, SPOT5 AND GIS DATA – AN INTEGRATED APPROACH

M. Förster^a, B. Kleinschmit^a and H. Walentowski^b

^a Technical University Berlin, Franklinstr. 28/29 (FR 2-6), D-10587 Berlin,
email: michael.foerster@tu-berlin.de

^b Bavarian State Institute of Forestry, Am Hochanger 11, D-85354 Freising

ABSTRACT

In the context of the EU Habitats Directive, which contains the obligation of environmental monitoring, nature conservation authorities face a growing demand for effective and competitive methods to survey protected habitats. The presented research study aims at evaluating the use of earth observation data of different geometric resolution (ASTER, SPOT5) in order to determine forest habitat types at two test sites, situated in the pre-alpine area in Bavaria (southern Germany). Therefore two modelling approaches (rule-based method with applied Bavarian woodland types and multivariate technique of cluster analysis) were tested for the purpose of detecting potential habitat types. The results were subsequently compared to the terrestrial mapped habitat areas of the NATURA 2000 management plans. The first results show that these techniques are a valuable support in mapping and monitoring NATURA 2000 forest habitats.

Keywords: NATURA 2000, forest habitats, monitoring, SPOT5, ASTER

1 INTRODUCTION

NATURA 2000 sites cover approximately seven per cent of the territory of Germany. The EU Habitats Directive (council directive 92/43/ECC) requires a standardized monitoring of the habitat types and a reporting every six years. For this reason, an operational, economically priced and as far as possible automated application is required. The rapidly developing remote sensing sensor technique and also new image processing methods offer new possibilities to apply remote sensing data for NATURA 2000 monitoring.

The possibilities of remote sensing techniques for detecting and monitoring of biodiversity within the scope of the Habitats Directive has already been proven by EU projects SPIN (Langanke *et al.*, 2004) and EON 2000+ (Sell *et al.*, 2004). Consequently this study is a contribution to an effective implementation of these methods for operational use at the regional level (of German federal states). Therefore these methods have to be cost-effective and highly standardized to be a support in terrestrial mapping processes.

2 SPATIAL MODELLING OF POTENTIAL NATURAL FOREST SOCIETIES

The estimation of the spatial distribution and the kind of NATURA 2000 habitat types is based on the modelling of potential natural forest societies. This is done by the use of expert knowledge about the requirements of the habitat types in respect to site specific factors, such as soil type or relief. Up to now, area-wide information about the potential natural vegetation is available only at a small scale (Seibert, 1968) or in form of statistical data (Walentowski, 2001). In order to derive spatially more detailed facts, a rule-based and a multivariate approach was applied.

As test areas two sites (Angelberger Forst and Taubenberg) in the pre-alpine region of southern Bavaria were chosen. These test sites sized 650 and 1850 hectares, respectively are nearly completely wooded. Within these NATURA 2000 sites, different semi-natural mixed forest types exist, including Beech forests (9110, 9130), Alluvial forests with *Alnus* and *Fraxinus* (91E0), and bog woodland (91D0). To evaluate the results of this study forest management plans including mapped woodland habitats were available for the two selected areas.

For the modelling of the potential natural forest societies a digital terrain model (DTM 5 and DTM 25), a conceptual soil map (1 : 25.000) as well as a forestry site map were used. To exclude regional knowledge and receive results transferable to other NATURA 2000 sites, the data was processed without a field survey.

2.1 RULE-BASED METHOD

The purpose of this easily comprehensible and rule-based method was to model potential natural forest societies in areas with identical natural woodland composition (Walentowski *et al.*, 2004). The test sites are situated in the region of mountainous mixed forest (Taubenberg) and beech forest (Angelberger Forst). For habitat types, which could exist in this natural woodland composition, a register of location factors was developed, including soil type, relief type, water balance, and site related additional attributes, such as the location of very dry areas. Furthermore sites with a high (H) suitability and sites where the existence of the habitat type is generally possible (P) or excludable (E) were distinguished. Based on the existing suitability the geo data were combined to a set of rules (see Figure 1):

1. Modelling of locations with high (H) and excludable (E) suitability for a habitat type. As an example, the possibility of the existence of bog woodland (91D0) on brown soils will be excluded, while the existence on peaty soils is highly possible.
2. Modelling of the possible occurrences (P) of a habitat type. A case in point would be the suitability of Luzulo-Fagetum beech forest (9110), which has a wide range of possibilities to occur. This habitat type can grow on different relief types (southern exposition, steep slopes, hilltops) as well as on different soil types (sand, gravel). These geo factors are not spatially exclusive. Therefore a site can be chosen as a possible habitat type because of one or several parameters. In this case, the number of possible occurrences was summed up.
3. The calculated site qualities for the habitat types were combined. Firstly a potential forest society is chosen for areas with high suitability (H). At sites without H the habitat type with the largest number of possible occurrences was selected.
4. At each grid cell the dominating habitat type will be chosen as the potential natural forest society. The result is a complete spatial database of the potential natural vegetation.

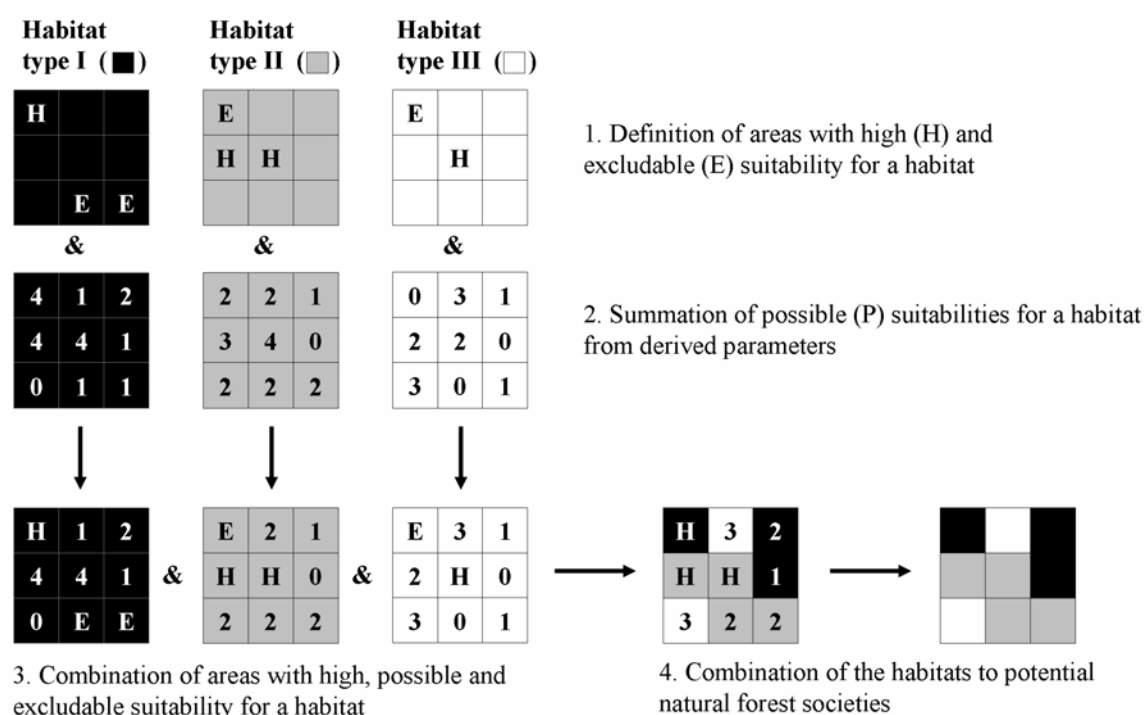


Figure 1. Scheme of the rule-based model

2.2 MULTIVARIATE CLUSTERING OF RELIEF TYPES

The multivariate clustering method of relief types (developed from SciLands inc.) assumes that distribution patterns of various site specific factors are significantly affected by changes of the terrain. Classifying the terrain into landscape ecologic relevant morphographic units can give important evidences concerning the distribution of woodland societies. Therefore three autonomous relief type categories of the terrain were calculated:

- Subdivision of the relief in summit areas, bottom areas, slopes, and closed depressions.
- Differentiation of areas with convergence of discharge (e. g. vales), areas with divergence of discharge, and intermediate areas, which mediate between convergence and divergence.
- The terrain is organized into areas with a slope perpendicular to contour lines and into areas which are steepened or flattened relative to the extremes.

The calculated subdivisions were summarized to clusters. The applied approach is a combination of the iterative minimum distance method (Forgy, 1965) and of the hill climbing routine (Rubin, 1967). The differentiated clusters were assigned to the possible habitat type in the pre alpine area. The method relies exclusively on the digital terrain model (DTM). Hence only habitat types which strongly depend on relief parameters, such as Tilio-Acerion forests of slopes, screes and ravines (9180), can be processed. Because of the higher intensity of the relief only the NATURA 2000 area Taubenberg was used for the multivariate clustering.

3 CLASSIFICATION OF FOREST TYPES

In order to identify forest habitat types, the modelled potential natural forest societies had to be combined with a classification of the in situ vegetation. To guarantee a possible implementation into the workflow of the local authorities, the acquired data had to be cost-effective and processable using standard methods. The satellites SPOT5 and ASTER were considered suitable for the differentiation of coniferous, deciduous, and mixed forest as they offer a spatial resolution of 5 m to 15 m and spectral bands in the infrared and near infrared region.

The satellite scenes of SPOT5 (acquisition date: 07.09.04) and ASTER (acquisition date: 14.09.04) were georeferenced using the DTM. Furthermore a Minnaert correction was carried out with the objective of topographic normalization (Kleinschmit *et al.*, 2005). On the basis of scanned and georeferenced true color air photographs, training areas for all existing principal tree species and types of mixture were defined. On the basis of these training areas a Maximum Likelihood classification was carried out. The result was summarized to higher-ranking classes and validated by the use of forest management maps and local knowledge of forest officials.

For the test site Angelberger Forst the classes deciduous (young), deciduous, coniferous, and mixed forest were detected, while for the test site Taubenberg the classes deciduous, coniferous – non fir, coniferous – fir and mixed forest were recognized. The results correspond to the experiences of Blaschke *et al.* (2003), who detected similar classes with ASTER and SPOT4. Finally the scenes were segmented in two levels of detail. The majority class in one object was assigned as attribute to each polygon in order to receive more homogeneous classes.

4 IDENTIFICATION OF HABITAT TYPES

The results of the satellite classification were combined with the methods described in section 2 in order to calculate the potential natural forest societies. The potential forest habitat type is only selected if the classified real vegetation corresponds to the topmost forest society (see Figure 2). An example would be the detection of a potential beech forest society, which is not considered a potential habitat type if the satellite classification result is coniferous forest.



Figure 2. Potential natural forest societies (left) and derived potential habitat types for the test site Taubenberg (example from rule-based method)

5 COMPARISON OF MODELLLED AND TERRESTRIAL MAPPED HABITAT TYPES

The results were compared to the terrestrial mapped habitat types from the forest management plans (see table 1). A great part of the considered habitat types could be detected by the use of the model approaches. The best results were achieved by using the satellite data with higher spatial resolution (SPOT5). An additional factor for good correlation of the results is the utilization of a forestry site map (FSM) within the model. It was evident that the models tend to underestimate habitat types with distinct site specific growing conditions, such as Alluvial Forest (91E0). In such cases the clustering technique yields better results. Thus it was possible to detect 80 % of the areas with Tilio-Acerion forests of slopes, screes and ravines (9180), while the rule-based method achieved only 40 % of this habitat type.

Table 1. Exemplary comparison of the results of the rule-based method with the forest management plan. Additionally the test site Angelberger Forst was processed with and without a forestry site map (FSM). The term “n.a.” stands for “not available” because the habitat type is inexistent in this area.

Habitat types	Angelberger Forst with FSM (in %)	Angelberger Forst without FSM (in %)	Taubenberg without FSM (in %)
9110	95.4	92.3	62.5
9130	70.6	66.0	76.9
9180	n.a.	n.a.	40.0
9410	n.a.	n.a.	66.7
91D2 – 91D4	n.a.	n.a.	80.0
91E0	91.7	58.8	46.2
9160	88.2	73.9	n.a.

A more detailed comparison, which involved the detected area size of model results and forest management plans, resulted in increased differences. As an example, the areas of Alluvial Forest of the data sets intersect only to 33.2 %. This is due to the model assumption that a habitat type is necessarily in the condition of the potential natural vegetation. However, a modelled and detected deciduous beech forest can be an artificial planted forest of neophytes. Therefore, it would be impossible to detect these kinds of errors by the use of the methods described above.

Another point of discussion are the methods of terrestrial mapping. Often the transitions in between habitat types or between a habitat type and other forest are not clearly to detect. In addition habitat types will be mapped even if the main species covers only a small proportion of an area. The intention in

mapping these habitat types is to ascribe a higher importance to them, because there is a potential development of the forest society. As a consequence terrestrial mapped habitat types cannot be detected with remote sensing techniques (especially in cases of low percentages of deciduous forest).

The results could be improved if a buffer area surrounding the modelled habitat type would be applied as “suspected potential habitat type”. For the rule-based method the accuracy could significantly be increased by using the second dominant modelled habitat type (see figure 1). When combining potential habitat type and potential natural forest society the subordinated forest society could be kept as possible result.

6 SUMMARY AND OUTLOOK

The consultation with local forest authorities showed that the developed method could be a support for terrestrial mapping processes. However the applied methods leave still room for improvement. The results of the modelling will be improved using the model BERN (Schlutow *et al.*, 2004), which detects the site related factors of habitat types as fuzzy relations using fuzzy logic methods.

Furthermore, the classification goodness of the real forest vegetation strongly depends on the spatial and spectral resolution of the sensor. Thus with the use of air photographs and high resolution satellite imagery (Quickbird) the possibilities of remote sensing techniques for detecting and assessing habitat types will be analysed in more detail.

ACKNOWLEDGMENTS

The Authors would like to thank Mr. Seitz and Dr. Fischer from the Bavarian Forest Management Authority who critically evaluated our results. Additionally we would like to thank SciLands inc. (Mr. Bock) for the processing of the cluster analysis.

REFERENCES

- Blaschke, R. and Felbermeier, B. 2003. Satellitendaten für die Forstwirtschaft - Nutzeranforderungen der forstlichen Praxis und Potential der optischen Satellitenfernerkundung. *Forstliche Forschungsberichte*. 196:93-108.
- Forgy, E.W. 1965. Cluster Analysis of Multivariate Data: Efficiency versus Interpretability of Classifications (abstract.). *Biometrics*. 21: 768.
- Kleinschmit, B. and Coenradie, B. 2005. Entwicklung von Methoden zum Monitoring von Trockenheitsfolgeschäden unter Nutzung von hochauflösenden Fernerkundungsdaten. Report, TU-Berlin, Germany, pp.9-10.
- Langanke, T., Blaschke, T. and Lang, S. 2004. An Object-Based GIS / Remote Sensing Approach Supporting Monitoring Tasks in European-Wide Nature Conservation. First Mediterranean Conference on Earth Observation (Remote Sensing), April 21-23 2004, Belgrade, pp 245-252.
- Rubin, J. 1967. Optimal Classification into Groups: An Approach for Solving the Taxonomy Problem. *J. Theoretical Biology*. 15:103-144.
- Schlutow, A. and Hübener, P. 2004. The BERN Model. Bioindication for Ecosystem Regeneration towards Natural Conditions. Research Report 200 85 221, Federal Environmental Agency, Berlin, Germany.
- Seibert, P. 1968. Übersichtskarte der natürlichen Vegetationsgebiete von Bayern 1:500.000 mit Erläuterungen. *Schriftenreihe für Vegetationskunde*. 3.
- Sell, E., Grubinger, J. and Eisl, M. 2004. Earth Observation for Natura 2000+ - Eine Bottom-up-Studie – Ergebnisse aus Deutschland und Österreich. AGIT, July 7-9 2004, Salzburg, Austria, pp 626-631.
- Walentowski, H., Ewald, J., Fischer, A., Kölling, C. and Türk, W. 2004. *Handbuch der natürlichen Waldgesellschaften Bayerns*. Geobotanica, Freising, pp. 441.
- Walentowski, H. 2001. Waldland Bayern – ein Buchenmeer? *LWF-Aktuell*. 30:1-3.

Session 4a

HIERARCHICAL COMBINATION OF DATA FROM SATELLITE IMAGERY AND AERIAL PHOTOGRAPHS FOR MULTI SOURCE FOREST INVENTORY

Sakari Tuominen

Finnish Forest Research Institute, National Forest Inventory
Unioninkatu 40 A, FIN-00170 Helsinki. Finland

Phone: +358 10 2111

Email: sakari.tuominen@metla.fi

ABSTRACT

Medium resolution (e.g. Landsat TM) satellite imagery has been utilized in many multi-source forest inventory applications for estimating forest characteristics. The main advantages of the satellite imagery are large area coverage and wide spectral range. However, due to the relatively coarse spatial resolution the accuracy of satellite-based forest estimates has been relatively poor at the stand and sample plot levels. Another widely available data source for forest inventories is digital aerial photography. The main advantage of the aerial photographs is their superior spatial resolution compared to satellite imagery. The digital interpretation of aerial photographs, however, is complicated because of factors such as bi-directional reflectance, which cause the spectral values of the image pixels to depend, in part, on their location in the image. On the other hand, the high spatial resolution of aerial photographs makes it possible to utilize the textural information in predicting the forest parameters. In this study Landsat 7 ETM satellite image spectral features and color-infrared aerial photograph texture features were combined applying different hierarchical approaches for estimating forest stand characteristics. The methods that were developed and studied for combining satellite and aerial photograph information were based on pre-stratification and nearest neighbor estimation within the derived strata, and hierarchical nearest neighbor estimation approaches. The results indicate that the estimation accuracy of forest characteristics can be significantly improved by utilizing a hierarchical approach in combining different remote sensing imagery.

Keywords: multi-source forest inventory, satellite imagery, aerial photographs, image texture

1 INTRODUCTION

In Finland, the data acquisition for forest management planning has been traditionally based on stand-level field inventories. In its present form the method is based on the use of aerial photographs for delineating the inventory units (i.e. stands) and a field assessment of the stand characteristics, including its final delineation, by a forestry professional. Typically, all stands in the inventory area are visited in the field, and thus, the method requires high amount of fieldwork and skilled professional staff. An apparent drawback is the subjectivity of the method. Delineation of the stands and selection of measurement points within the stands depends on the person carrying out the inventory, which may result in bias in the inventory data. Furthermore, the stand borders also tend to change between consecutive inventories due to silvicultural operations and natural disturbances, which makes the compartments unsuitable for monitoring. However, the main problem to be solved in the present Finnish forest management planning system is the disparity between the required amount of fieldwork and the resources allocated for the work. Thus, new methods have to be introduced in order to increase the efficiency of the forest management planning. One option that is strongly brought forward for rationalizing the forest management planning system is reducing the fieldwork through the increased use of remote sensing imagery.

The Multi Source National Forest Inventory (MSNFI) of Finland has utilized optical satellite imagery (mainly Landsat TM) to produce estimates and statistics of forest attributes for national and provincial level (Tomppo, 1991; Tomppo 1993). The forest inventory method used in MSNFI is based on combining the information of a systematic sample of field plots, satellite images and digital map data. Utilizing the forest estimates produced by MSNFI has been examined as one option for enhancing the forest management planning work. However, it has been concluded that the accuracy of the optical satellite

image based estimates has been relatively poor at a stand level (e.g. Tokola *et al.*, 1996; Holmgren & Thuresson, 1998; Poso *et al.*, 1999; Mäkelä & Pekkarinen 2001). One reason suggested for the high stand- and plot-level estimation errors is the limited spatial resolution of the satellite image material in relation to the typical forest characteristics. The average stand size in Finland is relatively small, approximately 1.5 - 2 ha in Southern Finland, and therefore a typical stand consists of a small number of satellite pixels. Because of the small size of the stands a considerable proportion of these pixels are mixed, i.e. they also carry spectral information from adjacent stands, and thus, they may represent poorly the stand spectral properties. On the other hand, for forest management planning inventories described earlier, full aerial photograph coverage is typically acquired. This would make it possible to combine the high spectral resolution of satellite imagery used for MSNFI and the high spatial resolution of aerial photographs used for management oriented inventories. If combining the data from satellite images and aerial photographs in forest inventory could significantly improve the accuracy of local forest estimates of the current MSNFI, this would make it possible to reduce to amount of expensive fieldwork in forestry planning.

Until now, the National Forest Inventory (NFI) has gone through entire Finland with an interval of approx. ten years - the areas of one or two provincial Forestry Centers have been inventoried each year. This means that in some areas the inventory data can be a decade old. This is now considered to be unacceptable, and at present the NFI is being developed towards more up-to-date inventory system. This includes shorter intervals in the measurement of the field sample plots of the traditional NFI, and also shorter intervals of updating the MSNFI estimates based in satellite image interpretation. This would improve the applicability of the MSNFI data for forest management planning purposes.

2 MATERIALS AND METHODS

2.1 STUDY AREA AND FIELD DATA

The study area is located in Northern Karelia, in the municipalities of Kontiolahti and Eno. The size of the study area is 10 x 10 km. The area consists mainly of managed forest in private ownership. The landscape in the study area is North-Karelian hill country with elevation between 100 and 250 m a.s.l. The terrain depressions are covered by lakes or peatlands and higher elevations by wooded hills. The field data in the study area was measured from a systematically located grid of approximately 600 field sample plots. The field plots were measured in year 2000 (Katila & Tomppo, 2005). The field measurement generally followed the guidelines of the Finnish NFI.

2.2 REMOTE SENSING IMAGERY

The study area was covered by Landsat 7 ETM image (path 186, row 16) acquired on 10 June 2000. The image was geo-referenced on the basis of digital map data. The spatial resolution of the geo-referenced image was 25 m.

Color-infrared aerial photographs covering the study area at scale 1:30000 were acquired in 2000 and 2001. The photographs were digitized using red, green, and blue filters representing near-infrared (NIR), red (R) and green (G) channels. The images were orthorectified by the National Land Survey of Finland (using ground control points and a raster DEM), and resampled to a pixel size of 0.5 m. Finally, a photo mosaic covering the study area was composed of the images.

The spectral values of all Landsat ETM channels were utilized as satellite image features in this study. The aerial photograph features that were utilized in this study included the spectral averages of aerial photograph pixels. Additionally textural features consisting of features based on image gray level co-occurrence matrices and gray level standard deviations within pixel blocks were utilized (Haralick *et al.*, 1973; Haralick, 1979; Wang *et al.*, 1997). The aerial photograph features were extracted from square-shaped windows (size 20 x 20 m) surrounding the sample plots.

The textural features of aerial photographs were considered suitable for estimation of forest attributes, since the aerial photograph resolution is finer or similar in relation to the size of the objects of interest (i.e. trees) in the image. Thus, the local variation of the pixels should correspond to the forest structure. In Landsat satellite images the resolution is coarse in relation to the objects, and single pixels contain many objects causing low local variance of pixels (Woodcock & Strahler, 1987). Thus, the textural features of satellite imagery have generally performed poorly in forest inventory (e.g. Shang & Waite, 1991; Cohen & Spies, 1992).

For the estimation of forest attributes the aerial photograph features were standardized to a mean equal to 0 and standard deviation equal to 1, since the original image features had very diverse scales of variation due to the different ways of extracting the features. Without the standardization, the variables with large variation would have higher weights in the estimation, regardless of their correlation with the estimated forest attributes.

2.3 ESTIMATION OF FOREST ATTRIBUTES

Different combinations of satellite image and aerial photograph data were used the estimation of forest attributes of the field sample plots. K nearest neighbors (k-nn) method was applied in the estimation (e.g. Muinonen & Tokola 1990, Tomppo 1991). The following forest attributes were estimated: mean diameter at breast height (D), mean height (H), basal area (BA) and volume (VOL) of total growing stock. The nearest neighbors were determined by the Euclidean distances between the observations in the feature space determined by the satellite image and aerial photograph features. The estimates were calculated as the average values of the stand variables of the k nearest neighbors.

Along with the k-nn, stratification based on the image features was utilized in the estimation of forest attributes. The stratification was carried out using k-means clustering algorithm (MacQueen, 1967). Different numbers of strata (16-20-24) were tested since the result of stratification may be sensitive to the number of strata employed. Mean vector estimator was applied within the strata for estimating forest attributes.

The accuracy of the estimates derived using different estimation methods was tested using leave-one-out cross-validation technique, in which each sample plot in turn was left out from the set of reference plots and estimated with the aid of the remaining plots. The estimates of the field plots were compared with the corresponding ground truth values of the plots. The accuracy of the estimates was measured by their root mean square errors (RMSE).

Alternative methods studied for utilizing satellite imagery and aerial photographs in the estimation of forest attributes:

- A. K-nn estimation based on satellite image spectral features
- B. K-nn estimation based on aerial photograph spectral and textural features
- C. K-nn estimation based on a combination of satellite image and aerial photograph features
- D. Stratification based on satellite image spectral features and k-nn estimation based aerial photograph spectral and textural features within the strata
- E. Stratification based on a aerial photograph spectral and textural features and k-nn estimation based on satellite image spectral features within the strata
- F. K-nn estimation based on satellite image spectral features and k-nn estimation based on aerial photograph spectral and textural features. The final estimates were calculated as averages of the individual estimates produced by each data source.
- G. K-nn estimation based on satellite image spectral features and k-nn estimation based on aerial photograph spectral and textural features. The final estimates were calculated as weighted averages of the individual estimates weighting them by the inverse values of their mean square errors (MSE) (Tuominen & Poso, 2001)

In the k-nn estimation the k was set to 5. When applying estimation methods D and E, the k was set equal to the number of field plots in those strata, where the number of field plots was < 5 . It has been shown that increasing the number of nearest neighbors (value of k) generally improves the accuracy of the estimates (Tokola *et al.*, 1996, Nilsson, 1997), until the accuracy of the estimates stabilises at some level, and increasing the k after that does not produce further gains in the estimation accuracy. In this study, the number of field plots was quite low, and thus a relatively small value of k was considered appropriate.

3 RESULTS

The results show that for estimating local forest attributes the aerial photograph features perform significantly better than Landsat ETM features. The accuracy of the estimates is higher than presented by e.g. Poso *et al.* (1999), who used Landsat 5 TM and color-infrared aerial photograph data. Furthermore, when combining the aerial photograph and satellite image features into a single data set for k-nn estimation (estimation method C), the estimation accuracy is lower or approximately similar than when using only aerial photograph features.

When using hierarchical approaches in combining satellite image and aerial photograph data (estimation methods D-G) the estimation results were divergent. In methods combining stratification based on one data source and k-nn estimation within the strata based on another data source (methods D, E), the estimation accuracy was significantly lower than in other methods utilizing both data sources. The best estimation results judged by RMSE values were obtained in methods, where k-nn estimation was carried out separately within data sources (satellite or aerial origin), after which the estimates were combined together (methods F, G). The root mean square errors for estimates of stand diameter, height, basal area and volume are presented in Table 1.

Table 1. Root mean square errors of forest attribute estimates different estimation methods

Estimation / data combining method	Diameter, cm	Height, m	Basal area, m ² /ha	Volume, m ³ /ha
A	8.03	5.50	7.21	76.07
B	7.33	5.03	6.75	71.17
C	7.40	5.13	6.68	72.24
D	7.79	5.37	7.03	73.16
E	7.87	5.37	6.88	73.64
F	6.96	4.80	6.33	67.32
G	6.94	4.78	6.32	67.21

4 CONCLUSIONS

The results show that using a combination of different remote sensing data sources may result in more accurate estimates than those obtained using individual data sources. The typical problem when combining different data sources for estimators like k-nn is that it is difficult to control the weighting of different input variables. The usefulness of any input variable can be studied by measuring the correlation between image features and forest attributes, but, on the other hand, the image features are often highly correlated with each other, and adding extra variables having high correlation with the other variables does not improve the estimation. Furthermore, when using estimation methods based on measuring distances between the sample plots in the feature space (such as k-nn or k-means clustering), the features having large variation receive high weights, unless the image features are standardized for similar scales of variation. However, after standardization the input variables receive similar weight regardless of their potential in estimating forest attributes.

Determining the correct weights of the input variables in k-means or k-nn estimation while simultaneously taking into consideration their variation and usefulness for the estimation is a complicated task. The problem can be managed at least in some extent by a hierarchical approach in combining the data sources. This kind of approach, for example, makes it possible to ensure that a dissimilarity of some features (e.g. aerial photograph texture) is not drowned by a general similarity of other features (e.g. spectral properties of aerial photographs and satellite imagery).

In this study, the methods combining k-nn estimation and k-means clustering did not perform well. A part of the problem may be the selection of nearest neighbors within the strata, which is problematic in small strata where the k approaches the total number of field plots in a stratum. In that case the data used for stratification gets a major weight leaving minor or none weight for the data used for k-nn. This kind of combination would require an adaptive k-nn procedure allowing the user to adjust the k in relation to the size of the strata. The best performing methods in this study were the ones applying separate k-nn estimation for each data source, and combining the individual estimates. These may be given equal weights or weights derived as inverse mean square errors of the individual estimates. The 1/MSE weighted estimates had the highest accuracy for all forest attributes, confirming the results Tuominen and Poso (2001).

REFERENCES

- Haralick, R. M., Shanmugan, K. & Dinstein, I. 1973. Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics*. Vol. SMC-3, no. 6: 610-621.
- Haralick, R. 1979. Statistical and structural approaches to texture. *Proc. IEEE*, 67(5),786-804.
- Holmgren, P. & Thuresson, T. 1998. Satellite remote sensing for forestry planning - a review. *Scandinavian Journal of Forest Research*, 13: 90-110.
- Katila, M. & Tomppo, E. 2005. Sampling simulation on multi-source wall-to-wall forest maps - an application for small areas. *Manuscript*.
- MacQueen, J. 1967. Some methods for classification and analysis of multivariate observations. Volume 1 of *Proceedings of the Fifth Berkeley Symposium on Mathematical statistics and probability*, pp 281-297. Berkeley, 1967. University of California Press.
- Muinenen, E. & Tokola, T. 1990. An application of remote sensing for communal forest inventory. In: *The usability of remote sensing for forest inventory and planning. Proceedings from SNS/IUFRO workshop 35-42*. Umeå, Sweden: Remote Sensing Laboratory, Swedish University of Agricultural Sciences (Report 4).
- Mäkelä, H. & Pekkarinen, A. 2001. Estimation of timber volume at the sample plot level by means of image segmentation and Landsat TM imagery. *Remote Sensing of Environment*, 77:66-75.
- Poso, S., Wang, G. & Tuominen, S. 1999. Weighting alternative estimates when using multi-source auxiliary data for forest inventory. *Silva Fennica* 33: 41-50.
- Tokola, T., Pitkänen, J., Partinen S. & Muinenen, E. 1996. Point accuracy of a non-parametric method in estimation of forest characteristics with different satellite materials. *International Journal of Remote Sensing* 17: 2333-2351.
- Tomppo, E. 1991. Satellite image-based national forest inventory of Finland. *International archives of photogrammetry and remote sensing* 28: 419 - 424.
- Tomppo, E. 1993. Multi-source national forest inventory of Finland. In *Proceedings of Ilvessalo Symposium on National Forest Inventories*, 17-21 Aug. 1992, Finland. Edited by Aarne Nyysönen, Simo Poso and Johanna Rautala. *The Finnish Forest Research Institute Research Papers* 444: 52 – 60.
- Tuominen, S. & Poso, S. 2001. Improving multi-source forest inventory by weighting auxiliary data sources. *Silva Fennica* 35: 203-214.
- Wang, G., Waite, M.-L. & Poso, S. 1997. SMI user's guide for forest inventory and monitoring. University of Helsinki, *Department of Forest Resource Management Publications* 16.
- Woodcock, C.E. & Strahler, A.H. 1987. The factor of scale in remote sensing. *Remote Sensing of Environment* 21: 311-332.

FOREST STRUCTURAL PARAMETERS DESCRIPTION THROUGH THE USE OF VARIOGRAMS OF REAL AND SIMULATED IMAGES

S. Merino-de-Miguel^{1,2}, J. Solana-Gutiérrez¹, F. González-Alonso²

¹ Department of Forest Economy and Management, UPM. Ramiro de Maeztu, s/n Madrid 28040 Spain

² Remote Sensing Laboratory-INIA. Ctra. A Coruña, km 7.5 Madrid 28040 Spain.

Email: smerino@montes.upm.es, smerino@inia.es

ABSTRACT

This study presents a quantitative analysis of the sensitivity of textural information of high resolution aerial photographs and artificial images to a number of forest structural parameters (basal area, density, forest cover, diameter at breast height, stem wood volume with bark...) as derived from the Spanish Second National Forest Inventory. Artificial images were simulated using an geometric-optical model. Influence of topographic effects, quantified through the use of slope and aspect, was also investigated. Image texture was quantified using the experimental variogram (calculated as an average of four directional experimental variograms), one of the most reliable and easy-to-compute geostatistic tools. Reached results revealed that there is a complex dependency of variogram parameters (range, sill, nugget effect, slope of the linear part of the variogram) on forest structural parameters. However, this dependency seemed to appear simpler when topographic effects were removed. Variogram parameters were extracted from the mathematical functions that were previously fitted to the experimental variograms.

Keywords: variogram, texture, topographic effects, aerial photographs.

1 INTRODUCTION

The estimation of forest structural parameters is one of the main components of studies dealing with vegetation dynamics and land cover changes (O'Dwyer, 2002). If that estimation has to be undertaken using remotely-sensed data, the study of such parameters usually refers to those that can be assess from above (O'Dwyer, 2002). Remote sensing data and techniques might be considered as one of the premier observational and analytical tools for use in scientific understanding and managing of the world's forests (Franklin and Wulder, 2003).

Much of the work done in the remote sensing of forested ecosystems has been in the spectral domain (Cohen et al., 1990) which typically involves the analysis of multispectral images using statistically based decision rules. Since this approach tended to produce bad results it was suggested that the extraction of information from remote sensing images of vegetated surfaces would greatly benefit from the use of spatial data (Woodcock et al., 1988). This is based on the fact that the three-dimensional structure of vegetation (e.g. tree crowns in a forest canopy) is related to image texture (Bruniquel-Pinel and Gastellu-Etchegorry, 1998).

Spatial structure becomes apparent in remotely-sensed images when the scene contain discrete objects that are identifiable because their spectral properties are more homogeneous within than between them and other scene elements (Jupp et al., 1988). The spatial structure can be measured using different tools, commonly known as geostatistics. Geostatistics has been widely applied in the Earth sciences over the last decades and particularly widespread in the remote sensing community (Atkinson et al., 1996). For the present study, variogram was the selected geostatistical tool because of its straightforward computation, ease of interpretation and because the assumptions underlying its application are not very rigid (Woodcock et al., 1988).

The study presented herein investigates whether the spatial information (or texture) contained in high spatial resolution imagery and extracted using the variogram, can provide accurate estimations of individual forest parameters. This study presents a quantitative analysis of the structure of *Pinus pinaster* forests found in the province of Madrid (Spain) using: (i) real and simulated images, (ii) field data from the Spanish Second National Forest Inventory, and (iii) forest and topographic cartographies. In particular, this investigation is focused on analysing the sensitivity of image textural information to a number of forest parameters (basal area, density, forest cover, diameter at breast height, stem wood volume with bark...). Influence of topographic effects on image spatial information is also investigated through the analysis of

two of their most characterizing parameters: slope and aspect, as derived from a digital elevation model. Results reached so far revealed that there is a complex dependency of variogram parameters (range, sill, etc) on forest parameters. However, this dependency seemed to appear simpler when topographic effects were removed.

2 STUDY AREA

The present work involves the study of all the *Pinus pinaster spp mesogensis* patches found within Madrid province, in central Spain. This specie grows alone or mixed mainly with other *Pinus* species such as *Pinus sylvestris*, *Pinus pinea* or *Pinus nigra*. It is Mediterranean tree, capable of resisting long winters and dry and hot summers, and it appears in either even-aged or uneven-aged forests. The study area is as large as 18 000ha.

3 MATERIAL

In the present study four different sets of data were used: (i) aerial photographs, (ii) field data, (iii) a digital elevation model (DEM) and (iv) a forest map. Some topographic maps (1:50 000 and 1:5 000) were also used for supporting some processing.

Aerial photographs came from a 1:6 500 scaled flight dated in 1990-91. These photographs were scanned and geometrically corrected through the use of ground control points, camera parameters and a DEM. Once orthorectified, these photographs were subseted into mini-scenes (an example is shown in figure 1a) corresponding to those places where field work had been carried out. The mini-scenes were as big as 100 by 100m.

Field data came from the Second Spanish National Forest Inventory (2NFI) sample plots, measured during 1990. A total of 72 sample plots were considered within the present work. Mean values for each of the forest ground variables considered (basal area, density, forest cover, diameter at breast height, stem wood volume with bark) were calculated using individual tree measurements (NFI database).

The DEM, as derived from the 1:5 000 topographic map, was used for both geometrically correcting the aerial photographs and for deriving slope and aspect within the limits of the mini-scenes (an example is shown in figure 1b). The slope and aspect of any mini-scene were the slope and aspect of a plane fitted to the DEM. Forest cartography (that was provided together with the NFI sample plots) was used for general approximation to the study area and for stratification purposes.

Among these materials, more than 4 000 scenes (an example is shown in figure 1c) were artificially simulated using a geometric-optical model as designed and implemented by Lewis (1996; 2002). Variables defining simulated images fold into various categories: (i) variables defining scene objects (density - number of trees per unit area - and crown size), (ii) variables defining terrain characteristics (slope and aspect), (iii) variables defining sun position (azimuth and zenith angles), etc. Ranges of variation for such variables were extracted from real data (2NFI, DEM, time and date of aerial photographs).

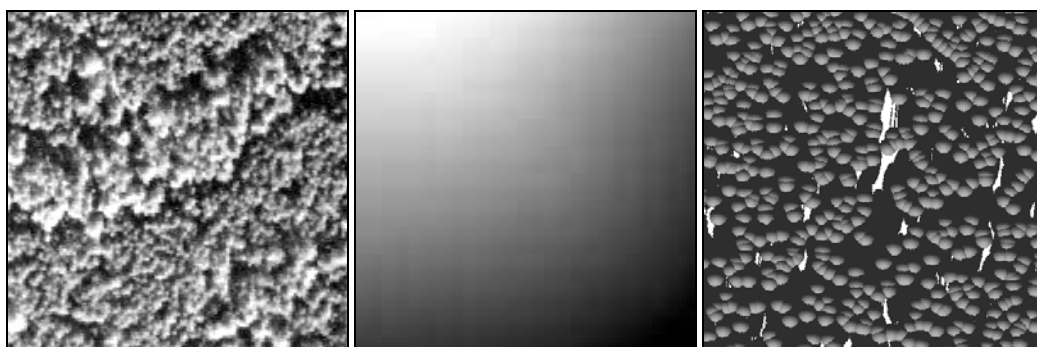


Figure 1. From the left to the right: (a) mini-scene aerial photograph around a 2NFI sample plot (numbered 110), (b) mini-scene DEM corresponding to sample plot 110, (c) example of simulated image.

4 METHODS

The textural measure used to study the structure of forest canopy from aerial photographs and simulated images was the variogram. The variogram was chosen because it offers a simple and efficient method by which to quantify textural patterns that in turn are strongly related to canopy structure (St-Onge and

Cavayas, 1997). The effect of the spatial resolution on the relationship between forest and texture parameters was also investigated, thus original images (either real or simulated) were degraded from the original 1m pixel size to a 10m pixel size by an averaging method. Finally, the influence of topography in image texture was also analysed through the use slope and aspect as derived by a plane fitted to the corresponding mini-scene DEM. Slope and aspect in the case of simulated images were those values defining the terrain characteristics of the scene and therefore they need not to be calculated.

All the experimental variograms (calculated as an average of four directional variograms) were fitted with a theoretical model in order to quantify the texture parameters of interest: range, sill, nugget effect and slope of the linear part of the variogram (from now on referred to as slope). The appropriate model was chosen on the basis of its simplicity and its goodness of fit which can be visual and quantitative analysed (McBratney and Webster, 1986), in this case through the use of the root mean squared error and the Akaike information criterion. Used models include: spherical and exponential plus linear, with or without considering the nugget effect. The fitting to the experimental variogram was done by the recommended method of least squared approximation (Oliver *et al.*, 1989). The maximum lag or distance at which the model was fitted to the experimental variogram was finally set to 25m (experimental variogram was calculated over one third the size of the mini-scene, that was 33m). Both sets of data (real and simulated images) were processed in the same way.

5 RESULTS AND DISCUSSION

The first step in the analysis of the results consisted on looking for any kind of relationship or correlation between field and geostatistical data as derived from the analysis of aerial photographs. The best results were found between range and stem volume with bark, although in general there was not any evident relation. Provided that this low correlation might be due to a great variability within field data (Hudak and Wessman, 1998), some clustering was carried out. However results did not improve that much. The analysis of simulated images neither produced, at least in an easy and apparent manner, a strong correlation between any of the variables defining the scene (crown radius, density and canopy cover) and any of the variogram parameters (range, sill, nugget, slope).

The spatial structure of an imaged scene is a function of the relationship between the size of the objects within the scene and the size of the resolution cells (Jupp *et al.*, 1988). Therefore, if the spatial resolution varies, the spatial structure of the image also varies. The pattern of variation was identified by Woodcock *et al.* (1988) and Bruniquel-Pinel and Gastellu-Etchegorry (1998): as the spatial resolution of the image is degraded the height of the sill decreases and the range increases, leading the variogram to a simpler shape. Results reached in the present work confirmed this point.

Some authors realised about the potential effect of the topography on the variogram and its interaction with viewing and illumination conditions of the scene (Atkinson *et al.*, 1996). The effect of slope and aspect on real and simulated images was clear when visual analysis was performed on them. The key point was to find out whether or not it could be demonstrated that slope and aspect affect the way the spatial structure appears in the imaged scene and if it was possible to quantify such an effect. These analyses were just performed over simulated images since this kind of data allowed us to vary only the factor of interest while the other factors remain constant.

In order to carry out the analysis of the influence of slope and aspect on the variogram, simulated images were clustered into groups where crown size and number of trees (density) stayed constant. This way of showing the results revealed that range and sill were affected by slope and aspect in a quantifiable and apparent manner, as shown in figure 2a. As expected, for a slope of 0°, sill stayed constant (because is independent of aspect) while for a slope of 30°, the largest value considered, the variation of sill is maximum. Once this effect was known, we investigated again whether it was possible or not to find any robust relationship between forest and structural parameters. Cohen *et al.* (1990) stated that the sill of the variogram was related to forest cover. Figure 2b shows such relationship as a function of terrain aspect.

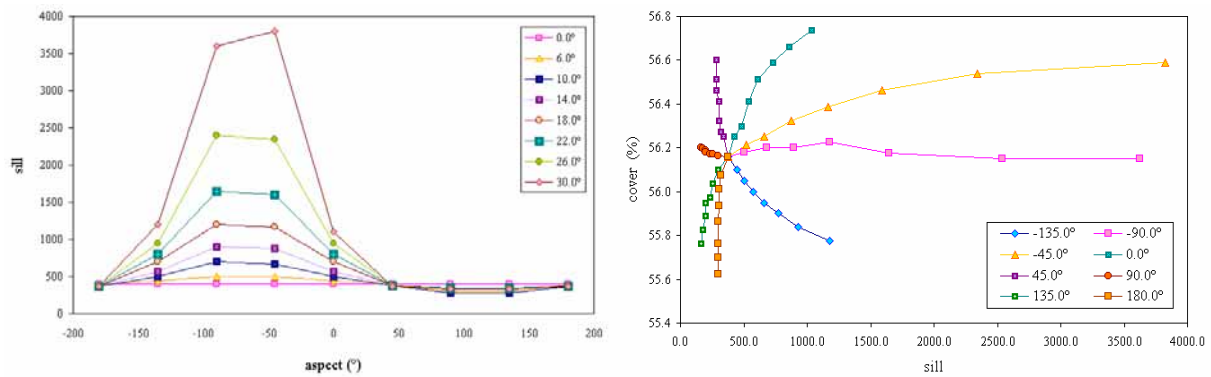


Figure 2. From the left to the right: (a) sill as a function of slope and aspect (simulated images), (b) forest cover (%) as a function of sill for varying aspect (simulated images).

6 CONCLUSIONS

Results reached so far show the difficulty of extracting reliable texture information from real aerial photographs. Similar results appeared when simulated images were analysed. The effect of slope and aspect on the image structure may be responsible of those results.

Forest structural and biophysical parameters influence the spatial structure of high resolution imagery (Bruniquel-Pinel and Gastellu-Etchegorry, 1998) but their influence may be totally masked by ground conditions such as slope and aspect. Therefore, the analysis of spatial structure in high resolution images may benefit from the removal of topographic effects. In fact, when considered constant slope and aspect conditions, an easy relationship was found between sill and canopy cover percentage. The results also show that the variogram is an effective tool with which to analyse this structure.

With the availability of high resolution remotely-sensed imagery and computer-based texture analysis tools, traditional methods of image analysis based on the spectral properties of the image are no longer a suitable approach to the identification of vegetation distribution patterns, and an alternative method is therefore required. The use of geostatistic methods to analyse the spatial domain of high resolution imagery may offer a suitable alternative. However, more analyses and experimentation are needed.

ACKNOWLEDGMENTS

The authors want to make acknowledgment to the Department of Geography of the University College of London for providing the simulation code for the generation of artificial images as well as the computation facilities. We are also grateful to the Banco de Datos de la Naturaleza (Ministry of Environment) for providing data and cartography from the Second National Forest Inventory and to the Consejería de Medio Ambiente (Comunidad de Madrid) for the topographic maps.

REFERENCES

- O'Dwyer, S., 2002. Scene modelling using geostatistics and local variance for forest parameter mapping. Transfer thesis, Department of Geography, University College of London.
- Franklin, S.E. and Wulder, M.A., 2003. Remote Sensing of Forest Environments, Conclusions. In: *Remote Sensing of Forest Environments*, Kluwer Academic Publishers, Massachusetts, USA.
- Cohen, W.B., Spies, T.A. and Bradshaw, G.A., 1990. Semivariograms of digital imagery for analysis of conifer canopy structure. *Remote Sensing of Environment*, 34: 167-178.
- Woodcock, C.E., Strahler, A.H. and Jupp, D.L.G., 1988. The use of variograms in remote sensing. I. Scene models and simulated images. *Remote Sensing of Environment*, 25: 323-348.
- Bruniquel-Pinel, V. and Gastellu-Etchegorry, J.P., 1998. Sensitivity of texture of high resolution images of forest to biophysical and acquisition parameters. *Remote Sensing of Environment*, 65: 61-85.
- Jupp, D.L.G., Strahler, A.H. and Woodcock, C.E., 1988. Autocorrelation and regularization in digital images. I. Basic theory. *IEEE Transactions on Geoscience and Remote Sensing*, 26(4): 463-473.
- Atkinson, P.M., Dunn, R. and Harrison, A.R., 1996. Measurement error in reflectance data and its implications for regularizing the variogram. *Int. J. Remote Sens.*, 17(18): 3735-3750.
- Lewis, P., 1996. Botanical Plant Modelling for Remote Sensing Simulation Studies, PhD thesis, University of London.
- Lewis, P., 2002. DRAT – The Advanced Radiometric Ray Tracer (<http://www.geog.ucl.ac.uk>).

- St-Onge, B.A. and Cavayas, F., 1997. Automated forest structure mapping from high resolution imagery based on directional semivariogram estimates. *Remote Sensing of Environment*, 61: 82-95.
- McBratney, A.B. and Webster, R., 1986. Choosing functions for semi-variograms of soil properties and fitting them to sampling estimates. *Journal of Soil Science*, 37: 617-639.
- Oliver, M., Webster, R. and Gerrard, J., 1989. Geostatistics in physical geography. Part I: theory. *Trans. Inst. Br. Geogr. N.S.*, 14: 259-269.
- Hudak, A.T. and Wessman, C.A., 1998. Textural analysis of historical aerial photography to characterize woody plant encroachment in South Africa savanna. *Remote Sensing of Environment*, 66: 317-330.

ASSESSMENT OF QUICKBIRD IMAGES TO MONITOR EARLY POST-FIRE VEGETATION DYNAMICS APPLYING SPECTRAL MIXTURE ANALYSIS

L.M Reithmaier, P. Barbosa, J. San-Miguel-Ayanz^a

^a European Commission, Directorate General Joint Research Centre,
Institute for Environment and Sustainability, Via E. Fermi, TP 261, 21020 Ispra (VA), Italy;
email: {lucia.reithmaier, paulo.barbosa, jesus.san-miguel}@jrc.it

ABSTRACT

The mapping of the spatial development of vegetation cover during the first years after forest fires implies to delineate precise structural parameters of re-growing vegetation by high to very high spatial resolution. This study tested the ability of spectral mixture analysis (SMA), applied to high spatial resolution Quickbird and SPOT data, to produce realistic and meaningful endmember fractions for the study of post-fire vegetation regeneration in the early stage. The analysis was performed on a large burned area in Anchuras, Spain. The affected pre-fire vegetation of the selected burned area was comprised of *maquis* shrublands and pine woodlands. Special emphasis was put on the pre-processing of the data to retrieve accurate vegetation reflectance values. A linear SMA was applied to all data sets using image derived endmembers, which were identified as green vegetation (GV), soil and shade. Spatio-temporal patterns of vegetation cover changes could be explained on the basis of shade-normalized GV fraction images. Vegetation cover and its change could be successfully detected with high accuracy through the use of the Quickbird and SPOT derived GV fraction images. The preliminary results obtained in this study show promising perspectives on the use of SMA for the analysis of vegetation re-growth after forest fires; however, in order to fully validate the results obtained in this study, it would be necessary to perform this analysis on longer time frame.

Keywords: Post-fire vegetation dynamics, Spectral Mixture Analysis, Quickbird, SPOT

1 INTRODUCTION

Monitoring dynamics of the vegetation after fire is essential to establish post-fire resource management and to evaluate the necessity of reforestation programs in order to reduce soil erosion (Riaño *et al.*, 2002). In general, at the stand level our understanding of the changes induced by fire in plant abundance is well documented for the Mediterranean region (Trabaud, 1987, Moreno and Vallejo, 1999). At a regional scale, remote sensing techniques have been increasingly used to evaluate the spatial pattern of vegetation recovery after wildfire (Riaño *et al.*, 2002, Peterson and Stow, 2003). During the first years after forest fires, the mapping of the spatial dynamic requires very high spatial resolution imagery such as Quickbird or SPOT to delineate precisely spatial patterns and to assess the swiftness of recovery.

The purpose of this study, which was part of the EC research project Fire Spread Prevention and Mitigation (SPREAD; *Contract number* EVG1-CT-2001-00043), was to study the ability of the Spectral Mixture Analysis (SMA) in order to gain information on vegetation re-growth in the early stage after the forest fire.

2 DATA AND METHODS

2.1 STUDY AREA AND DATA

The study was carried out in a large forest fire that took place on 1st of August 2002, in the surroundings of Anchuras, between Toledo and Ciudad Real provinces (Central Spain). The burned area (1400ha) is within a meso-mediterranean climate type. The affected area is very uneven, with four ravines and mean altitude of 500m. Biogeographically, the affected vegetation is comprised within the dehesas of *Pyro-Quercetum rotundifoliae*. The shrublands fall within *Genisto-Cistetum ladanifer* association. The fire affected several vegetation types, including crops, grasslands, shrublands, *Pinus pinaster* afforestations and several *Quercus* spp woodlands.

Two field plots (90m x 180m) were selected on two west-facing slopes, in the north of the burned area, for the comparison with remote sensing derived vegetation cover. These areas were occupied by a *Quercus*

suber open forest with shrubs mainly composed of *Cistus* and *Cytisus* species. Within these plots (named Valley A and Valley B) a stratified nested random sampling scheme was chosen to collect vegetation properties; in particular herbaceous, shrub, tree, and total vegetation cover data which were related to the remote sensing data. Unfortunately, this field sampling approach, which was also defined for other ecological studies, proved to be inappropriate for the comparison with the image data. Field sampling in relation with satellite data is usually more efficient if the study area is subdivided into regions or land use cover types; then several plots should be distributed in each selected land cover type across the study area; hence, choosing a stratified sampling approach with a reasonable number of plots would have been more likely to be sufficiently representing the area of interest (Justice and Townshend, 1981).

Cloud-free Quickbird and SPOT satellite data were chosen due to their high spatial resolution. The Quickbird images represent with its spatial resolution of 2.4m (multispectral) the very high spatial resolution image within this study. In contrast to its detailed spatial information, only four spectral bands in the visible and near-infrared wavelengths are available. The first image was acquired on 10.09.2003, more than one year after the fire, and the second image was acquired on the 22.09.2004. Additionally, a SPOT5 image with a spatial resolution of 10m and a spectral resolution of four bands within the visible-infrared region was used. The SPOT image was acquired one year after the forest fire (05.08.2003).

2.2 METHODS

2.2.1 Pre-Processing

Monitoring vegetation cover over time requires an adequate consistency between multi-temporal data sets and different sensors. Comparability becomes even more critical when referring to fractional cover estimates such as those derived from SMA. Several pre-processing steps were performed in order to assure comparability and to obtain land-surface reflectance.

Optimizing the geometric location accuracy, displacement errors, also caused by topographic relief, were reduced performing orthorectification of all acquired images. The images were further radiometrically calibrated using specific calibration coefficients of each sensor type. Since remote sensing in uneven terrain is limited by topographic effects on spectral signatures, a topographic normalization approach (Minnaert correction) was applied to reduce topographically induced illumination effects. As last step of the pre-processing chain, atmospheric correction was performed on the basis of the atmospheric modelling code 6S (Vermote *et al.*, 1997).

2.2.2 Spectral Mixture Analysis (SMA)

Numerous remote sensing techniques have been developed for change detection studies such as vegetation regeneration after forest fires. In several studies, the SMA has proved to be superior to widely used spectral indices (Elmore *et al.*, 2000, Riaño *et al.*, 2002); this is the case of quantitative analysis of vegetation cover in sparsely vegetated areas, in which confusion with soil background reflectance can be minimized through SMA. Therefore, SMA was chosen as the appropriate remote sensing technique for this project.

The idea of SMA is based on the theory of mixed pixels, which states that every pixel consists of a relative proportion of dominant features which have relatively constant spectral properties. In order to address this problem, the SMA extracts the abundances of the main contributors of reflectance (referred to as endmembers) at a sub-pixel scale. As a result abundance maps are calculated representing the fraction of the chosen endmembers (Adams *et al.*, 1993). Possible endmembers, however, are restricted to the number of the bands of the image data plus one (Hill, 1993). In case of Quickbird and SPOT, the number of endmembers is limited by the inherent dimensionality of the data to three endmembers in order to obtain realistic results.

As unmixing model for the SMA, linear spectral unmixing, which allows a simple interpretable connection to be made between the physical characteristics of the field data and the model was chosen. Here, the spectra are the linear summation of the spectrum of each endmember multiplied by the surface fraction they cover. As constraints of the model, the fraction of each endmember is restricted to a value between 0 and 1, and all fractions have to sum up to unity. This creates meaningful endmember fractions and a total cover that are physically realistic.

The crucial and most difficult part of the SMA is the actual performance of selecting the endmembers, because it is the influencing key factor of the SMA module. The definition of appropriate spectral endmembers can be either done using pure field-spectra from spectral libraries (*reference endmember*) or from the image itself (*image endmember*). As appropriate reference endmember were not available for the

study site, an approach to extract pure pixels from the image was applied to retrieve image endmembers. As commonly used in other studies, spectral scatterplots of image band combinations were built. The image endmembers were found at the vertices of the polygon that bounds the data space. These pixels of the data cloud were masked and then identified in the original reflectance image as soil, green vegetation (GV) and shade. For SPOT the common band combination NIR/red, where differences of vegetation and soil are mostly pronounced (Peterson and Stow, 2003), was applied. For Quickbird, additional procedure was necessary prior to the scatterplot to be able to bind the data space. The Principal Component Analysis (PCA) improved the ability to derive image endmembers, since over 99% of the spectral variability could be mapped in the first two components. Usually, the first two components were used to build the scatterplot in order to determine the image endmember. The GV endmember was mainly derived from the canopy of cork oak trees (*Quercus suber*) while the soil endmember was located on agricultural areas. The shade endmember which is supposed to account for illumination and albedo effects (Elmore and Mustard, 2003) was not taken from dark water surfaces but from ‘real’ shades (canopy shades) since, after the topographic correction, only unresolved canopy shadows should dominate the shade fraction.

Later, the model fitness or accuracy was assessed by the following assessment criteria: Fraction criterion, RMS criterion and Residual criterion. Overall, very fine results were obtained from all criteria. Additionally, shade normalization was applied to all endmembers in order to eliminate the effects of the shadow-casting canopies, strongly correlated to GV fraction in high resolution imagery. The normalized GV fraction is then representing the vegetation cover which is further compared with field data.

3 RESULTS

The normalized GV fraction images were investigated concerning differences between sensors and acquisition years in case of Quickbird. The example of a valley (Figure 1) shows similar vegetation patterns for all sensors in both years: scattered trees or shrubby vegetation on both slopes (dark shades) and smaller patches of high soil fraction (bright shades) right at the bottom of the valley. Since the spatial resolution of SPOT is only 10m, single trees or shrubs could not be resolved as in the case of the of Quickbird (2m resolution). Those areas where patches of single trees, soil and canopy shade could be identified in the Quickbird image are only defined as large homogenous areas with high GV fraction. Regarding the time difference between the two Quickbird images, a distinct increase of vegetation is visible on both slopes. This increase could be explained by a higher vegetation cover of re-sprouting individuals or shrubs, as the distribution of trees remained mainly unchanged.

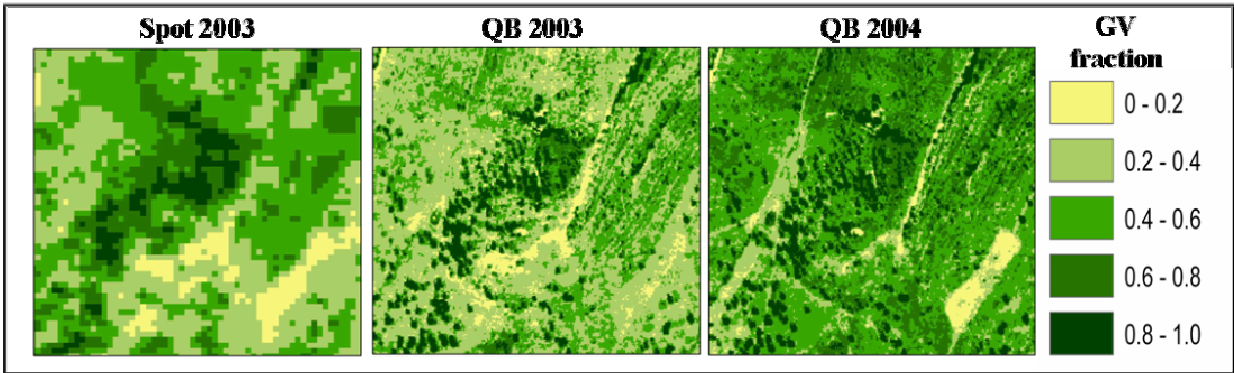


Figure 1. Example subsets of Green Vegetation (GV) fraction images of SPOT and Quickbird (QB). In green shades different fraction classes are illustrated.

Table 1. Average Green Vegetation fraction estimates for the entire burned area of Quickbird (QB) and SPOT (before and after shade normalization).

Date	Green vegetation (mean)	
	<i>no shade norm.</i>	<i>shade norm.</i>
QB_2003	14.0%	32.1%
QB_2004	14.6%	32.5%
SPOT_2003	18.1%	33.1%

In order to obtain physically meaningful vegetation cover estimates, the GV fraction was shade normalized. After the shade normalization the mean GV cover within the burned area was about doubled in all cases (Table 1). Comparing the burned area mean GV of Quickbird 2003 with the one of Quickbird 2004, the increase of vegetation was less than one percent in both cases, with and without shade normalization. Regarding the comparison of the two sensors, non normalized mean vegetation values of SPOT 2003 were already higher than the one of Quickbird 2004. However, after the normalization, the vegetation cover difference between SPOT 2003 and Quickbird 2004 estimates was significantly reduced, highlighting the importance of shade normalization, in particular of the Quickbird images.

The low mean increase of vegetation between the first two years after the fire could be explained by different human activities such as intensive pastures, road or other constructions. Figure 2 shows an example of vegetation change due to natural regeneration (positive change illustrated in green colours) or human activities (negative change in red colours). In this subset a general slight re-growth can be noted throughout the burned patches; negative trend of vegetation change is caused by a construction site (south-east corner) and roads, as well as pasturing, which result in grassy vegetation (northern part).

In general, from 2003 to 2004 distinct spatial patterns with similar regenerative capability or decline of vegetation regeneration can be noted throughout the entire burned area. The comparison with pre-fire land cover information revealed strong correlation between the re-growth swiftness and pre-fire vegetation type. A substantial re-growth of about 20 to 30% can be noted in areas of former dense shrublands while the regeneration of pine woodlands showed more than 20% decrease of vegetation cover. Negative trends can be usually noted in the large area of dehesas and pastures that could be linked to land management practices. Here, the vegetation change is dominantly negative between 10 and 30% decrease. Hence, re-growth and decline of vegetation are near-balanced, resulting in the aforementioned slight mean increase of vegetation regarding the entire burned area.

Comparison of the GV fraction derived from the satellite imagery with field properties was attempted. However, several factors made the correlation between field and remote sensing data rather difficult; among others, the inappropriate field sampling design and the time lag between the acquisition time and the field sampling lead to only low correlation values.

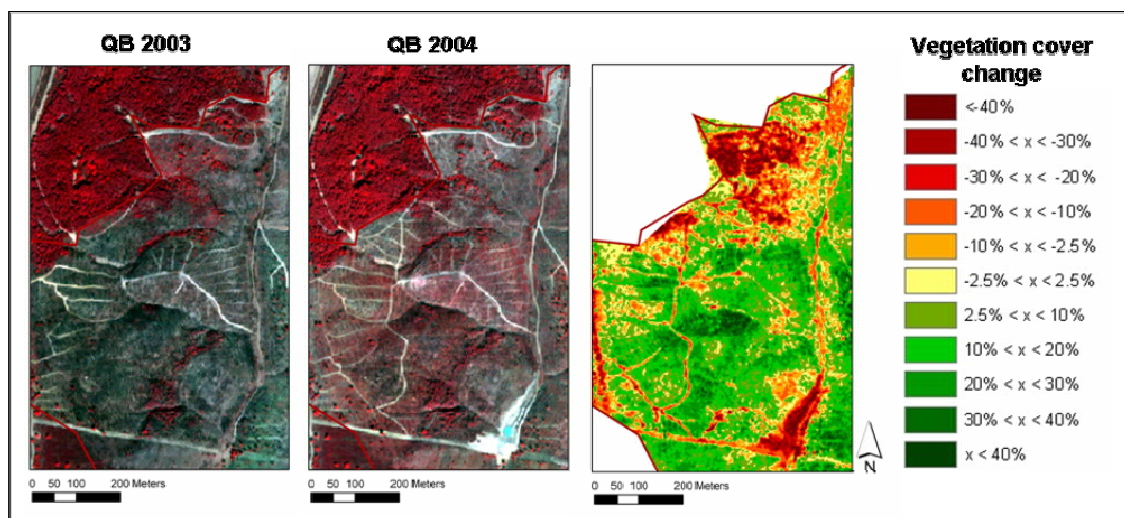


Figure 2. Vegetation cover change between the first two years after the fire, illustrated on two false colour subsets of Quickbird (QB) of both years and on the change map of vegetation cover derived by SMA.

4 CONCLUSIONS

In this study, the use of SMA as tool for detecting vegetation regeneration proved its strength, which lies in the fact that it explicitly refers to physical parameters. Even with the limited spectral resolution (four bands), the pre-processed Quickbird images were able to successfully detect vegetation with high accuracy using the image endmember approach. Due to its high spatial resolution even single trees could be identified.

The vegetation cover derived from SPOT using SMA reached high accuracy; these results agree with Souza's *et al.* (2003), where derived fractions from SPOT4 proved to be useful for regeneration studies. Since its spatial resolution is limited to 10 meters, the overall vegetation cover is slightly overestimated compared to Quickbird estimates. As stated before, the resolution of Quickbird could resolve trees and their canopy shade. SPOT, however, only delineates patches of several trees with high vegetation cover. This slight overestimation is then further increased after the shade normalization. Therefore, it can be stated that after a thorough topographic normalization, no shade normalization is advisable in case of SPOT image.

The mapping of regeneration dynamic within the first two years after the forest fire was obviously correlated to the pre-fire land cover type. This confirms the observation of the regeneration of resprouting shrubs which are adapted to burning and therefore contribute to a relative quick increase in vegetation cover (Trabaud, 1987). In pine woodlands, on the contrary, an average decline of vegetation cover of more than 20% was detected between the first and the second year after the fire. Also, the dying back of single cork trees (*Quercus suber*) could be detected by Quickbird. Some cork trees were decorked before the fire, but they survived the first year. However, without their thick corky bark they lost their passive resistance (Trabaud, 1987) and died within the second year. Statements about the regeneration within dehesas, pastures or sparse shrublands can not be made, since they are highly influenced by land use practices.

The preliminary results obtained in this study show promising perspectives on the use of SMA for the analysis of vegetation re-growth after forest fires; however, in order to assess the recovery of the original vegetation before forest fires, it would be necessary to perform this analysis on a longer time frame. Moreover, the sampling design of the field measurements of future studies should take into consideration the applicability for satellite imagery, in order to be able to fully validate the results obtain in this study.

REFERENCES

- Riaño, D., Chuvieco, E., Ustin, S., Zomer, R., Dennison, P., Roberts, D. and Salas, J. 2002. Assessment of vegetation regeneration after fire through multitemporal analysis of AVIRIS images in the Santo Monica Mountains. *Remote Sens Environ.* 79: 60-71.
- Trabaud, L. 1987. Fire and survival traits of plants. In: *The role of fire in ecological systems*. L. Trabaud (Ed.), SPB Academic Publishing, pp 65-89.
- Moreno, J. M. and Vallejo, R. 1999. Fire impacts on the ecosystem and restoration: summary of the main findings from the Delfi-fire database. In: *Proceedings of the International Symposium on Forest Fires*, CINAR, S.A.
- Peterson, S.H. and Stow, D.A. 2003. Using multiple image endmember spectral mixture analysis to study chaparral re-growth in southern California. *Int. J. Remote Sens* 24: 4481-4505.
- Justice, C.O. and Townshend, J.R.G. 1981. Integration ground data with remote sensing. In: *Terrain analysis and remote sensing*. J.R.G. Townshend, (Ed.), London.
- Vermote E., Tanré, D., Deuzé, J.L., Herman, M., and Morcrette, J.J. 1997. Second Simulation of the Satellite Signal in the Solar Spectrum (6S): an overview. *IEEE Trans. Geosci. Remote Sensing* 35: 675-686.
- Elmore, A.J. and Mustard, J.F. 2003. Precision and accuracy of EO-1 Advanced Land Imager (ALI) data for semiarid vegetation studies. *IEEE Trans. Geosci. Remote Sensing* 41: 1311-1320.
- Adams, J.B., Smith, M.O., and Gillespie, A.R. 1993. Imaging Spectroscopy: Interpretation based on spectral mixture analysis. In: C.M. Pieters and P.A.J. Englert: *Remote geochemical analysis: Elemental and mineralogical composition*. Press Syndicate of University of Cambridge., pp. 145-166.
- Hill, J. 1993. High precision land cover mapping and inventory with multi-temporal earth observation data. European Commission, EUR 15271.
- Souza Jr., C., Firestone, L., Silva, L.M. and Roberts, D. 2003. Mapping forest degradation in the Eastern Amazon from SPOT4 through spectral mixture models. *Remote Sens Environ.* 87: 494 – 506.

THE PERFORMANCE OF DIFFERENT LEAF MASS AND CROWN DIAMETER MODELS IN FORMING THE INPUT OF A FOREST REFLECTANCE MODEL: A TEST ON FOREST GROWTH SAMPLEPLOTS AND LANDSAT ETM IMAGES

M. Lang^{ab}, T. Nilson^{ab}, A. Kuusk^a, A. Kiviste^b, M. Hordo^b

^a Tartu Observatory, 61602 Tõravere, Estonia, e-mail: lang@aai.ee

^b Estonian Agricultural University, Kreutzwaldi 5, 51014 Tartu, Estonia

ABSTRACT

Five different foliage mass models and six crown diameter models were tested on the preparation of the reflectance model input for 246 forest growth sampleplots in Estonia. Spectral database of forest understorey vegetation was used to characterize the forest ground cover. Three atmospherically corrected Landsat Enhanced Thematic Mapper (ETM) scenes were used to compare with the simulation. Best correlation between the simulated and measured reflectance factors was found in the shortwave infrared band (ETM5). In the near infrared band (ETM4) predicted reflectance factors were systematically lower for brighter stands. There was virtually no correlation in the visible region. The causes of the undescribed variability are related to the characterization of forest understorey vegetation, foliage optical properties, and the lack of adequate foliage mass and tree crown diameter models for Estonian conditions.

Keywords: reflectance modeling, tree foliage mass, tree crown dimensions

1 INTRODUCTION

Many variables from regular forest inventory data are directly utilized for running the reflectance model (FRM) of Kuusk and Nilson (2000). On one hand, information about age (A), stand relative basal area (T), tree species composition, site type, number of trees per unit area (N), tree height (H), stem diameter at breast height ($D_{1.3}$) usually exists in stand description records, while these parameters are traditionally used in forest management planning. On another hand, forest stand reflectance depends in addition on other variables that are not used in everyday forest management practice (Nilson et al., 1999; Nilson et al., 2003). One must have estimates for total green leaf mass per tree, tree crown dimensions and forest understorey vegetation characteristics to run the forest reflectance model of Kuusk and Nilson (2000). The link between the variables in interest and forest inventory data can be established using regression models (Widlowski et al., 2003).

The studies about relationships between tree stem and crown characteristics have been driven by the development of the aerial photography based forest inventory (Dmitriev, 1981; Howard 1991). In a simplest approach linear dependence between crown diameter (D_{cr}) and $D_{1.3}$ is used (Jakobsons, 1970). To account for the influence of stand density, H and $D_{1.3}$ ratio (Widlowski et al., 2003) or average distance between trees (L) is included into the model (Kuz'mitšev, 1977).

Foliage mass models are usually based on the $D_{1.3}$, H and live crown length (L_{lc}) (Widlowski, 2003; Marklund, 1988; Vares, 2004). Difficult to measure, but a descriptive variable is the stem diameter at live crown base (D_{lcb}) (Hoffmann and Usoltsev, 2002).

Forest canopy cover, LAI and soil properties determine the light regime and growth conditions for the plants growing beneath the trees (Kull et al., 1995). The two-layer homogeneous canopy reflectance model of Kuusk (2001) considers many structural and spectral properties of the ground vegetation. The drawback of the flexibility is the need for many parameters. To overcome this problem Kuusk et al. (2004) propose the inversion of the model on the measured ground vegetation spectra.

We focused our study on two aspects: 1) performance of the regression functions that predict tree crown diameter and leaf mass per tree and 2) influence of the forest understorey vegetation properties on the stand reflectance.

2 MATERIAL AND METHODS

2.1 THE IMAGES AND SAMPLEPLOTS

We used three atmospherically corrected Landsat Enhanced Thematic Mapper (ETM) images from World Reference System (WRS) path 186 row 19 (Figure 1). Images are dated to 10.07.1999, 10.06.2000 and 31.05.2002. Image from summer 1999 was cloudfree, on the image from year 2000 north-east corner of Estonia was covered with cumulus clouds and on the image from year 2002 were some cumulus clouds.



Figure 1. Study area, image frame of the Landsat ETM+ scene 186-19 and locations of the sampleplot clusters

We used the data from forest growth sampleplots (Kiviste and Hordo, 2002). Plots are circular with a radius of 10 ... 30 m depending on the forest age. Map coordinates of the plots are determined by GPS, from digital maps or orthophotos. All trees having $D_{1.3}$ larger than 2 cm are tallied, distance and azimuth from the plot centre is measured. For every fifth tree the height and height to live crown base (H_{lcb}) is measured.

Out of 381 sampleplots within satellite image frame, 246 were suitable for our reflectance simulation study. We excluded some plots because of clouds or cloud shadows, recent management events and too large variability of the reflectance. Plot reflectance was calculated using the area of pixels overlapping with the plot as weight. Plot reflectance variability was estimated from 100 random samples 10 meters around the measured location of the centre. The plots having sampling error larger than 10% in ETM bands 3, 4 and 5 were left out.

According to species distribution we had 45 spruce dominated plots, 88 pine dominated plots, 89 deciduous dominated plots and 24 mixed plots. Stand age varied from 15 to 146 years, and density varied from 0.023 to 0.494 trees per m^2 in the upper layer. In 64 plots the second storey tree layer was denser than 0.05 trees per m^2 . Nemoral, meso-euthrophic and mesotrophic site types dominated in the sample.

2.2 PREPARATION OF THE REFLECTANCE MODEL INPUT

A detailed description of the input data structure for the FRM can be found from its user guide (Kuusk and Nilson, 2002). We run the FRM using different models for foliage mass (Marklund, 1988; Cermák et al., 1998; Gower et al., 1993; Hoffmann and Usoltsev, 2002; Vares et al., 2004) and crown diameter (Jakobsons, 1970; Kuz'mitčev, 1977; Nilson et al, 1999; Lang, 2000; Pretzsch et al., 2002; Nagel et al., 2004;). Marklund's (1988) regression functions for deciduous trees estimate the total mass of living shoots. Foliage mass was assumed to be $1/[1+0.5\exp(0.25\ln(D_{1.3}^2)H)]$ of the estimated shoot mass. For Hoffmann and Usoltsev (2002) D_{lcb} was calculated by a function of Ozolins (2002) which was adapted for Estonia by Padari (2004).

The used tree species specific values for optical and structural constants are given in Table 3. The tree distribution parameter c was calculated for each crown diameter model from resultant crown closure (CRCL) and canopy closure (CACL) as $-\ln(1-CACL)/CRCL$.

To characterize forest understorey vegetation for FRM, we inverted the two-layer model of Kuusk (2004) on the site type specific spectra from the database (Lang et al.2002). For each test we generated two upper understorey layer (LAI_u) scenarios: 1) constant LAI_u and 2) $LAI_u = 1.6092 - 0.2064 * LAI_{trees}$.

Table 1. Tree species specific values for other FRM input variables

FRM input variable	<i>Species</i>		
	Scots pine	Norway spruce	Deciduous
Crown form	ellipse	cylinder+cone	ellipse
Leaf weight per area (SLW, g m ⁻²)	168	254 (Sellin, 2000)	Niinemets and Kull (1994)
Clumping index of needles or leaves	0.60	0.56	1.0
Leaf chlorophyll content (% of SLW)	0.24	0.37	0.65
Leaf water content (% of SLW)	100	100	120
Brown pigments (% of SLW)	0.2	0.2	0.2
Lignin+cellulose+sugars etc. (% of SLW)	99.6	99.7	99.4
Leaf structure parameter, n	3/SLW	3/SLW	2/SLW
Shoot length	0.15	0.2	0.40
Branch area index	0.15*LAI	0.12*LAI	0.1*LAI

3 RESULTS AND DISCUSSION

A comparison of simulated and measured reflectance factors for summer 1999 is shown in Figure 1 for the green (ETM 2), red (ETM 3), near-infrared (ETM4) and shortwave-infrared (ETM5) spectral bands when understorey vegetation leaf area index (LAI_u) is a site type specific constant, crown diameter is predicted according to Jakobsons (1970) and foliage mass is estimated using $D_{1.3}$, H and L_{lc} (Marklund, 1988).

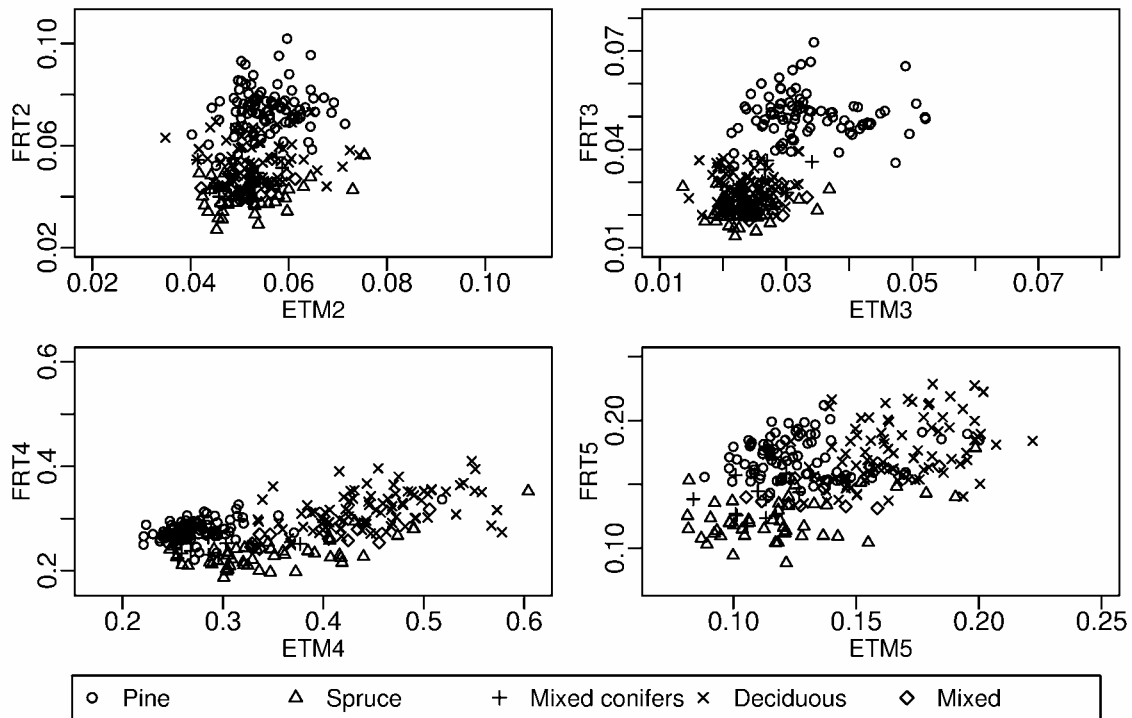


Figure 2. Simulated and measured (Landsat ETM, 10. July 1999) reflectance in case of constant LAI_u , $D_{1.3}$ based crown diameter model of Jakobsons (1970) and leaf mass model of Marklund (1988) based on $D_{1.3}$, H and L_{lc}

The best agreement between the simulated and measured reflectance coefficients was in band ETM 5 in all combinations of models. There was virtually no relationship in the red and green spectral bands irrespective of the crown diameter- or foliage mass model used. This is rather surprising and has probably caused by the wrong estimate of chlorophyll content of the foliage. In addition to the variability in leaf biochemistry a pure technical problem with the red channel is the low dynamic range of signal over the forests. In the near infrared spectral band (ETM4) the simulated reflectance coefficients for brighter stands were systematically lower than measured reflectance factors.

Using the second understorey LAI scenario, i.e. when the understorey LAI is related to the LAI of tree-layer, made the stands slightly brighter in all spectral bands except for the band ETM4. In ETM4 the effect

of the overstorey-understorey LAI relation was opposite – the stand reflectance decreased. This effect is partially caused by the fact that soil and debris are usually brighter than vegetation in all ETM bands except band 4. Decrease in abundance of forest understorey green vegetation decreases multiple scattering in the near infrared spectral region and resulting in smaller stand reflectance.

When comparing simulated reflectances against measured for other two Landsat ETM images, which are from the beginning of the summer, the results were in general similar to those presented in Figure 1. We did not adapt the leaf biochemistry nor leaf mass or SLW to simulate the effects of differences in vegetation period because lack of suitable models and data.

One issue when using the regression models, is the inclusion of predictive variables. Usually, when developing the model, more variables give a better fit. When using these functions on the material where correlation between the parameters is different, we can get systematic differences. For example, using the models of Marklund (1988) and only $D_{1,3}$ as predictive variable we get systematically bigger leaf mass estimates than with model that includes also tree height. By including the live crown length, the estimates are smaller than using the model based on the $D_{1,3}$ and H . This makes it rather difficult to assess whether one model is better than other. When averaged over all plots the resultant LAI derived from $D_{1,3}$ based model was 5.4, using $D_{1,3}$ and H LAI decreased to 4.6 and by including L_{lc} LAI dropped to 3.8. Predicted average LAI ranged for deciduous stands from 4.3 to 6.4, for pine dominated stands from 2.6 to 4.6 and for spruce dominated stands from 4.1 to 8.8.

The differences in estimated crown diameters depending on the model were large. The resultant canopy closure varied 30...50% depending on model. Prediction of the tree crown diameter based on the inventory data remains problematic because of significant influence of the stand density and even species composition (Dmitriev et al., 1981). Laser scanning of forests is probably the way to collect large amounts of data for the development of tree crown and canopy structure models.

Beside issues related to leaf biochemistry, canopy closure and foliage mass, a set of problems has to be solved to characterize the forest understorey vegetation for the FRM. The rough information about the forest ground cover is usually in the databases as site type code. The total leaf mass per unit area, leaf properties like SLW, species composition of the moss, lichen, grass and shrub and hence the spectral properties of the tree layer background vary in a great deal inside one site type. Good fit between the simulated and measured reflectances in earlier tests (Nilson et al., 1999) can partially be explained by the facts that the ground reflectances of studied forest stands were explicitly measured using a field radiometer and crown closure of the stands was measured.

4 CONCLUSIONS

None of the tested foliage mass and tree crown diameter model combinations was found to be significantly better than others. The problems of the regression models are basically their relation to the material that was used for parameterization. Besides the LAI and crown closure the variability caused by the forest understorey vegetation can be significant. Unless we are able to predict the forest understorey vegetation properties depending on the site productivity and tree layer LAI or canopy cover, it is not possible to get better correlation between the simulated and measured forest stand reflectances.

ACKNOWLEDGMENTS

This work has been supported by the Estonian Science Foundation grants nos. 4696 and 5768, State forest management centre (RMK) and Environmental Investment Centre (EIC).

REFERENCES

- Kuusk, A., Nilson, T. 2000. A directional multispectral forest reflectance model. *Remote Sensing of Environment* 72(2), 244-252.
- Nilson, T., Kuusk, A., Lang M., Lökk, T. 1999. Forest reflectance model as an interface between satellite images and forestry database. In: *IUFRO conference on remote sensing and monitoring. Held in Rogow, Poland, June 1-3.* Office of Official Publications of the European Communities, Luxembourg.
- Nilson, T., Kuusk, A., Lang, M., and Lökk, T. 2003. Forest reflectance modeling: Theoretical aspects and applications. *Ambio* 32(8), 535-541.
- Widlowski, J-L., Verstraete, M., Pinty, B., Gobron, N. 2003. Allometric relationships of selected European tree species. EC Joint Research Centre, Italy.

- Dmitriev, I., D., Murahtanov, E., S., Sukhikh, V., I. 1981. *Lesnaja aërofotos'emka i aviacija*. Lesnaja Promyšlennost', Moskva, pp. 150-162 (In Russian).
- Howard, J., A. 1991. *Remote sensing of forest resources*. Chapman & Hall, London, pp. 296-307.
- Jakobsons, A. 1970. The correlation between the diameter of tree crown and other tree factors – mainly the breastheight diameter. *Research notes 14*. Department of Forest Survey, Royal College of Forestry, Stockholm.
- Kuz'mitšev., V., V. 1977. Zakonomernosti rosta drevestoev. Nauka, Novosibirsk, pp. 115-129 (In Russian).
- Marklund, L., G. 1988. Biomass functions for pine, spruce and birch in Sweden. *Rapporter Skog 45*. Swedish University of Agricultural Sciences, Umeå.
- Vares, A., Lõhmus, K., Truu, M., Truu, J., Tullus, H., Kanal, A. 2004. Productivity of Black alder (*Alnus glutinosa* (L.) Gaertn.) plantations on reclaimed oil-shale mining detritus and mineral soils in relation to rhizosphere conditions. *Oil Shale* 21(1), 47-62.
- Hoffmann, C. W., Usoltsev, V. A. 2002. Tree-crown biomass estimation in the forest species of the Ural and of Kazakhstan. *Forest Ecology and Management* 158, 59-69.
- Kull, O., Aan, A., Sõelsepp T. 1995. Light interception, nitrogen and leaf mass distribution in a multilayer plant canopy. *Functional Ecology* 9, 589-595.
- Kuusk, A. 2001. A two-layer canopy reflectance model. *Journal of Quantitative Spectroscopy & Radiative Transfer* 71, 1-9.
- Kuusk, A., Lang, M., and Nilson, T. 2004. Simulation of the reflectance of ground vegetation in sub- boreal forests. *Agricultural and Forest Meteorology* 126(1-2), 33-46.
- Kiviste, A., Hordo, M. 2002. Network of permanent forest growth plots in Estonia. *Forestry studies XXXVII*, 43-58 (In Estonian with English summary).
- Kuusk, A., Nilson, T. 2002. Forest Reflectance and Transmittance FRT User Guide version 09. Tartu Observatory. Manuscript.
- Cermák, J., Riguzzi, F., Ceulemans, R., 1998. Scaling up from individual tree to the stand level in Scots pine. I., Needle distribution, overall crown and root geometry. *Annales des sciences Forestières* 55, 63-88.
- Gower, T. S., Reich, P. B., Son, Y. 1993. Canopy dynamics and aboveground production of five tree species with different leaf longevities. *Tree physiology* 12, 327-345.
- Lang, M. 2001. Satelliidipildi ja metsandusliku andmebaasi sidumise metsa heleduse mudeli abil. *MSc thesis*, Estonian Agricultural University, Tartu. Manuscript (In Estonian).
- Pretzsch, H., Biber, P., and Durský, J. 2002. The single tree-based stand simulator SILVA: construction, application and evaluation. *Forest Ecology and Management* 162, 3-21.
- Nagel, J., Albert, M., Schmidt, M. 2002. Das waldbauliche Prognose und Entscheidungsmodell BWINPro 6.1. *Forst und Holz* 27, 486-493.
- Ozolins, R. 2002. Forest stand assortment structure analysis using mathematical modelling. *Forestry studies XXXVII* 33-42.
- Padari, A. 2004. Metsahindamisprogramm "RaieWin". *MSc thesis*, Estonian Agricultural University, Tartu. Manuscript (In Estonian).
- Lang, M., Kuusk, A., Nilson, T., Lökk, T., Pehk, M., Alm, G. 2002. Reflectance spectra of ground vegetation in sub-boreal forests. Web page: available on-line <http://www.aai.ee/bgf/ger2600/> from Tartu Observatory, Estonia.
- Sellin, A., 2000. Estimating the needle area from geometric measurements: application of different calculation methods to Norway spruce. *Trees* 14, 215-222.
- Niinemets, Ü., Kull K. 1994. Leaf weight per area and leaf size of 85 Estonian woody species in relation to shade tolerance and light availability. *Forest Ecology and Management* 70, 1-10.

Session 4b

TOP HEIGHT ESTIMATION USING COMMERCIAL AIRBORNE REMOTE SENSING PRODUCTS IN BRITISH FORESTRY

J. C. Suárez^{1*} and E. D. Wallington²

¹Forest Research, Northern Research Station, Roslin, Midlothian, EH25 9SY, UK.

²Edinburgh Earth Observatory, School of GeoSciences, The University of Edinburgh, Drummond Street, Edinburgh, EH8 9XP, UK.

*Corresponding author: e-mail: juan.suarez@forestry.gov.uk, Tel: + (44) (0)131 445 6949

ABSTRACT

Top height is an important parameter in production forestry and ecological studies. Top height can also be used to describe forest structure, bio-fuel estimations, and the amount of carbon sequestration. The work presented in this paper is a comparative study based on the use of commercial airborne LiDAR and Radar data. The aim of this work is the development of a cost-effective method for the estimation of top height to enhance traditional field methods in operational forest inventories and related applications. The results are compared to ground measurements in a group of four study areas in Northern England, SW and Northern Scotland and North Wales.

Comparison between these two airborne data capture techniques and field observations showed retrieved height underestimates of between 20-30% for Radar and <10% for LiDAR. The R^2 values were about 0.81 for Radar and 0.92 or 0.98 for LiDAR. The discrepancies are mainly attributed to slope angles, stocking density, tree height and edge effects inside and at the periphery of stands. Results are focused primarily on Sitka spruce (*Picea sitchensis*), but other species are discussed within the height retrieval context. A brief discussion of the potential to differentiate heights at the sub-stand level is also introduced. The ability to differentiate heights will have beneficial consequences for more detailed management practices like the transition to continuous cover forestry, monitoring health or stocking density; in this case by assessing height variation across a stand.

Keywords: SAR interferometry, LiDAR, top height retrieval, operational forestry.

1 INTRODUCTION

Top height is an important parameter in production forestry that can also be used in studies looking at the forest stand structure, bio-fuel estimations or carbon sequestration. The work presented in this paper is a comparative study based on the use of commercial airborne LiDAR and Radar data. The study aims at the development of a cost-effective method for the estimation of top height in forest stands. The method seeks to enhance traditional field routines for forest inventory and to explore operative issues for its implementation in the industry. The results are evaluated by comparing to the information collected on the ground in a group of study areas in Great Britain.

The perceived advantages of these two airborne systems are the high resolution of the data and their versatility to capture data at will. Additional advantages are the increasingly competitive costs and the possibility to process the data into a product compatible with off-the-shelf GIS systems. Intermap Technologies have recently used their Star-3i airborne X-band SAR Interferometer to capture elevation data for the entire UK, under their Nextmap Britain campaign (Intermap Technologies Inc, 400 Inverness Parkway, Suite 330, Englewood, CO USA). Such data are rapidly being acquired for a large number of countries, and so, these datasets are becoming available to industrial users. The UK Environment Agency acquired recently the capability of collecting LiDAR data for the production of high resolution digital terrain models. This government agency also works as a contractor for other agencies throughout the UK.

2 STUDY SITES AND DATA

2.1 STUDY SITES

Four study sites were considered: Coed Y Brenin Forest District in North Wales (N52:49:12 W3:53:27 lat/long), Kielder Forest District in Northumberland, North England (N55:11:44 W2:32:11 lat/long), Glen

Affric in the Scottish Highlands (N57:17:00 W4:54:49 lat/long) and Aberfoyle in SW of Scotland (N56:10:00 W:4:22:00 lat/long). Forest stands consisted of a number of species including Sitka spruce (*Picea sitchensis*), Norway spruce (*Picea abies*), Japanese larch (*Larix kaempferi*), Western hemlock (*Tsuga heterophylla*) and Scots pine (*Pinus sylvestris*).

2.2 DATA SETS

2.2.1 InSAR digital surface model

The InSAR (Henderson and Lewis, 1998) Digital Surface Model (DSM) was supplied by Intermap Technologies (Intermap, 2005), and was geo-referenced to OSGB36 prior to delivery. The DSM is produced from the first return of the signal. It represents the first surface the signal came into contact with, whether ground or vegetation canopy. The DSM is presented in a pixel size of 5x5 m and a RMSE of between 0.5 – 1.0m (95% = 1.0 - 2.0m), dependent on flying height. These accuracies are quoted for moderately sloped, unobstructed terrain (Intermap, 2003).

2.2.2 Ordnance Survey DEM

The Ordnance Survey (OS) Profile 10m Digital Elevation Model (OSDEM) was used as a ground reference surface. The OSDEM has a pixel size of 10x10 m (re-sampled to 5x5 m to match the DSM; this was done purely to ease comparison and not as an attempt to improve resolution). This DEM is assumed to be a representation of the true ground surface for the purposes of this study.

2.2.3 LiDAR

Airborne LiDAR was obtained from the Environment Agency using an Optech ALTM2033 scanner (Optech Incorporated, 100 Wildcat Road Toronto, Ontario, Canada). The first survey was undertaken in Aberfoyle in September 2002 at a flying altitude of 1,000 m a.s.l. The sampling intensity was 3-4 returns per m² and a beam divergence of 10 cm. The scanning angle was 20 degrees. A second area in Kielder forest was surveyed in April 2003. Sampling intensity varied from 6 to 23 returns per m². Beam divergence was up to 1 m and the scanning angle was 10 degrees. Data were distributed in ASCII XYZ format, where first and last returns, with their corresponding intensities, were located to the OS national grid. RMSE were ± 40 cm in X and y and $\pm 9 - 15$ cm in Z.

2.2.4 Study sites and ground reference data

The SAR data were analysed over three study sites: Kielder (N. England), Coed Y Brenin (Wales) and Glen Affric (Scottish Highlands), whilst the LiDAR were analysed over the same Kielder sites as well as Aberfoyle (SW Scotland). Standard forest inventory techniques (Philip, 1994; Husch *et al.*, 2003; Hamilton, 1998) were used to establish the top height of the sample stands. This height was then used as the true top height of the stand.

3 HEIGHT RETRIEVAL METHODOLOGY

The technique for top height retrieval used in this study was similar for both the InSAR and LiDAR data sets, but a number of different steps were taken in each. The generic height retrieval algorithm was:

$\text{Height}_{\text{Tree}} = \text{Height}_{\text{Canopy}} - \text{Height}_{\text{Ground}}$	(1)
---	-----

3.1 INSAR

In all the stands, height values from the DSM and OSDEM were retrieved from 50 x 50m plots. Subtraction of the OSDEM from the DSM was performed to recover the height per pixel within the plot (Wallington *et al.*, 2004). Retrieved top height was estimated by averaging the highest 25 pixel heights. This technique follows standard forest practice of taking the tallest 100 trees/ha (Philip, 1994). This top height was then compared to the measured top height per plot.

3.2 LiDAR

LiDAR instruments can generate accurate canopy height models that can be used to estimate other forest parameters such as canopy heights, stand volume, and the vertical structure of the forest canopy. A normalised forest canopy model is obtained by subtracting bare ground values from the canopy layer. In commercial airborne systems, the canopy layer is retrieved from the first laser return that measures the intensity of the signal as it first encounters an object in the ground. The last return will provide

information about the location and height of the mid-point of the last strong waveform that is normally associated with the terrain.

In order to approximate a model of the ground surface, the last returns were filtered to eliminate those hits being intercepted by the forest canopy and, therefore, not reaching the underlying terrain. The method involved an iterative process of selection of points within kernels of variable size according to the local minima (Suárez et al., 2005).

A normalised canopy height model was calculated for every laser hit as the difference between the first return and the nearest return classified as terrain. Mean distances between canopy and ground returns varied between 0.5 and 1.5 m, with standard deviations less than 1 m. Individual tree heights were accurately predicted in 73% of the cases within ± 1 m. The largest underpredictions were observed in sub-dominants. The suppressed and dead trees were missed completely. Generally, individual tree heights were 7-8% shorter than observed due to the low number of laser hits intercepted by the apices. In Kielder, the higher density of returns per m^2 (6 to 23) reduced underestimations to less than 4%.

4 TOP HEIGHT RETRIEVAL RESULTS

4.1 INSAR TOP HEIGHT RETRIEVAL

Top height predictions across all the test sites (and a number of species) produced a strong Pearson's correlation of 0.79 (Figure 1, left). As expected, the retrieved height was systematically underestimated due to signal penetration through the canopy and the shadowing effects generated from the scanning angles. At the shorter wavelengths (X-band), penetration is limited to the upper canopy, and as such the resultant height of the scattering phase centre is predominately dominated by scattering from the smaller scatterers in the canopy. The resulting average underestimation is 34.7%. Regression analysis gave a R^2 value of 0.62 (Figure 1, right).

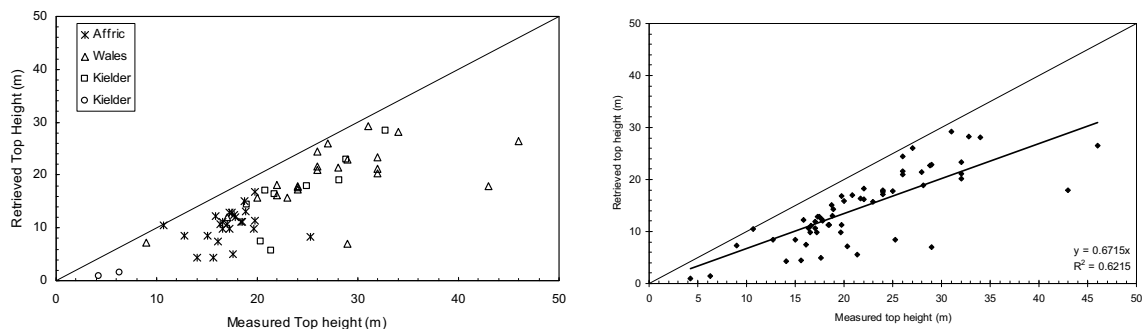


Figure 1. Left: Measured top height against retrieved top height for a range of species and sites. Average error of 7.23 m (34.7%), observations = 59, Pearson's $R = 0.79$. Note: Kielder (○) data were collected independently of Kielder (□). **Right:** Regression analysis of measured top height plotted against retrieved top height, $R^2 = 0.62$.

To allow comparison with the LiDAR data and to remove any effects of species, Sitka spruce (*Picea sitchensis*) stands were investigated in Coed Y Brenin (Wales) and Kielder. Underestimation reduced to between 21% in Kielder and 24% in Coed Y Brenin, with R^2 values of 0.94 and 0.84 respectively (Figure 2). The results showed that X-Band does not easily penetrate the forest canopy in Sitka spruce stands due to the compactness of their canopies. In contrast, other species such as Japanese larch and Scots pine present less dense canopies and produce some permeability of the signal through them. This would explain the increase in the underpredictions compared to spruce and hemlock stands.

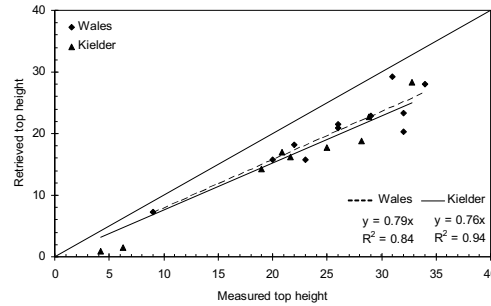


Figure 2. Retrieved top height for Sitka spruce (*Picea sitchensis*) stands in Kielder and Coed Y Brenin.

4.2 LiDAR TOP HEIGHT RETRIEVAL

Top height retrieval followed different methods in the two monitoring areas. In Aberfoyle, individual tree tops were initially estimated following an Object-Oriented classification method, as described in Suárez et al. (2005). The method aimed at the identification of individual tree tops using a rule based system implemented in eCognition (Definiens Imaging GmbH, 2001; Trappenstreustrasse 1, 80339 Muenchen, Germany). The classification combined digital aerial photography and an interpolated image of the canopy height model. After this, top height estimates were calculated as a mean of the 25 higher trees in each of the $\frac{1}{4}$ ha plot (Figure 3). Results produced a general underestimation of 7%; in line with the predictions of individual tree heights, and R^2 of 0.91.

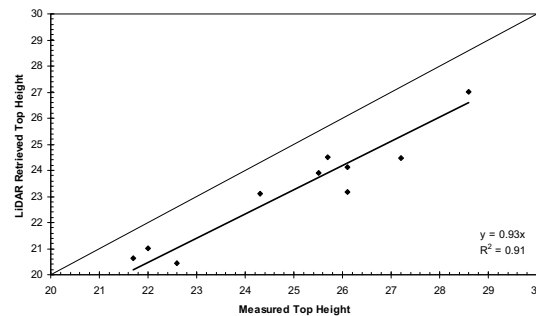


Figure 3. LiDAR Retrieved top height against measured top height in Aberfoyle.

In Kielder, top heights were estimated using percentiles. The 99th percentile produced predictions very close to the observed values in the field (Table 1). The variability in the results could not be related to stand density, age, density of LiDAR returns or percentage of ground hits.

Plot ID	Observed Top Height	Estimated Top Height	Accuracy
6	32.81	32.53	0.99
7	28.15	28.15	1.00
8	20.33	19.48	0.96
9	21.35	19.9	0.93
10	17.92	17.82	0.99
11	20.78	20.42	0.98
12	21.24	21.01	0.99

Table 1. LiDAR predictions of top height against measured top height in Kielder.

5 CONCLUSIONS

In this paper, the use of commercial airborne X-band SAR and LiDAR have been compared. Both systems require an estimation of the tree canopy models as a difference between data obtained from the canopy tops and the signal from the underlying ground. In the case of SAR estimates, this operation required the use of

ancillary data such as the OS DTM. After this, the tree canopy models were classified differently to estimate top heights.

The findings are summarised as:

- Radar retrieved heights showed an underestimation of around 30%; this is explained by the signal penetration and attenuation in the canopy, especially in the less opaque species such as Scot pine and Japanese larch.
- LiDAR produced underestimations of around 7% in Aberfoyle, in line with the estimations of individual tree heights. This percentage was significantly reduced in Kielder by means of percentiles (99th) to less than 2%.
- Top Height estimations correlated well in both systems with R^2 values of between 0.81-0.92 for InSAR and 0.91 for LiDAR. In the percentile method with LiDAR data, the R^2 value approximated 0.99.
- Reasons for height under prediction were investigated. Canopy characteristics (shape and density) as well as viewing angle, tree height and density were seen as potential factors. The influence of edge effects was also examined, especially in SAR. The results indicated that better height retrievals were achieved far from edges in SAR. In LiDAR, all these comparisons proved inconclusive in the two areas with the two methods.
- The general conclusion is that both systems are close for operational use. The cost of the data and the availability of data providers will be the driving factors that will determine their business use. At present, there is no indication that SAR data will be periodically produced by Intermap. In contrast, LiDAR measurements can be produced on the same basis as the standard photogrammetric surveys. Additionally, their cost can be reduced altering mission parameters such as density of returns, flying height, scanning angle or size of the beam divergence. The new configurations will reduce the number of returns to the minimum necessary to achieve the desired results.

ACKNOWLEDGMENTS

The authors would like to thank Intermap Technologies and the Environment Agency for kindly supplying the data for the test sites. Thanks are also due to Stuart Snape (UK Forest Enterprise) for supplying ground reference data for Coed Y Brenin and Kielder respectively; Daniel Donoghue and Pete Watt of Durham University for supplementary ground data for Kielder, and to Iain Woodhouse and Izzawati of Edinburgh University for interpretation of radar modelling. This study was funded by the NERC and the UK Forestry Commission.

REFERENCES

- Hamilton, G.J. 1998. *Forest Mensuration*. Forestry Commission Booklet 39. Fifth impression. Her Majesty's Stationery Office, UK.
- Henderson, F.M., and Lewis, A.J. 1998. Principles and applications of imaging radar. Edited by Henderson, F.M., and Lewis, A.J. Manual of Remote Sensing, Third Edition, Volume 2. John Wiley & Sons Inc., New York.
- Intermap. 2003. Product Handbook and Quick Start Guide. Intermap Technologies Corp.
- Intermap. 2005. <http://www.intermap.com>. Intermap Technologies. Last accessed 4th April 2005.
- Suárez, J.C., Ontiveros, C., Smith, S. and Snape, S. (2005). 'The use of airborne LiDAR and Aerial Photography in the estimation of individual tree heights in forestry'. Computers and Geosciences. Vol 31, issue 2, March 2005, pp 253-262.
- Philip, M.S. 1994. *Measuring Trees and Forests*. Second Edition. CAB International, Oxon, UK.
- Wallington, E.D., Izzawati and Woodhouse, I.H. 2004. Forest Height Estimation from X-band SAR. In Proceedings of IGARSS 2004, 20th-24th September 2004, Anchorage, Alaska.

USING AIRBORNE LIGHT DETECTION AND RANGING (LiDAR) TO IDENTIFY AND MONITOR THE PERFORMANCE OF PLANTATION SPECIES MIXTURES

Pete Watt^a and Jimmy Wilson^b

^a Department of Geography, University of Durham, Durham, DH1 3LE, England

email: p.j.watt@durham.ac.uk

^b Forest District Planning, Galloway Forest District, Creebridge, Newton Stewart, DG8 6AJ, Scotland

ABSTRACT

For the effective management of forest resources, managers require accurate and up-to-date information not only on forest growth but also on the condition of forest crops. In even-aged single species plantations this type of information is typically gathered using a combination of field survey and aerial photo interpretation. However, the same exercise becomes more difficult when tree species are planted in alternate rows or intimately mixed in varying ratios and combinations. This process can be quite subjective and is affected by factors such as the timing, resolution and quality of the photography the size and shape of area of mixture and the photo interpreter and/or field surveyor. This paper investigates the ability of airborne **Light Detection and Ranging (LiDAR)** data for identifying and mapping Sitka spruce (*Picea sitchensis*) and Lodgepole pine (*Pinus contorta*) plantation mixtures in upland conifer plantations in the UK. Two approaches are assessed; the first uses LiDAR intensity data to separate pine and spruce species and the second uses LiDAR-derived crown density measures, coefficient of variation (*Cv*), skewness (*skewfp*); the percentage ground returns (*pczero*) which provides a measure of canopy openness and the mean canopy height (*meanh*) which enables areas with height variations to be identified. The merits of each measure are assessed separately and then combined to produce classifications that include and exclude LiDAR intensity. The results show that Sitka spruce and lodgepole pine can be separated using either LiDAR intensity or a combination of LiDAR-derived crown measures. Where the two species are planted as a mixture, the dominant species can also be identified and mapped.

Keywords: LiDAR, conifer identification, forest mapping

1 INTRODUCTION

During the 1960's the UK Forestry Commission established extensive areas of forest plantations in remote upland areas. Whilst a majority of the plantations were established using Sitka spruce, some of the poorer waterlogged or heather infested sites were planted using structured species mixtures of Sitka spruce and Lodgepole pine as a nurse species. These were planted in alternate rows or intimately mixed in varying ratios and combinations. The rationale being that the pine species would improve the site characteristics and over time these sites would become dominated by Sitka spruce which is the commercial crop. In practice, as these crops have matured, two possible outcomes have emerged: self thinning crops dominated by one species – usually Sitka spruce; or crops where both species compete equally. It is important for both outcomes to identify and map areas of different species dominance in order to produce reliable growth estimates. Traditionally, a combination of ground survey and aerial photography has been used to help provide this type of information, with the ground survey component providing quantitative measurements and the aerial photography used for crop stratification purposes. The accuracy of this method depends on how closely changes in crop structure on the ground can be matched with those observed in the aerial photography.

While a number of studies have used LiDAR to provide estimates of height, volume and basal area (i.e. Naesset, 2002; Means et al, 1999) relatively few studies have assessed the potential of LiDAR for identifying and mapping forest species. In Swedish forests, Holmgren & Persson (2003) successfully identified individual trees from high-resolution LiDAR data by using LiDAR to delineate and classify each tree by species. At a stand-level, Maltamo et al. (2004) identified suppressed trees in multi-layered spruce forests by summarizing the height distribution of the laser data.

This is the first study to evaluate LiDAR as an alternative and unbiased method for mapping crop variability in conifer plantation mixtures in the UK.

2 MATERIALS AND METHODS

2.1 Study area

The area for this study was located in Galloway Forest District forest south-west Scotland. The selected area was approximately 300 ha and contained a range of plantation types from single species crops to intimate mixtures. The terrain is gently undulating with a maximum slope of 14° and a topographic range of 230 to 300 m above mean sea level (amsl).

2.2 Image data

LiDAR data was acquired on 15/06/2003 using an Optech ALTM 3033 system operating at a flight altitude of 1,250 m amsl using a scan angle of <9°, with a point density of 4 returns/m². Also acquired was cloud-free SPOT 5 HRG multi-spectral scene on 17/04/2003. The NIR band of the SPOT 5 HRG image was also used for the purpose of validating the radiometric consistency of the LiDAR intensity channel data. The SPOT image was clipped so it matched the LiDAR coverage and then geo-corrected using the LiDAR data to provide GCPs. The reported RMS error of the geo-correction process was 0.5 pixel or 5 m.

2.3 Field data

Accurate height, volume and species data was measured for 20 (0.02 ha) field sample plots, 17 were located in pure Sitka spruce stands and three in Sitka spruce /lodgepole pine mixed stands. Since no plots were measured in pure lodgepole pine, height measurements from two areas were extracted from the 10 m LiDAR height grid (62 pixels). Table 1 summarises the plot data for each crop type by top height and volume. Plots measured in the Sitka spruce show the highest variability in both height and volume. This is because these plots cover a range of forest conditions. Variation in top height and volume is also observed in the Sitka spruce/lodgepole pine mixtures. However, in terms of volume, lodgepole pine tends to be the dominant species.

Table 1. Summary of field plot measurements by crop type

Crop Type	Tree species	Obs	Variable	Mean	Std. Dev.	Min.	Max.
		(No. field plots)					
Pure	Sitka spruce	17	Top height (m)	19.5	3.1	13.5	23.9
			volume (m ³ /ha)	500.7	160.6	184.2	697.4
Pure	Lodgepole pine*	(62)	Top height	15.8	0.9	13.0	18.0
			volume	n/a			
Species mixture	Sitka spruce	3	Top height	16.9	1.6	15.3	19.0
			volume	146.4	21.3	132.7	177.3
	Lodgepole pine		Top height	16.5	1.5	14.4	17.9
			volume	184.0	86.7	60.4	256.6

* Top height calculated from 62 pixels extracted from the 10 m LiDAR height grid

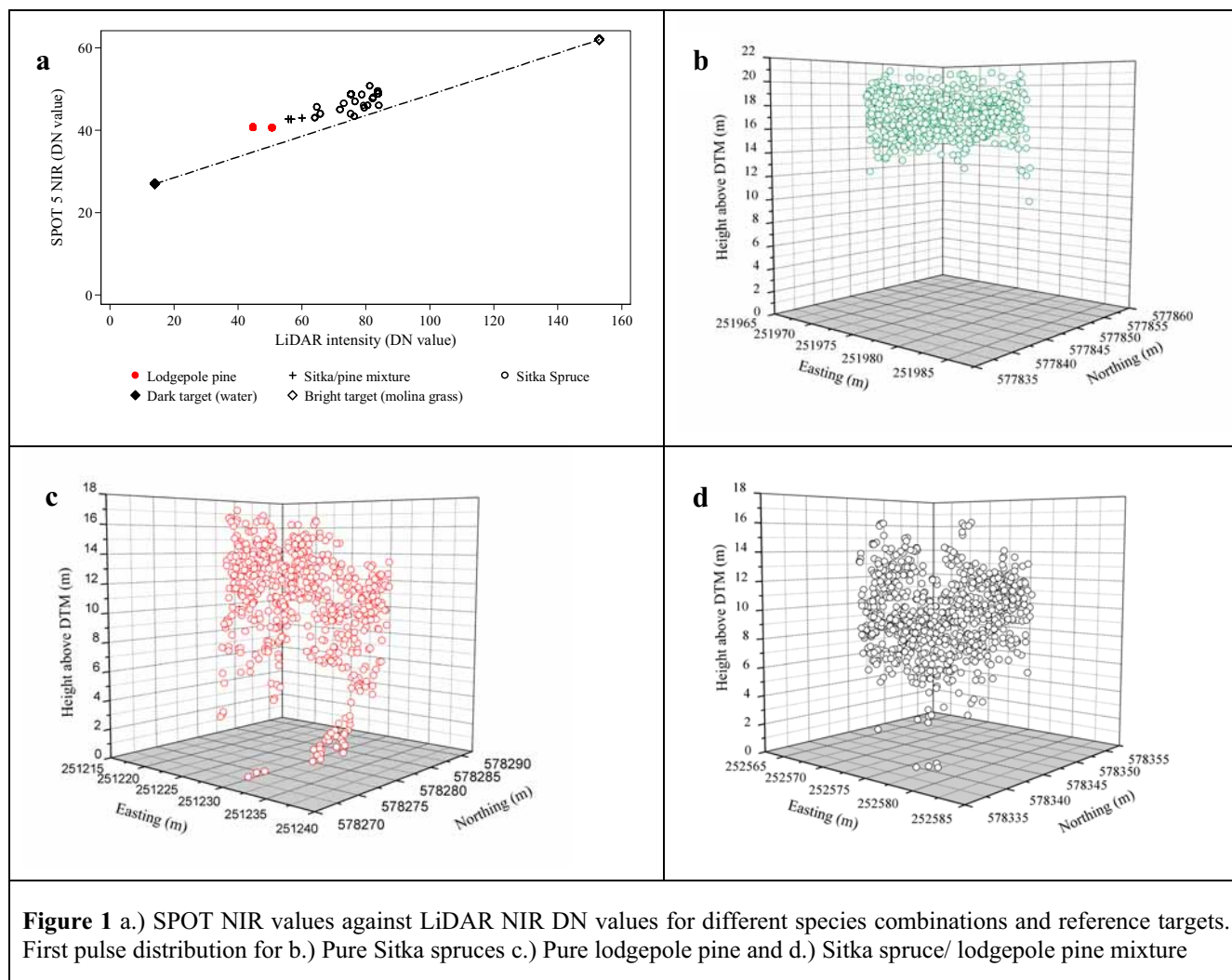
3 METHODS

Two approaches are used to identify plantation species composition; the first uses the LiDAR intensity data and the SPOT 5 NIR band. Here, the SPOT 5 NIR band is used as a means of checking the radiometric consistency of the LiDAR data. The second approach uses statistical parameters derived from distribution of the laser height data, namely the coefficient of variation (*C_v*), percentage of ground returns (*pczero*), skewness of height (*skewfp*) to try to identify differences in forest canopy structure. Mean height (*meanh*) is also included to provide a measure of crop variability. Each dataset was generated at a 10 m resolution, so that it matched the spatial resolution of the SPOT 5 image. Over each 0.02 ha sample plot the relevant pixel values were extracted for further analysis.

3.1 LIDAR INTENSITY

Over forest, the following factors may affect the recorded intensity: (i) variations in laser path length caused by changes in the distance between the sensor and target, (ii) the orientation of the target relative to the sensor which may change according to the laser scan angle or topography, (iii) the laser beam divergence which alters the footprint size, and (iv) attenuation of the signal by the atmosphere. Filtering of

the laser pulses by elimination of non-canopy returns, removes a substantial amount of the variation in the intensity values. Further correction for variations in laser path length also normalizes these differences between laser scan lines. Figure 1a shows the normalized LiDAR near infrared response plotted against the SPOT 5 HRG near infrared band for the different conifer species and reference targets. This shows that for similarly aged stands, pine and mixtures are spectrally separable from pure Sitka spruce in the near infrared.



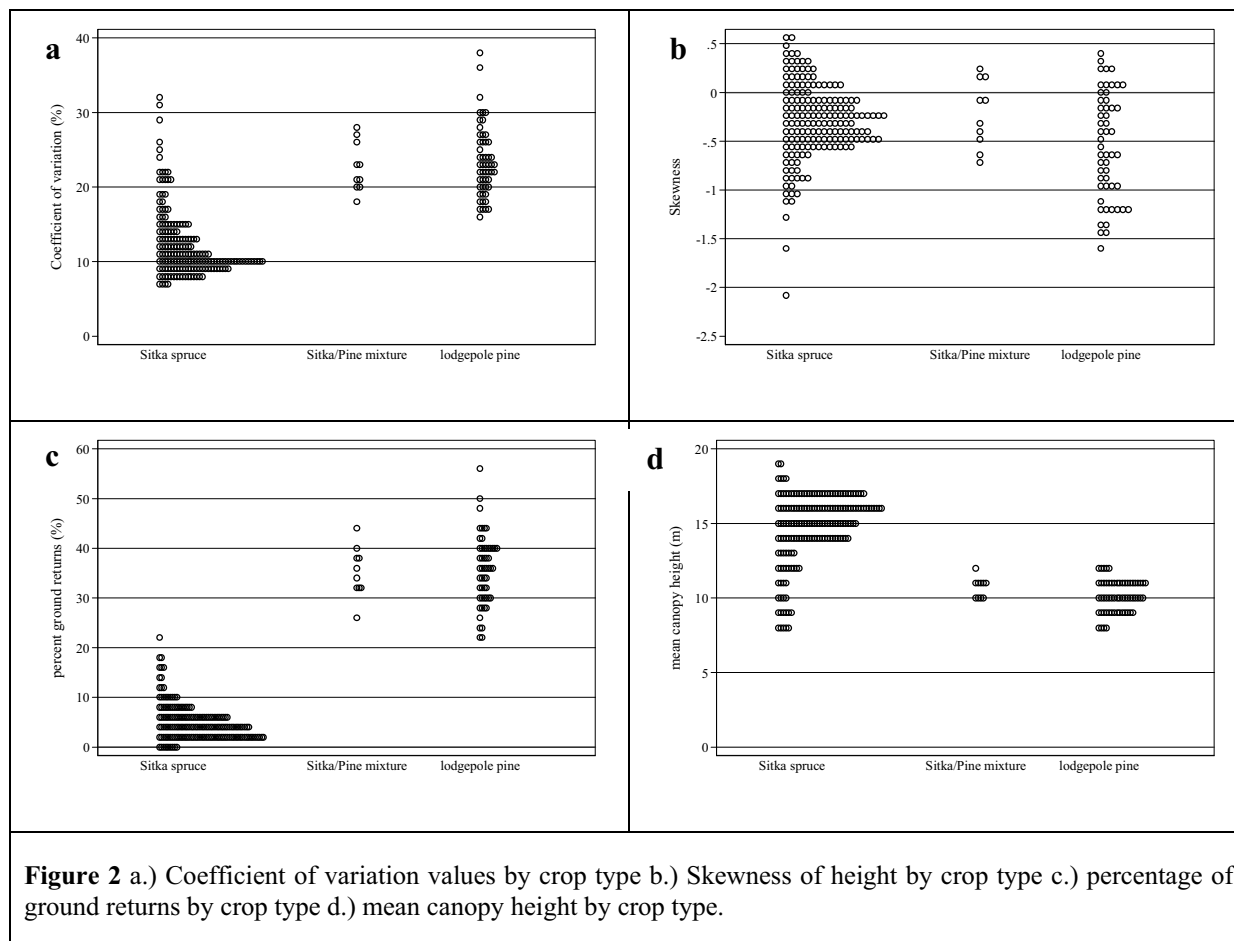
3.2 LIDAR CROWN DENSITY MEASURES

The growth and morphological characteristics of the conifer tree species established in this setting are different. For example, vigorous pure Sitka spruce crops form a dense forest canopy and reach canopy closure relatively early (approx 10 m). On the other hand, pure lodgepole pine or Sitka spruce/ lodgepole pine mixtures have open clumpy canopies. Figure 1 b to d shows that for Sitka spruce, lodgepole pine and species mixtures that the different canopy structures affect the distribution of the laser pulse (first pulse) through the canopy. The largest difference is observed between pure Sitka spruce and plots containing lodgepole pine. In this example, the density of the Sitka spruce canopy restricts the laser penetration to the top third of the canopy, whereas with the lodgepole and species mixture the laser penetrates almost to the ground. The distribution of laser pulses is similar for the pure lodgepole pine and the species mixture. The distribution in the pure lodgepole pine canopy is more homogenous than in the species mixture.

4 RESULTS

Quantitative analysis of laser intensity and measures of canopy density and height suggests that some measures are more effective at differentiating crop type than others. LiDAR intensity is the best single measure for identifying different species and species mixtures. Of all the species, Sitka spruce is the easiest to identify because of its distinctive laser height distribution (refer to Figure 1b). The most

promising variable, although not strictly a canopy density measurement is the percentage ground returns followed by the coefficient of variation. In pure lodgepole pine and Sitka spruce/ lodgepole pine mixtures, the same measures are effective, although the result is not as definitive as observed in Sitka spruce. Skewness of height and mean height are not very useful measures for discriminating between different tree species. Of all the measures assessed, mean height is the most effective for identifying wind damaged areas and areas of height variability.



In an effort to improve the accuracy of the species classification, all of the measures were combined in a supervised classification. The following classes were defined using a combination of the GIS compartment boundaries, SPOT and LiDAR data: (i) pure areas of Sitka spruce divided into two classes using mean height, (ii) areas of pure lodgepole pine and (iii) areas of wind damage. Non-forest areas were excluded from the classification. Two classifications were generated using the ENVI maximum likelihood classifier¹ algorithm. The first classification included canopy density measures and height and the second also included LiDAR intensity channel data. Figure 3 a and b show the results for each classification with the Forestry Commission's GIS compartment boundaries overlaid. Overall there is little difference between the two LiDAR classifications. According to the Forestry Commission's GIS the area of pure pine area is correctly classified. In plantation mixtures the classification provides more detail than the GIS data as it not only identifies the dominant species but also maps its spatial extent.

¹ Maximum likelihood classification assumes that the statistics for each class in each band are normally distributed and calculates the probability that a given pixel belongs to a specific class.



5 DISCUSSION

The objective of this paper was to assess the potential of LiDAR intensity data and crown density measures for identifying plantation mixtures and areas of variable growth. Although the study area is small (300 ha) the results strongly suggest that LiDAR data can be used to provide more than just estimates of forest height and volume. The methods developed show that in UK conifer forests species can be mapped using either LiDAR-derived crown density measures or LiDAR intensity data. Potentially, the intensity channel is more problematic to use, as the intensity is not radiometrically calibrated to a published standard. Also, for the reasons outlined in section 1 it is to be expected that intensity values will vary between different LiDAR flights. However, if the intensity of the returns is analysed in relative terms, or calibrated against another source, then it offers considerable potential for forest mapping. Filtering of the intensity data further improves the classification by removing unwanted noise after which it is possible to discriminate between different conifer species. A major limitation of the technique is in areas where topography is not constant (i.e. laser path length and footprint vary), although it appears that a more robust correction (i.e. using Luzum's method) can be applied if the position and angle of the laser pulse is known. However, this uncertainty is removed by using all the LiDAR variables in a statistical classification. This would allow the LiDAR variables to be combined with other remotely sensed data (i.e. IKONOS, SPOT, IRS or even Radar data), which could be used to further refine the classification or to validate classifications/predictions made using other data. In addition to identifying and mapping forest species, it is also possible to map areas of anomalous growth. Such areas are best identified using the crown density measures coupled with forest height. If compared with conventional methods that use manual interpretation of aerial photography, the LiDAR-based method is more accurate because it is much easier to identify areas of wind damage and variations in height that are not necessarily depicted on the aerial photography. The methods described provide an accurate means of identifying different forest species and areas of anomalous growth. This information will improve forest yield estimation by excluding non-productive areas, identifying areas of poor growth and quantifying and mapping species composition.

REFERENCES

- Holmgren, J., Persson, Å. 2003. Identifying species of individual trees using airborne laser scanning. *Remote Sensing of Environment* (90) 415–423.
- Maltamo, M., Packalén, P., Yu, X., Eerikäinen, K., Hyyppä, J., Pitkänen, J. 2004. Identifying and quantifying heterogeneous boreal forest structures using laser scanner data. In *Proceedings of the ISPRS working group VIII/2, Laser-scanners for forest and landscape assessment*, 3-6 October 2004, Freiburg, Germany.
- Means, J.E., Acker, S.A., Fitt, B.J., Renslow, M., Emerson, L. and Hendrix, C.J. (1999) Predicting forest stand characteristics with airborne scanning lidar. *Photogrammetric engineering and remote sensing* (66) 1367–1371.
- Naesset, E., 2002. Predicting forest stand characteristics with airborne laser using a practical two-stage procedure and field data. *Remote Sensing of Environment* (80) 88–99.

ASSESSMENT OF FOREST STAND PARAMETERS FROM LASERSCANNER DATA IN MIXED FORESTS

R. Wack^a and H. Stelzl^b

Joanneum Research, Institute of Digital Image Processing, Wastiangasse 6 8010 Graz/Austria,

^a email: roland.wack@joanneum.at

^b email: harald.stelzl@joanneum.at

ABSTRACT

This paper describes the results of an investigation on the extraction of forest stand parameters from LIDAR data in mixed forests. As a test site served a 500 ha forest dominated by spruce, pine, oak and six other tree species. The local forest management provided GIS forest inventory data as reference. In addition over 700 trees were measured. Results show that at least a separation coniferous and deciduous forest is required to achieve good results. Therefore the usage of high resolution multispectral satellite imagery is suggested. In addition a 3D visualization based on LIDAR data is discussed.

Keywords: forest stand parameters, data fusion, laser scanning, forest attributes, landscape visualization.

1 INTRODUCTION

The potential of satellite remote sensing for the classification of forests has been demonstrated by many projects. But since optical sensors are not able to provide information on the vegetation structure below the vegetation top layer, vertical forest structure, tree heights and timber volume can hardly be assessed. Furthermore, the classification of different development stages of alpine forests is only possible with an unsatisfactory accuracy. This is particularly true in inhomogeneous forests which are typical for the alps. Also forest openings are frequently classified as deciduous forest when dense vegetation predominates in the undergrowth. To overcome the above listed shortcomings of a classification based solely on multi-spectral satellite imagery, advantage can be taken from the synergetic effects of a combined use of satellite image data such as SPOT IV and V, Quickbird or IKONOS and airborne laserscanner data providing 3D information on the forest structure. Due to the fact that laser pulses are capable to penetrate the crown or vegetation layer, information on the vertical forest structure can be derived. On the other hand parameters such as forest types or tree species can be derived from multi-spectral satellite image data.

Based on high density airborne laser scanner data single tree modelling of forest trees is possible at a high success rate (Persson 2003). Especially coniferous trees require a high laser point density since the small tree tops need to be hit by the laser pulse to achieve good results for the tree heights. If the forest has been captured with lower point densities such as one point per m² and less, the modelling of individual trees becomes rather uncertain.

Since many large areas of several 1000 km² have been captured by laserscanning in Europe at such lower laser point densities, this data can be fused with satellite imagery to extract forest parameters. Not at single tree level but at stand wise level, using the information of the laser scanner data to extract parameters like vertical forest structure, average stand heights and timber volume. This approach was investigated based on data sets from a test site located near Burgau/Austria

2 EXTRACTION OF FOREST STAND PARAMETERS

2.1 TEST SITE

The test site Burgau for the extraction of forest stand parameters is located near Burgau in the eastern part of Austria, an area of 500 ha with sloppy terrain covered by a mixed forest. The dominant species are spruce, pine and oak, but about 6 additional species need to be considered. For this site LIDAR data, captured by Toposys, as well as a multispectral IKONOS scene were available. The local forest management provided GIS field data for 212 stands containing information on the species and the stem volume of each stand. For the investigation on the assessment of forest stand parameters like basal area, dominant height and stem number, 10 plots with 710 trees were measured additionally.

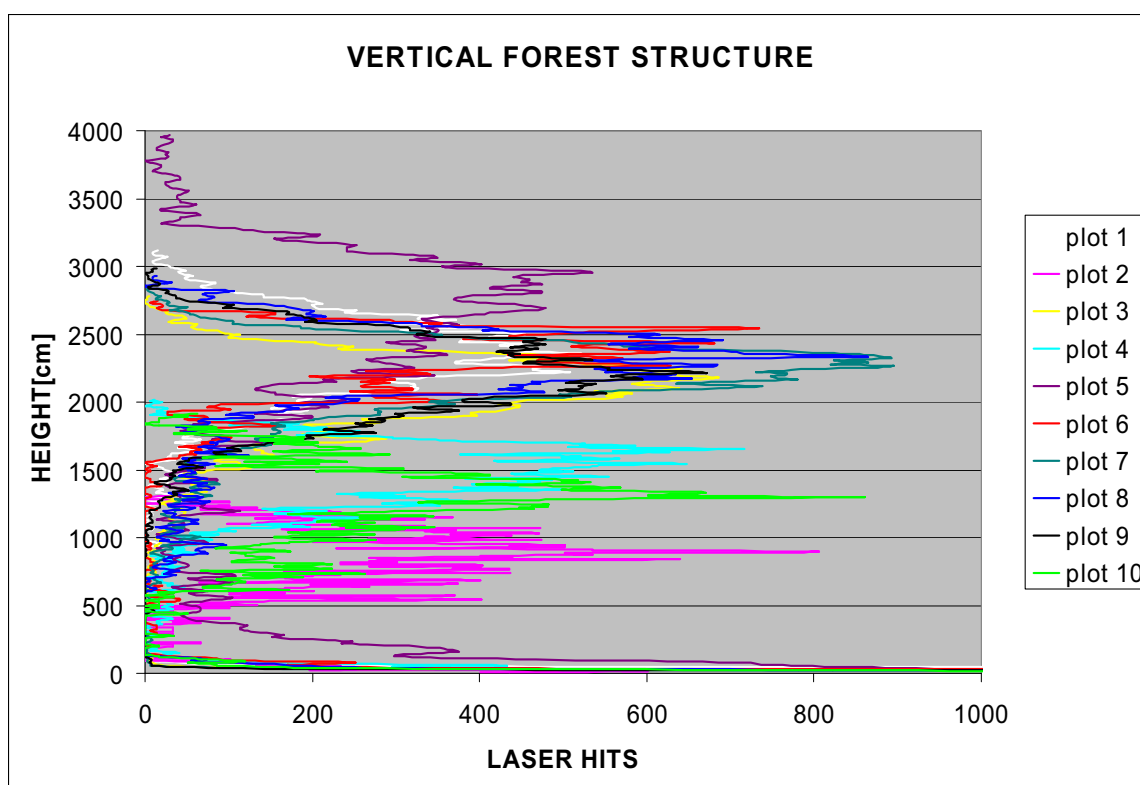
Table 1: results of the field work at test site Burgau – in total 710 trees were captured with position, height, dbh and species.

plot nr.	plot size	dominant species	plot stem nr.	stem nr. [trees/ha]	basal area [m ² /ha]	hdom [m ²]	timber vol. [m ³ /ha]
1	1291	spruce(83%), oak(10%)	41	318	26,76	24,97	356,99
2	135	spruce(95%) oak(5%)	20	1481	27,21	13,04	241,65
3	1863	spruce(64%) pine(25%)	108	580	25,86	20,84	322,01
4	838	pine(45%) spruce(27%) alder(17%)	83	990	29,63	16,09	285,40
5	1356	spruce(42%) oak(48%) maple(6%)	33	243	41,13	23,74	612,47
6	632	ash(71%) oak(9%)	58	918	31,07	19,43	359,76
7	1859	spruce(53%) pine(37%)	115	619	29,61	21,33	372,62
8	1365	spruce(56%) pine(30%)	104	762	32,92	19,24	407,90
9	2763	spruce(55%) oak(26%) pine(16%)	93	337	17,87	19,90	243,74
10	317	spruce(80%) birch(11%)	54	1703	36,31	16,07	331,14

2.2 VERTICAL VEGETATION STRUCTURE ANALYSIS

To extract information on the vegetation structure for each forest stand, a subset containing the laser points of the vegetation as well as the DTM of each stand was created, based on the GIS data of the stand boundaries used by the forest management. By subtraction of the DTM from the laser points of the vegetation, normalized vegetation heights could be generated for all points. To create a waveform like height distribution that shows the different stand characteristics, all laser points of a stand were accumulated according to their height above ground. Furthermore each data set was scaled by the stand area, thus making it possible to directly compare the waveform like distribution (see table 2).

Table 2: Vertical vegetation structure from LIDAR data - the different profiles show the distribution of LIDAR hits for each sample plot. The variations between the vertical forest structures are caused by different age classes, species and growing conditions at each site. Scrub growing beneath the trees clearly shows up in the data (plot 2).



Research on LIDAR data (Naesset, 2004) summarized to a waveform like data distribution that is describing the vertical vegetation structure of a stand has shown, that these profiles contain enough information to extract parameters like timber volume and basal area. To access the information of the waveform like profile a set of parameters need to be defined which help to characterise the form. Based on these parameters different predictive models can be set up and tested by regression analysis using ground truth data.

To characterise the shapes of the LIDAR profiles of the stands/plots Means et. al (Means et al, 1999) suggest: *mh* - mean canopy height; *qmh* – quadratic mean canopy height; *refl* – canopy reflectance. In addition the parameter *hdom* – dominant height of 95%; was used together with the first significant maxima from top described by its magnitude and height. These parameters were used before for the estimation of eucalyptus stand parameters and proved to be suitable for an accurate estimation of forest parameters (Wack 2004).

First, all stands of Burgau, regardless of their species composition, were used in a regression analyses to estimate stem volume. The correlation coefficient was 89% with a rmse value of 59. In relationship to the average stem volume of 185 m³/ha a rather pure result. Taking into consideration that the accuracy of the field data provided by the forest management may not be very accurate, the 10 sample plots were used in a next step for a regression analyses. The results are shown in table 3.

Table 3: Results of a regression analyses for 10 mixed plots; *mh* = mean canopy height; *refl* = canopy reflectance; *hdom* = dominant height(95%); *r*² = correlation coefficient; *max_a* = number of hits at first significant maxima from top; *max_h* = height at first significant maxima from top

reference parameter	used parameters from LIDAR data for regression analyses	<i>r</i> ²	Rmse
timber volume	<i>hdom</i> ^L , <i>max_a</i> , <i>max_h</i> , <i>refl</i> , <i>mh</i>	0.94	53.1
dom. height	<i>hdom</i> ^L , <i>refl</i> , <i>mh</i>	0.94	1.5
basal area	<i>hdom</i> ^L , <i>max_a</i> , <i>max_h</i> , <i>refl</i> , <i>mh</i>	0.85	5.7
stem nr.	<i>hdom</i> ^L , <i>max_a</i> , <i>max_h</i> , <i>refl</i>	0.90	274.4

Even though the differences of the vertical forest structures at the different plots are visible (table 2) the results of the regression analyses regarding the volume did not improve significantly. Either the selected parameter set for the description of the LIDAR profiles were not suitable or the species dependent tree shapes bias the parameterisation of the profiles. To investigate the influence of the species on the result of the regression analyses the forest stands were separated into 4 stratum. Stratum I consisted of all stands with a share of over 30% on deciduous trees. Stratum II covers all stands with 100% spruce. Stratum III contains all coniferous stands and stratum IV consist of all coniferous stands where the content of spruce was lower than 50%. The results of the regression analyses are shown in table 4.

Table 4: Results of a regression analyses for the estimation of tree volume using 4 stratum;

stratum	used parameters from LIDAR data for regression analyses	<i>r</i> ²	rmse
I(24stands)	<i>hdom</i> ^L , <i>max_a</i> , <i>max_h</i> , <i>refl</i> , <i>mh</i>	0.96	10.5
II(16 stands)	<i>hdom</i> ^L , <i>max_a</i> , <i>max_h</i> , <i>refl</i> , <i>mh</i>	0.89	17.3
III(134 stands)	<i>hdom</i> ^L , <i>max_a</i> , <i>max_h</i> , <i>refl</i> , <i>mh</i>	0.95	24,7
IV(25 stands)	<i>hdom</i> ^L , <i>max_a</i> , <i>max_h</i> , <i>refl</i> , <i>mh</i>	0,95	22,9

In contrast to the result achieved by a regression analyses using all forest stands with an rmse of 59, an improvement could be achieved by separating the stands according to tree species. The most significant distinction to be made is the one between coniferous and deciduous trees. In a second step the separation of spruce and pine would improve the result. Since many alpine areas are covered by a spruce and larch forest a classification of those two species may be helpful. For the extraction of forest stand parameters for larger areas, e.g. entire provinces, a coverage with GIS forest information will not be available. Therefore stands have to be created by using the vertical forest structure from LIDAR data in combination with classified tree species from multispectral satellite imagery like IKONOS or SPOT V. A simple distinction between coniferous and deciduous forest will help to improve the results of the estimated forest parameters. An

additional refinement of the classification of tree species will help to further improve the quality of the stand parameters.

3 3D-VISUALIZATION OF HIGH RESOLUTION LASER SCANNER RESULTS

3.1 GEO-DATA FOR 3D-VISUALIZATION

The results of the laser scanner data have been used to produce 3D landscape models in order to achieve an impressive and easy understandable visualization of the test area. The used data was a DTM and vector information of the classified vegetation as well as vector information about streets and lakes. Furthermore artificial ecosystems based on 2D- and 3D-objects have been used.

3.2 DEVELOPEMENT OF ARTIFICIAL ECOSYSTEMS

Beside the above mentioned geo-data from laser scanner and classification results virtual ecosystems had to be created. For this purpose an object library was developed, containing all necessary forms of vegetation like trees, scrub, meadows, but also textures for streets or lakes. The object library enabled the combination of particular objects and artificial ecosystems. The construction of this database, as well as the 3D visualization based on this database, was carried out using the animation software “Visual Nature Studio 2“ (VNS) of the American company “3D Nature“. The building of the object library was carried out using both real 3D objects and plain 2D images (mainly vegetation photographs). Vegetation can be simulated very appropriately by 2D objects which helps to save render time. The 2D images are always rotated in an upright position towards the optical axis of the camera and by means of shading a pseudo 3D visualization is accomplished (3D Nature, 2003).

For the test area of Burgau, textures for the specific local vegetation had to be created, because the textures available in VNS were not appropriate for the visualization of the test area. Respective photos of the local vegetation were integrated into the object library. A sufficient resolution of these photos and an exact release (clear separation from the background) had to be paid attention to. Thus, objects such as firs, pines, oaks, etc. were integrated. In order to design the ecosystems in a realistic way also textures for understorey and ground were included. To be able to model an uneven ground structure more appropriately, also “Bump Mapping“ was applied (Stelzl et al., 2004).

3.3 3D-VISUALIZATION ON THE BASIS OF VECTOR DATA

The classification results were stored in a GIS database which consisted of polygon objects with attributes about the distribution and age of the different species. In order to create a virtual landscape the attributes of the above mentioned polygons were connected with the objects in the library. By definition of the densities of the different species taken from the attributes in GIS realistic landscapes were created (Stelzl, Raggam, Sacherer, Almer, 2004).

In addition to the vegetation data vector information from streets and lakes were integrated and used for visualization. By applying lake effects and models for rural lanes the visualization was improved. The next figure shows a 3D-view of the test area based on the vector data and attributes:

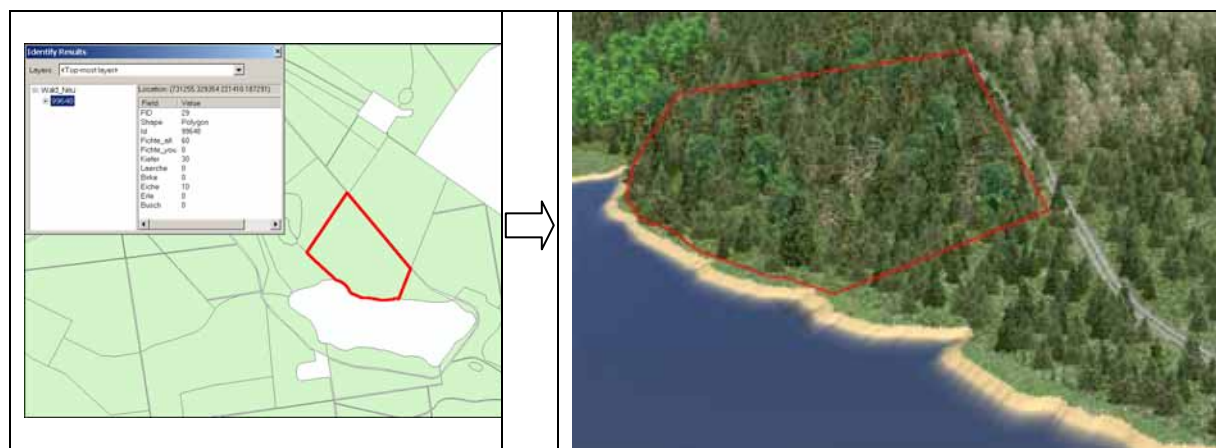


Figure: Rendering of a virtual landscape based on vector data in GIS data base. The rendered polygon defines an area with 60% fir, 30% pine and 10 % oak.

With this method a fully automatic 3D-visualization of the vegetation situation can be achieved. Every changes or planning tasks done in GIS can be visualized immediately by re-importing the changed shape-file.

4 CONCLUSIONS AND OUTLOOK

For the test site Burgau a stand wise estimation of forest parameters based on LIDAR data was carried out to investigate the suitability of such an approach within a mixed forest. The results achieved by regression analyses showed that a separation of tree species is necessary to achieve reliable results. A significant improvement can be achieved by separating coniferous and deciduous forest. A classification of additional tree species will again improve the result. Therefore a combination of LIDAR and a classified high resolution satellite multispectral imagery is suggested for the extraction of forest stands based on 3D structure and tree species. Such an approach will be tested in near future for the test site Burgau using IKONOS data. A second test site, South Tyrol (7500 km²), will soon be covered with LIDAR data and a SPOT V scene which will help to investigate the suggested approach on a larger scale.

Concerning the descript 3D visualization, an interface between the processing of airborne laser scanner data and 3D visualization was discussed.

The automation of the visualization plays a crucial role in the presentation of changes and planning tasks. By this means, an innovative tool for landscape planning and decision making is made available.

ACKNOWLEDGEMENT

This investigation on the extraction of forest stand parameters was partly financed by the EU within the project CarboInvent (Contract number: EVK2-2002-00157).

REFERENCES

- 3D Nature (2003): *Visual Nature Studio 2*. 3D Nature, LLC, 5740 Olde Wadworth, Suite C, Arvada, CO, USA 80002.
- Means J, Axcker St., Harding D., Blair B, Lefsky M, Cohen W., Harmon M, McKee W, 1999. *Use of Large Footprint Scanning Airborne Lidar To Estimate Forest Stand Characteristics in the Western Cascades of Oregon*, Remote Sensing of Environment 67, pp. 298-308.
- Harding D., Lefsky M, G. Parker, Blair B, 2001. *Laser altimeter canopy height profiles Methods and validation for closed-caopy, broadleaf forests*, Remote Sensing of Environment 67, pp. 298-308.
- Persson A., Holmgren J., Söderman U., 2003. *Detection, Measurement and species Classification of Individual trees for Forest Inventory and Visualization*. Proceedings Scandlaser Scientific Workshop Umea Schweden, pp223-234
- Stelzl H., Raggam H., Sacherer O. and Almer A. (2004): *Automatic 3D-Visualization of Geo-Information for Fireprevention in European Mediterranean Areas*. ISPRS - International Workshop on Processing and Visualization using High-Resolution Imagery; Technical Commission V/6 - Visualization and Animation, Pitsanulok, Thailand, 18-20 November 2004; in progress
- Stelzl H., Wimmer A., Raggam H., Sacherer O., Almer A., Banninger C. (2004): *Monitoring Mediterranean Wildlands with Very High Spatial Resolution Satellite Imagery for Fire Prevention and Control*. IGARSS 2004. Anchorage, Alaska. 20.-24.9.2004
- Wack R.,Oliveira T.,(2005) Analyses of vertical vegetation structures for fire management by means of airborne laser scanning. *Application in forestry- University of Göttingen*. pp 127-135

COMPUTER AUTOMATION OF A LIDAR DOUBLE-SAMPLE FOREST INVENTORY

Robert C. Parker

Department of Forestry, Forest and Wildlife Research Center, Mississippi State University,
Mississippi State, MS 39762, rparker@cfr.MsState.edu

ABSTRACT

The LiDAR (Light Detection And Ranging) Double-Sample automation system (LIDARDS) automates inventory and statistical computations for LiDAR tree height data and ground data in a stratified, double-sample inventory. The Windows based menu system provides a set of data formats and computational procedures that facilitate the rapid computation of a stratified, LiDAR-based, double-sample forest inventory. Sample tree diameters and heights from ground plots were used to obtain prediction equations for height and dbh of target trees identified on LiDAR surfaces. LiDAR heights in the Phase 1 data were allocated in a Monte Carlo simulation to species-product classes on each matching Phase 2 ground plot on the basis of percent distribution by numbers on the ground plot. Phase 1 LiDAR heights were randomly allocated to encountered species classes in each stratum and used to compute numbers of trees, basal area, and volume per acre. Phase 2 tree measures of dbh and height were used to compute LiDAR estimates of basal area (ft²) and volume (ft³) by using field derived dbh-height equations to predict dbh and volume. Comma delimited text files of Phase 1 and 2 estimates of trees, basal area, and volume on a per acre basis, double-sample regression estimates and associated precision and fit statistics, and partitioned volumes for each user-defined stratum were written to text files.

Keywords: LiDAR, inventory, double-sample, computer, automation

1 INTRODUCTION

Light detection and ranging (LiDAR) is a remote sensing tool that has the potential for acquiring measurement data for inventories of standing timber. Parker and Evans (2004) used small-footprint, multi-return LiDAR (0.25 hits per m²) in a double-sample application with a ground-based forest inventory in central Idaho and achieved an un-stratified sampling error of 11.5% on mean volume per acre at a 95% level of confidence. Sampling error was defined as one-half the confidence interval on mean volume expressed as a percentage of mean volume. Parker and Glass (2004) used small-footprint, multi-return LiDAR (4 hits per m², with a footprint size of 0.122 m and 1 hit per m² with a footprint size of 0.213 m) in southeast Louisiana to achieve stratified sampling errors of 9.5% and 7.6% (95% level of confidence) on mean volume per acre with the high- and low-density LiDAR, respectively. The standard and sampling errors were not improved when the high- and low-density LiDAR surfaces were smoothed (Parker and Mitchel 2005) or when LiDAR heights were adjusted to ground values with a regression equation (Parker and Glass 2004). The double-sample models used for the LiDAR-based inventories were adapted from traditional ground-based point sampling models. The objective of this study was to develop a user-friendly, computer application of a double-sample forest inventory that allowed the user to simultaneously analyze data from two LiDAR data sets and ground data for multiple species stands.

2 METHODS

2.1 LIDAR AND FIELD INVENTORY DATA

LiDAR data sets were surfaced to produce 1st return canopy and last return digital elevation models (DEM) with 0.2 m cell sizes using a linear interpolation technique. A spatial filtering technique derived from image analysis called smoothing was used to reduce tree location errors by minimizing the abrupt elevation changes in the initial canopy surface that could be erroneously interpreted as tree locations. A 1 m² filter moved across the LiDAR canopy surface, pixel by pixel, averaged the z-values within the window, and placed the result in the center pixel (McCombs et al. 2003). Tree height was interpreted as the difference between canopy and ground DEM z-values at each identified tree peak location.

Inventory design for this double-sample application involved the use of a systematic grid of circular plots 0.05 ac in size on a 52.6 ft by 330 ft grid with every 10th plot as a Phase 2 ground plot and all plots

being Phase 1 LiDAR plots. Ground data on each Phase 2 plot included tree diameter at breast height (dbh) on all trees > 4.5 in. dbh and total height, azimuth, and distance on 2 sample trees.

2.2 PHASE 2 SAMPLE TREE REGRESSION MODELS

Sample tree diameters and heights from ground plots were used to obtain prediction equations for dbh and ground height of target trees identified on the LiDAR surfaces for each of the encountered species groups. The dbh-height models used were:

$$\text{dbh} = b_0 + b_1 [\text{Ln}(H_{\text{gr}})]^{b_2} \quad (1)$$

$$H_{\text{gr}} = b_0 + b_1 (H_{\text{Li}})^{b_2} \quad (2)$$

$$H_{\text{gr}} = b_0 + b_1 (H_{\text{Li}})^{b_2} \quad (3)$$

where H_{gr} is measured ground height of trees identified on LiDAR plots, H_{Li} is estimated height of the same trees from LiDAR surfaces, and b_i are regression coefficients. Cubic foot volume of single trees was estimated with the equation developed by Merrifield and Foil (1967) to predict Minor's (1950) cubic volume.

2.3 DOUBLE-SAMPLE, REGRESSION ESTIMATOR MODEL

Phase 2 tree measures of dbh and height were used to compute LiDAR estimates of basal area (ft^2) and volume (ft^3) by using field derived dbh-height equations to predict dbh and basal area, and using dbh and height in a standing tree volume equation to predict volume. Thus, the model used in this computer application involved per acre mean estimates of LiDAR derived basal area and volume for the double-sample model:

$$\bar{Y}_{lr} = \bar{y} + \beta(\bar{X} - \bar{x}) \quad (4)$$

with variance:

$$s^2_{\bar{Y}_{lr}} = \frac{s^2_{y..x}}{n_2} + \frac{s^2_{y..} - s^2_{y..x}}{n_1} \quad (5)$$

where $\bar{Y}_{lr}$ = linear regression estimate of mean volume per acre from double-sample, $\bar{y}$ = mean value of volume per acre (y_i) derived from Phase 2 plots (n_2), $\bar{X}$ = mean LiDAR derived basal area or volume per acre from Phase 1 plots (n_1), $\bar{x}$ = mean LiDAR derived basal area or volume per acre (x_i) from Phase 2 plots, and β = linear regression slope coefficient for y_i as a function of x_i (volume or basal area).

2.4 REQUIRED DATA FILES

The computer application required Phase 1 LiDAR tree heights, Phase 2 tree data including LiDAR heights of sample trees, Phase 2 regression coefficients for the dbh-height and volume models for each species, dbh file of minimum and maximum dbh limits for each species-product combination, and strata definition of plot numbers and tree age by stratum in comma delimited formats. A user-defined data set name was prefixed to all system generated files and users may name all input data files.

Phase 1 LiDAR tree heights from each of up to two LiDAR data sets, where the file format is (plot#, LiDAR height).

Phase 2 data from ground plots, where the file format is (plot#, species code, product code, dbh, height, age, LiHds1, LiHds2) and the file name could be *PH2Tree.csv*. Height is ground measured height if the tree was a sample tree, age is tree age, and LiHds1 and LiHds2 are LiDAR heights from data sets 1 and 2, respectively. Plot trees with a dbh and height were sorted and used to obtain the regression coefficients for equations (1)-(3). Plot center was recorded with a Differential Global Positioning System (DGPS) and calculated sample tree coordinates were used on the LiDAR surfaces to locate "trees" that match the ground sample tree locations.

Phase 2 regression coefficients for each encountered species are listed in the comma delimited file in the equation sequence (1)-(3) and the Merrifield and Foil (1967) volume equation, where the file format was (b_0 , b_1 , b_2 , b_3) for all equations for each encountered species, and the file name was *Coeffic.csv*.

DBH minimum and maximum values for each combination of species and product class where the file format was (species code, product code, minimum dbh, maximum dbh) and file name was *DBH.csv*.

Strata definition and average age lists the stratum number, beginning and ending plot numbers that define the stratum, and average age of the stratum, if used, where the file format was (stratum#, beginning plot#, ending plot#, average age) and the file name was *Strata.csv*.

3 RESULTS

3.1 COMPUTER APPLICATION MENU

The computer application was a Windows based menu driven system. The user must complete the menu items in sequence from steps 1 through 9. Menu items 5-9 can be executed individually or menu item 10 will execute items 5-9 in sequential order if menu items 1 through 4 have been previously fulfilled. The menu items are:

1. Data Set: The user enters a unique data set name that will be used as a prefix to all system generated (intermediate and output) files.

2. User Parameters: The user must define various parameters, conversion factors, and counts that are used by the system during the computations. These items are explained in the user manual.

3. Phase 2 Regression Coefficients: The regression coefficients in the *coeffic.csv* file can be entered and edited with this option or the file can be created in a spread sheet, edited, and saved as a comma delimited text file in the data directory. The coefficients are stacked in the file by equation number and species.

4. Assign Plots and DBH Files: The *Strata.csv* file is a comma delimited text file containing stratum number, beginning plot number and ending plot numbers to define the stratum, and average age. The *DBH.csv* file is a comma delimited text file containing species code, product code, minimum dbh, and maximum dbh for each species-product combination.

5. Assign Heights and Volume to Phase 2 Trees: This was the first computation step in the system where plot totals of trees, basal area, and volumes on a per acre basis and percentages by species-product class were computed and a summary of Phase 2 ground estimates was obtained. Single tree volumes were computed for the trees in the *PH2Tree.csv* file using the volume coefficients from Merrifield and Foil's (1967) equation in the *Coeffic.csv* file and written to the output file *PH2TreeV.csv*. Plot totals and percent distribution of numbers of trees, basal area, and volume for each species-product class were computed and results written to the text file *Ph2Plot.csv* with plot totals for all species-product classes written to the text file *PH2PlotT.csv*. A summary of species-product totals and percent distributions for the total data set was stored in a text file named *PH2Sum.csv*, which was used in menu option 9 to allocate volume estimates from the linear regression procedures to species-product classes in an unstratified and combined stratum.

6. Iteration: LiDAR heights in the Phase 1 data set were randomly allocated in a Monte Carlo simulation to species-product classes on each matching Phase 2 ground plot on the basis of percent distribution by numbers of trees on the ground plot. Percent distributions of trees/ac by species-product class were obtained for each Phase 2 ground plot from the *PH2Plot.csv* file and the probabilities of occurrence were computed and ordered (highest to lowest) for each species-product class. Phase 1 LiDAR heights for the same plot were obtained from the *PH1ds1.csv* or the *PH1ds2.csv* data sets. The Phase 1 heights were allocated to the species-product classes a total of "maxIteration" times and mean values were computed.

7. Compute Phase 1 Species, N, BA and Volume: LiDAR heights from Phase 1 data files *PH1ds1.csv* and/or *PH1ds2.csv* were randomly assigned to species classes based on the percent distribution by species (in terms of number of trees per acre) in each stratum from the Phase 2 ground plot data. As each LiDAR height from the Phase 1 data file was read from the file, a species code was "assigned" based on a random assignment to the probability distribution for the stratum. The coefficients for the "assigned" species in the *Coeffic.csv* file were used to compute adjusted ground height, dbh, basal area, and tree volume for each LiDAR height in the Phase 1 plots.

8. Combine Phase 2 ground and LiDAR estimates: The previous Phase 1 LiDAR and Phase 2 ground data files were combined and the results written to a text file named *Phase2.csv*. This file was used to compute the regression relationships between ground volume ($y_i = \text{GVol}$) and LiDAR basal area or volume ($(x_i = \text{LiBA or } x_i = \text{LiVOL})$). A comma delimited summary file of Phase 1 and Phase 2 estimates of trees, basal area, and volume on a per acre basis for each LiDAR data set was created with the name of *VolSum.txt*.

9. Compute and Allocate Double-Sample Estimates: Regression estimates were obtained for the ground volume as a linear function of LiDAR basal area and LiDAR volume for up to 2 LiDAR data sets. The double-sample, linear regression estimates of mean volume per acre were computed with volume and basal area variables in linear function (4) for each stratum and combined strata. Regression estimates of mean volume per acre were obtained for non-stratified data, for each defined stratum, and combined strata estimates. The text file *Regress.txt* has the regression slope, index of fit, linear regression volume estimate, standard error of the linear regression estimate, sampling error at the 95% level of confidence, Phase 1 and 2 means for the independent variables, numbers of plots used for Phase 1 (N1) and Phase 2 (N2) estimates, and the number of random samples (NRS) versus double-sampling (N1 and N2 plots). Index of fit was defined as the proportion of total sums of squares explained by regression or $(1 - SS_{\text{error}} / SS_{\text{total}})$. The combined strata estimates for volume and standard error of each model and LiDAR data set were obtained by summing the weighted stratum estimates. Estimated samples sizes for Phase 1 and 2 of the double-sample were calculated with precision statistics from the current analysis and equations from Johnson (2000). The best regression estimate in terms of lowest sampling error was selected from each of the strata (non-stratified, user defined stratum, and combined strata) estimates and used to partition the mean volume estimate to the species-product classes. The mean volume was partitioned to the species-product classes on the basis of percent distribution of basal area and volume on the Phase 2 ground plots in each stratum. Each of the text files created by the system may be manipulated with a spreadsheet or viewed and printed with menu item 11.

10. Do Steps 5 thru 9 Consecutively: This menu option executes menu items 5 through 9 in a sequential manner. Menu items 5-9 were the consecutive steps necessary to compute the double-sample after all input items such as regression coefficients and strata plots are defined.

11. View Output Files: Summary files may be viewed or printed to paper with this menu option. The summary files available for viewing or printing are:

Volume Sum Text File (*VolSum.txt*): The volume sum text file is a summary from Phase 1 LiDAR and Phase 2 LiDAR and ground computations of per acre trees, basal area, and volume by species-product class for each sampling phase from menu item 8. The text file allowed the user to observe the differences and/or similarities between the Phase 1 and Phase 2 estimates of variables of interest.

Regression Text File (*Regress.txt*): The regression text file contains summary data relative to the computation of linear regression, double-sample estimates of mean volume per acre and associated precision statistics for user-defined strata with double-sample model (4). The dependent variable is mean volume per acre (ft^3/ac) for each of two independent variables, mean LiDAR basal area per acre (ft^2/ac), and mean LiDAR volume per acre (ft^3/ac) for up to two LiDAR data sets. Results were presented in two tiers for each user-defined stratum, un-stratified combined data, and stratified combined strata. First tier results included estimates of the slope coefficient, index of fit for the regression equation, adjusted linear regression estimate of mean volume per acre, standard error of the regression estimate, correlation coefficient between the dependent and independent variables, sampling error at the 95% level of confidence, Phase 1 sample size, and Phase 2 LiDAR and ground sample sizes. The second tier of results for a stratum lists Phase 1 and Phase 2 means of the independent variables LiDAR basal area and volume and the dependent variable ground volume, coefficient of variation, and sample size estimates for simple random sampling versus Phase 1 and Phase 2 sample sizes for a double-sample. Sample sizes were computed for a double sample assuming a 10:1 cost ratio between Phase 2 and Phase 1 plots and the calculated precision statistics for the stratum (Johnson 2000). The last tier in the regression text file contains the combined strata estimates of the linear regression estimate of volume per acre and its associated standard error and sampling error (at $\alpha=0.05$). The regression text file allowed the user to determine sampling gains with stratification for up to two sets of LiDAR data.

Strata Volume Text File (*StratVol.txt*): The strata volume text file used the “best” regression estimate in terms of lowest sampling error from each stratum at the 95% level of confidence and partitions the volume estimate into species-product classes using the average percent distribution of basal area and volume within the stratum. The strata volume text file allowed the user to see the best stratum estimate of mean volume per acre and its associated errors and the best estimates of species-product volumes.

12. Review Instructions: The brief instructions on data file contents and structures and output files were stored in a HTML help file and can be reviewed with menu item 12.

4 DISCUSSION

The LiDAR Double-Sample automation system automates the inventory and statistical computations for LiDAR data that have been previously processed to yield tree heights by plot and ground data that have been analyzed to yield regression coefficients for tree dbh and height relationships. Surfacing raw LiDAR data to produce a canopy and ground surface, interpreting the canopy surface for tree locations, and obtaining tree heights by plot location (McCombs et al. 2003) is an enormous task and this automation system does nothing to reduce the work load on the LiDAR data side of the process. It does, however, provide a set of data formats and computational procedures that facilitate the rapid computation of a LiDAR-based double-sample forest inventory. The LIDARDS system allows the user to set the number of iterations for the Monte Carlo simulation of species-product distribution on Phase 2 plots and whether to adjust LiDAR heights to ground heights with equation (3) before estimating dbh with equation (1). Parker and Glass (2004) and Parker and Mitchel (2005) found the LiDAR height to ground height adjustment process for high- and low-density LiDAR on smoothed and unsmoothed surfaces increased the sampling error of the volume estimates. By turning the “height adjustment” procedure on and off, the user can determine the effects of adjusting LiDAR height to ground height for each data set. The LiDAR Double-Sample automation system assumes the user can manipulate the raw data in spreadsheet software and save required data files in a comma delimited, text format. The LIDARDS system also assumes the user has an appropriate regression package that can produce the coefficients for the required tree dbh-height equations (1)-(3) from the ground data. An inherent disadvantage to the current system is the fixed models for the tree equations. Future versions of the LIDARDS system will offer other dbh-height models.

REFERENCES

- Johnson, E.W. 2000. *Forest Sampling Desk Reference*. CRC Press, Boca Raton, FL. 985 pp.
- McCombs, J.W., S.D. Roberts, and D.L. Evans. 2003. Influence of fusing lidar and multispectral imagery on remotely sensed estimates of stand density and mean tree height in a managed loblolly pine plantation. *Forest Science* 49(3):457-466.
- Merrifield, R.G. and R.R. Foil. 1967. Volume equations for southern pine pulpwood. Louisiana State University, North Louisiana Hill Farm Experiment Station Bulletin Forestry 7. Homer, Louisiana.
- Minor, C.O. 1950. Tables for estimating cubic volume of timber. Louisiana. Agriculture Experiment Station Bulletin 445. 39 pp.
- Parker, R.C. and D.L. Evans. 2004. An application of LiDAR in a double-sample forest inventory. *Western Journal of Applied Forestry* 19(2):95-101.
- Parker, R.C. and P.A. Glass. 2004. High versus low-density LiDAR in a double-sample forest inventory. *Southern Journal of Applied Forestry* 28(4):205-210.
- Parker, R.C. and A.L. Mitchel. 2005. Smoothed versus unsmoothed LiDAR in a double-sample forest inventory. *Southern Journal of Applied Forestry* 29(1):40-47.

Session 5a

MAPPING OF CANADIAN NORTHERN BOREAL FOREST BIOMASS USING QUICKBIRD HIGH SPATIAL RESOLUTION IMAGERY

A. Beaudoin^a, A. Leboeuf^b, L. Guindon^a, R.A. Fournier^b, J.E. Luther^c and M.-C. Lambert^a

^a Canadian Forest Service - Laurentian Forestry Centre, Ste-Foy, Canada, email: abeaudoin@nrcan.gc.ca

^b Département de Géomatique Appliquée, Université de Sherbrooke, Canada

^c Canadian Forest Service - Atlantic Forestry Centre

ABSTRACT

The limited availability of ground sample plots in Canadian northern regions requires a greater dependency of biomass mapping methods on modeling and remote sensing. We have developed and tested a method to map above-ground biomass of black spruce (*Picea mariana*) stands in northeastern boreal regions of Canada using high resolution satellite images. The method relies on QuickBird images transformed into biomass maps through image processing algorithms using the shadow fraction of tree crowns (SF). Three images provided test sites in diversified areas. A bitmap layer comprising the tree shadow was derived from the images. SF was calculated from the fraction of shadow area over the total area of reference square areas of fixed dimension overlaid on 108 ground sample plots (GSP). A linear regression was applied between GSP biomass and SF. The regressions for each of the test sites produced R^2 , RMSE and bias in the range of 85 to 88% (except one case at 44%), 14 to 18 t/ha and -3 to 8 t/ha, respectively. As statistical tests showed that there is no evidence that the 3 sites regressions are different, a global regression was calculated with all GSPs and produced R^2 , RMSE and bias of 82%, 15.3 t/ha and 4.2 t/ha respectively. Generalization of these results to extended areas of the boreal forest remains to be assessed in more details. However, the method appears to be an efficient mean to map biomass of black spruce stands in Canadian northern boreal forests.

Keywords: remote sensing, mapping, biomass, subarctic forest, black spruce, QuickBird, shadow fraction.

1 INTRODUCTION

The importance of carbon balance accounting related to climate change and increasing pressure for forest exploitation in the northern boreal forests of Canada has led to a demand for better estimates of total above-ground tree biomass at the stand level (hereafter referred to as biomass) and related carbon stocks. Unfortunately, the number of ground sample plots (GSP) in northern Canada is very small. Furthermore, GSP are very expensive to establish in northern regions due to the remoteness and lack of infrastructure. On the other hand, northern forests cover huge areas and are relatively homogenous in terms of species composition and structure. For instance, the boreal and taiga forests of eastern Canada, north of the 49th parallel in Quebec and Labrador, are largely dominated by black spruce (*Picea mariana*) stands. This context requires the development and use of satellite remote sensing methods to map biomass and calculate carbon stocks. Therefore, we investigated the usefulness of QuickBird (QB) high-resolution spatial imagery (HSRI) to provide biomass maps as surrogates to GSP.

Segmentation of HRSI has been used successfully to assess timber volume (Pekkarinen, 2002) and to identify land cover units in ecology (Lobo, 1997). Peddle and Johnson (2000), Peddle *et al.* (2001) and Seed and King (2003) found significant statistical relationships between tree shadow fraction (SF) and biomass or leaf area index of boreal forest stands. This relationship has been explained in large part by a model using spectral mixture analysis (Peddle and Johnson, 2000; Peddle *et al.*, 2001). The results of these studies suggest that SF may be a suitable variable for estimating biomass, provided that a reliable image processing algorithm to measure SF can be developed.

The goal of this study was to develop and test a SF-based method to map biomass of northern black spruce stands which can be applied to HRSI. The method provided biomass maps, which were subsequently used as surrogate GSPs to scale up biomass at the regional scale (Guindon *et al.*, 2005). This study was conducted in the larger context of the Earth Observation for Sustainable Development of Forests

(EOSD) project whose goal is to map the forest biomass of Canada using Landsat imagery (Luther et al. 2002).

2 TEST SITES AND DATA

Three test sites were selected within the eastern taiga shield and boreal shield ecozones (Fig. 1). The sites are near the towns of Chibougamau (CH) and Radisson (RA) in Quebec, and Wabush (WA) in Labrador. Their location and extent provided a good sample for the spatial variability of biomass range and forest stand conditions typical of northeastern boreal forests in Canada. Forest stands in the test sites were largely composed of black spruce (*Picea mariana*) with the presence of a few other species like jack pine (*Pinus banksiana*) (Rowe, 1972). The understory of the black spruce stands was generally covered by lichen, moss and shrub in various proportions. Topography was flat or gently rolling.



Figure 1: Location of the three test sites in two northern ecozones

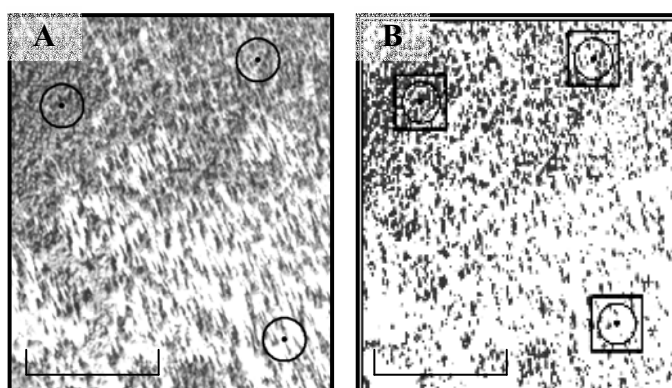


Figure 2: GSP and squares used to estimate SF, overlaid on A) the HRSI and B) the tree shadow bitmap layer

The spatial extent of each test site corresponded to the coverage of a single QB image (Table 1). A total of 108 circular GSP (400m²) dominated by black spruce (basal area > 75%) were established between 2002 and 2004 across the three test sites (31 in CH, 49 in RA, and 28 in WA) using a stratified semi-random sampling design. The position of each GSP was calculated from differential GPS with a precision better than 5m. Plots measurement followed the methods of the Canadian National Forest Inventory (Gillis, 2001). For a given GSP, the diameter at breast height (DBH) was measured for all trees (DBH > 5cm) and in a 4m sub-plot for seedlings (DBH < 5cm). Above-ground, oven-dry biomass of each tree (kg) including seedlings was estimated from DBH using the allometric equation developed by Ouellet (1983). Plot-level biomass, obtained by summing biomass values of all trees in the GSP, ranged from 4.8 t/ha to 163.0 t/ha with a mean value of 50.0 t/ha. Crown closure and mean height was also measured for each GSP with values ranging from 5 to 85% and 0 to 18m.

Table 1. QB image acquisition parameters for each test site

	CH	RA	WA
Date, local time	07/10/2003, 10:43	08/12/2003, 10:57	08/10/2002, 10:35
Sun elevation / azimuth (°)	58.6 / 143.2	49.0 / 152.9	51.1 / 157.7
Sensor elevation / azimuth (°)	78.5 / 96.4	76.7 / 172.3	80.4 / 225.6
Coverage (km²)	147	93	107

QB panchromatic (PAN) and multispectral (MS) images were acquired under clear-sky conditions with similar sun/sensor-viewing geometry for each test site and during the growing season in 2002 or 2003 (Table 1). Each QB image was geometrically corrected using at least 12 ground control points accurate within 5m. A first order polynomial resulted in an average positional RMSE of 10m. Since results achieved with both the PAN and MS-pansharped bands were similar (Leboeuf, 2005), here we report only on the results obtained using the PAN band (400-900nm) (subsection seen in Fig. 2a).

3 METHOD

The development of a method to estimate stand biomass from SF was divided into four separate sets of procedures that were designed to: (1) estimate SF values for GSPs from HRSI, (2) generate regression functions between GSP biomass and SF values for each test site, (3) calculate a global regression function for all sites and (4) map biomass as a grid layer for each site. The first set of procedures involved four steps. First, each HRSI was segmented with eCognition software (Definiens, 2002). Five scale factors were tested: 0, 5, 10, 20, and 40, whereby a scale of 0 implies no segmentation and an increasing scale factor produced objects of higher dimension (i.e. larger clusters of pixels). A tree shadow (TS) bitmap layer was generated from the segmented HRSI using the mean radiance value of pixels or objects and an image-specific threshold. The image threshold was determined by visual comparison of the threshold image with an enhanced image of the HRSI. Reference squares of fixed size were aligned on the 108 GSPs and provided a reference area from which SF (shadow area / reference square area) was calculated (Fig. 2b). Five sizes of squares were tested: 10, 30, 60, and 90m. Finally, SF values were geometrically normalized to a common sun-terrain-sensor viewing geometry to reduce variability associated with various image acquisition parameters.

A linear regression relationship was then established between GSP biomass and normalized SF for each of the three test sites. The regressions were calculated using a 70% random selection of GSP (GSP_{cal}) and adjusted R^2 values were reported. The remaining 30% GSP (GSP_{val}) were used to calculate error statistics: RMSE and bias. For each test site, linear regressions were calculated for each of the 20 potential configurations: five segmentation scale factors, four reference square sizes, and a constant image-specific classification threshold. The optimal configuration was identified as the one which maximised the R^2 , minimized RMSE and bias, while meeting practical considerations.

A weighted analysis of covariance was performed to assess if the relationships described by the regression equations for each test site were functionally the same. The residual normality was tested with Shapiro-Wilk test and the homogeneity of its variance was tested with residual graphics. The Fisher test was used to evaluate the coincidence and parallelism hypothesis with a threshold of $\alpha=0.05$ (Milton and Arnold, 2003) using GSP_{cal} data set. If no evidence was found that the three regressions were significantly different from each other, a global regression was calculated with a pooled dataset of all GSP_{cal} . Evaluation of these regressions relationship used the pooled dataset of GSP_{val} .

Finally, mapping biomass over the extent of each HRSI involved the derivation of an optimal TS bitmap layer using the optimal segmentation scale parameter and threshold. Post-classification was used to remove false TS such as water bodies having similar low radiance values. Then, for each cell of a grid layer where the cell size was given by the optimal reference square size, we calculated SF, normalized it and estimated biomass using the global regression equation. The resulting map was a grid layer where each cell with its biomass value can be used as a surrogate to traditional GSP. We assessed overall the precision of the three maps using the pooled GSP_{val} dataset.

4 RESULTS

A range of small scale parameters (0 to 10) and reference square sizes (10 to 30m) provided the best statistical results and were similar within these parameter ranges. Therefore, a segmentation scale parameter of 0 and a square size of 30m were selected as the optimal input parameters, considering the following practical considerations: (i) ease of application and short processing time of a pixel-based threshold approach and (ii) estimation done at the resolution of Landsat imagery appropriate for stand-level mapping and further biomass scaling-up at the regional scale using Landsat imagery. The three site-specific linear regressions $Bio = a + b \cdot SF$ with optimal input parameters are shown in Fig. 3a (dotted lines) with related statistics reported in Table 2. High R^2 values were obtained (85-88%) except for WA site (44%). This poorer result was partly explained by the more rugged relief, resulting in higher GSP-HRSI co-registration errors and additional variance of the SF values. Error statistics were similar between the three sites, with RMSE ranging from 14 to 18 t/ha and absolute bias value below 7.5 t/ha. There was no evidence that the three linear regressions were different ($F=0.89$ (p-value=0.41) for the slope and $F=2.14$ (p-value=0.12) for the intercept). The global regression using a pooled dataset of all GSP_{cal} (Fig. 3a) resulted in a RMSE value of 15.29 t/ha, providing a relative error of about 20%, and a low bias of 4.18 t/ha. Biomass estimated by the global regression equation corresponded well with biomass measured at GSP for the full range of conditions sampled (Fig. 3b).

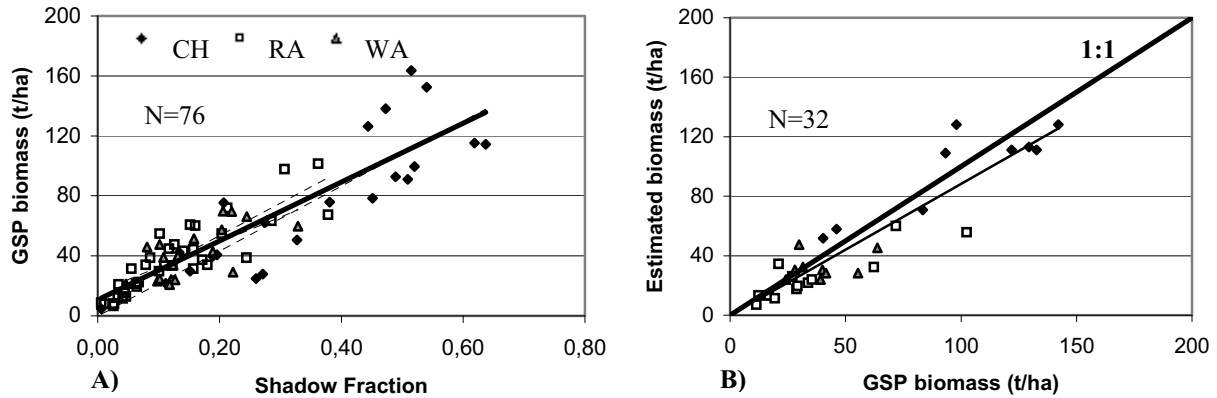


Figure 3. A) Linear regressions calibrated using 76 GSP_{cal} (dotted lines: local regressions; thick line: global regression; B) estimated vs GSP biomass using remaining 32 GSP_{val}.

Table 2. Local and global linear regressions with intersect *a* and slope *b*, and statistics from GSP_{cal} and GSP_{val}

GSP set	Nb GSP _{cal}	a	b	Adj. R ²	Nb GSP _{val}	RMSE	BIAS
CH	22	1.94	211.89	88 %	9	17.30	-2.50
RA	35	6.57	238.92	85 %	14	14.46	6.47
WA	19	16.19	177.40	44 %	9	14.39	7.31
Pooled	76	7.36	215.13	82 %	32	15.29	4.18

Assessment of the maps generated for the three test sites using the GSP_{val} pooled dataset provided an overall RMSE of 9.97 t/ha and a bias of 4.66 t/ha (Biomass map for Radisson in Fig. 4). Despite the low number of GSP used to validate the maps, biomass values were spatially consistent considering the strong influence of drainage and topography on forest biomass which is well reflected in the map (high biomass in valleys, low biomass in wetlands and on dry hill tops). In addition, mapped biomass values were consistent and close to zero over non-forested areas (shrublands, bare soils). However, water bodies of all sizes had to be masked out carefully through post-classification to avoid estimating biomass over them.

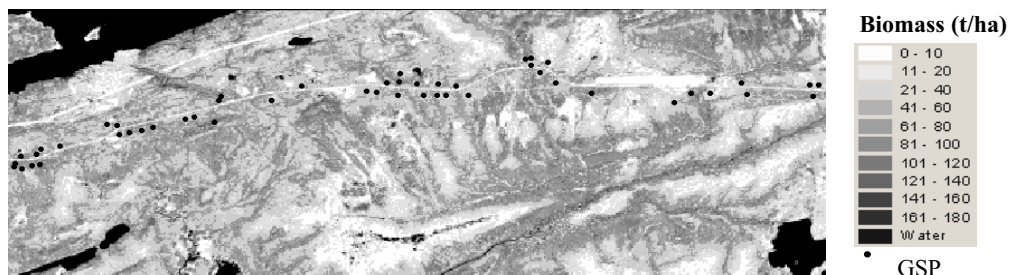


Figure 4. Biomass map derived from shadow fraction for Radisson test-site with GSP overlaid.

Several factors that may have contributed to the residual sources of variance between SF and biomass estimates and the reliability of the biomass mapping method were explored. For example, our analysis showed that the method can tolerate a large range of threshold values for each HRSI without increasing significantly the variance of SF. Preliminary statistical tests study showed that stratification based on understory type did not reduce significantly the variance of the resulting biomass. However, further investigation of the impact of understory type may be warranted given the highly variable combination of lichen, moss and shrub in black spruce stands. Geometric normalization taking into account the variability in sun-terrain-sensor viewing geometry was shown to be important for generalisation of the method across the three test sites. In particular, normalization for the variable sun elevation angle between HRSI acquisitions reduced significantly the variance of resulting biomass values. However, normalization of terrain effects were not addressed due to low slope and aspect angles for all GSP and for most test sites.

5 CONCLUSIONS AND PERSPECTIVES

We successfully developed and tested a SF-based method to estimate and map stand biomass of black spruce stands dominating the northeastern subarctic forests of Canada. This approach relied on normalized SF estimated from HRSI and used a pixel-based threshold approach. SF was then used as an independent variable in a global regression equation relating biomass to shadow fraction with a R^2 value of 82%. The sole use of the QB PAN band makes the method simple and cost effective relative to other approaches using HRSI. The method proved to be efficient to estimate biomass for: black spruce stands associated with lichen or moss; a range of biomass from 0 to 170 t/ha; and stand density and height up to 85% and 18m, respectively. Future work will assess further the validity domain of the method: (i) over a larger range of stand conditions (species composition, biomass range, understory type) and relief situations; (ii) using different HRSI types and related acquisition parameters; and (iii) for estimating other inventory attributes of interest.

ACKNOWLEDGMENTS

We acknowledge support for this study provided by the Canadian Space Agency's Government Related Initiatives Program and the Earth Observation for Sustainable Development Project. The CFS Ecoleap Project supported the establishment of the test sites and GSP. We thank Chhun-Huor Ung for his advice related to allometry and sampling design and Ron Hall for insightful discussions. We also thank Philippe Villemaire for his support in GIS and image processing. Finally we thank our field crew members: L. Guindon, S. Dagnault, J.-P. Bérubé, P. Villemaire, S. Côté, J. Donnelly and G. Strickland who collected ground truth data in challenging conditions.

REFERENCES

- Definiens inc. 2002. eCognition, version 3.1.
- Gillis, M.D. 2001. Canada's national forest inventory (responding to current information needs). *Environ. Monit. Assess.* 67: 121-129.
- Guindon, L. Beaudoin, A., Leboeuf, A., Ung, C.-H., Luther, J.E., Côté, S. and Lambert, M.C., 2005. Regional mapping of canadian subarctic forest biomass using a scaling-up method combining QuickBird and Landsat imagery, *In Proc. FORESTSAT 2005 workshop*, 31 May – 1 June 2005, Borås, Sweden.
- Leboeuf, A. 2005. Utilisation d'images Quickbird et Ikonos pour estimer la biomasse aérienne forestières de la forêt subarctique (boréale et taïga) basée sur une méthode de fraction d'ombre, M.Sc. Dépt. géomatique appliquée, Université de Sherbrooke, March 2005.
- Lobo, A. 1997. Image segmentation and discriminant analysis for the identification of land cover units in ecology. *IEEE. Trans. Geosc. Rem. Sens.* 35: 1136-1145.
- Luther, J.E., Fournier, R.A., Hall, R.J., Ung, C.-H., Guindon, L., Piercey, D.E., Lambert, M.-C., Beaudoin, A., 2002. A strategy for mapping Canada's forest biomass with Landsat TM imagery. *In Proc. Int. Geosc. Rem. Sens. Symp. and the 24th Can. Symp. Rem. Sens.* (CD-ROM), June 24-28, Toronto, ON.
- Milton, J.S. and Arnold, J.C. 2003. Introduction to probability and statistics: Principles and applications for engineering and the computing sciences. 4th edition.
- Ouellet, D. 1983. Biomass equations for black spruce in Quebec. *Environ. Can., Can. For.Serv., Laurentian Forestry Research Centre, Info. Report.* 11p.
- Peddle, D.R. and Johnson, R.L. 2000. Spectral mixture analysis of airborne remote sensing imagery for improved prediction of leaf area index in mountainous terrain, Kananaskis Alberta. *Can. J. Rem. Sens.* 26: 177-188.
- Peddle, D.R., Brunke, S.P. and Hall, F.G. 2001. A comparison of spectral mixture analysis and ten vegetation indices for estimating boreal forest biophysical information from airborne data. *Can. J. Rem. Sens.* 27: 6, 627-635.
- Pekkarinen, A. 2002. Image segment-based spectral features in the estimation of timber volume. *Rem. Sens. Environ.* 82: 349-359.
- Rowe, J.S. 1972. Forest Regions of Canada. *Can. For. Serv., Depart. Environ.*, Publication No. 1300.
- Seed, E.D. and King, D.J. 2003. Shadow brightness and shadow fraction relations with effective leaf area index: importance of canopy closure and view angle in mixedwood boreal forest. *Can. J. Rem. Sens.*, 29: 324-335.

REGIONAL MAPPING OF CANADIAN SUBARCTIC FOREST BIOMASS USING A SCALING-UP METHOD COMBINING QUICKBIRD AND LANDSAT IMAGERY

^aL. Guindon, ^aA. Beaudoin, ^bA. Leboeuf, ^aC.H. Ung, ^cJ. E. Luther, ^aS. Côté and ^aM.-C. Lambert

^a Canadian Forest Service, Ste-Foy, QC, Canada, G1V 4C7, e-mail: lguindon@cfl.forestry.ca

^b CARTEL, Université de Sherbrooke, Sherbrooke, QC, Canada, J1K 2R1

^c Canadian Forest Service - Corner Brook, NL, A2H 6J3

ABSTRACT

Mapping forest biomass of the subarctic region of Canada is important for inventory and reporting purposes. However, the limited availability of ground sample plots (GSP) in huge and remote regions requires mapping methods with greater dependence on modeling and remote sensing. We developed and tested a method to map the biomass of black spruce (*Picea mariana*) stands of Canadian subarctic forest using GSP, satellite sample plots (SSP) derived from QuickBird imagery, and Landsat Thematic Mapper imagery. Local scale biomass was mapped using three QuickBird panchromatic images and a global regression model of above ground biomass as a function of shadow fraction (SF) of the images (reported in Beaudoin et al. 2005). Herein, the QuickBird-derived biomass maps provided surrogates to traditional GSP for mapping biomass at the regional scale. Biomass grid cells of the local-scale maps were randomly sampled (n=600) and the K nearest neighbour (KNN) method was applied to the full extents of three Landsat images, representing three test sites. RMSE and bias were calculated using (i) GSP to estimate the combined error of scaling from the plot to the regional level and (ii) SSP to estimate the scaling error from application of the local biomass maps to the regional level. The RMSE of the regional scale biomass maps ranged from 10.1 to 19.6 t/ha with an overall RMSE of 17.2 t/ha based on the GSP. Bias estimates were only slightly positive with an overall bias of 2.8 t/ha for the three test sites. Overall, the scaling up method produced good estimates of biomass over the three test sites with very low biases and relative errors in the order of 20-30% depending on the test-site. Further developments will consider extension of the method across large areas of Canadian subarctic forests.

Key words: remote sensing, subarctic forests, biomass, Landsat, KNN, shadow fraction.

1 INTRODUCTION

The importance of carbon accounting in relation to climate change and the pressure of forest exploitation in northern boreal forests of Canada have led to an increased demand for estimates of the above-ground biomass and carbon stocks. Northern boreal forests in Canada cover about 4 billion km² across five ecozones. These forests are relatively homogenous in terms of species composition and structure, yet largely non-inventoried with limited accessibility. This context requires the development and use of satellite remote sensing methods to map biomass and calculate carbon stocks.

A variety of approaches exist to map biomass using Landsat imagery (Fazakas et al., 1999; Fournier et al., 2003; Hall et al., 1997; Pilger et al., 2002). In particular, the K-Nearest Neighbours (KNN) method is a non-parametric method that has been used extensively in the Scandinavian boreal forest context (Fazakas et al., 1999; Reese et al., 2002). KNN, as many other methods, requires ground sample plots (GSP) representing the full range of forest conditions found within the mapping extent. When a good representation of ground sample plots exists, this method has performed well for biomass estimation. However, in the absence of a consistent network of GSP in northern regions of Canada, an alternative approach for implementation of the KNN is required.

In earlier work, we mapped biomass at a local scale using QuickBird (QB) high resolution spaceborne imagery (HRSI) with a shadow fraction (SF) based approach (Beaudoin et al., 2005). This effort provided local scale biomass maps from QB images. In this paper we report on the use of Satellite Sample Plots (SSP) derived from QB biomass maps to test and implement the KNN method at a regional scale for mapping the biomass of northern black spruce (*Picea mariana*) stands. This study was conducted in the larger context of the Earth Observation for Sustainable Development of Forests (EOSD) project whose

goal is to map the forest biomass of Canada using Landsat imagery (Luther et al. 2002; Labrecque et al., submitted) and the Ecoleap project (Bernier et al. 2000; Ung et al. 2002).

2 TEST SITES AND DATA

Application of the KNN method using SSP derived from QB HRSI for regional scale biomass mapping was developed and tested over 3 test sites in Northern Canada (Fig. 1). The test sites (Chibougamau: CH; Radisson: RA; Wabush: WA) were located within two subarctic ecozones in the eastern taiga shield and the boreal shield (Quebec and Labrador). The sites were selected to cover a wide biomass range and represent the dominant stand conditions of Northeastern boreal forests.

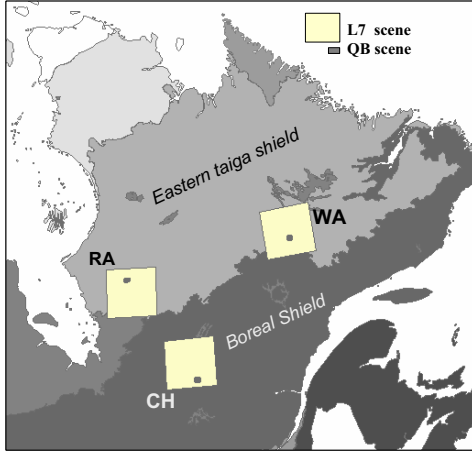


Figure 1. Location of the three test-sites in two northern ecozones, imaged by nested sets of QB and L7 images

Table 1. Landsat and QuickBird acquisition parameters

	CH	RA	WA
L7 Path/Row	16/25	18/23	12/23
Acq. date	08/28/2002	08/20/2000	09/25/1999
QB acq. date	07/10/2003	08/12/2003	08/10/2002

Each site was imaged by a QB HRSI nested in a L7 image. Each QB and L7 image pair was acquired under clear-sky conditions during the growing season between 1999 and 2003, except for one L7 image acquired in the Fall (Table 1). In each case, the coverage of the QB images (50-100 km²) was selected to optimize the representativeness of forest stands relative to those within the L7 coverage (34,000 km²). This was done by comparing the relative spatial occurrence (%) of five coniferous classes having different range of crown closure (CC) values (very sparse: 0-10%; sparse: 10-25%; open: 25-60%; dense: 60+%; treed bogs) provided by a land cover map derived from the L7 image using the ECM classification method (Beaubien et al., 1999). The location of QB images provided relative occurrences of those forest classes always within 15% and most of the time within 10% of the class occurrences within the L7 extent.

Each QB image was geometrically corrected with a 1st order polynomial and a resulting average positional RMSE (root mean square error) of 10m. Each panchromatic (PAN) band was transformed to local scale biomass maps using a global regression relationship established between biomass measured at each GSP plot and the shadow fraction extracted from each QB image (Beaudoin et al., 2005). Finally, each L7 scene was orthorectified with a resulting positional RMSE in the order of 15-20m and transformed into top-of-atmosphere reflectance.

3 METHOD

For each test site, we randomly sampled each local scale QB biomass map made to obtain N SSP used as a training set (SSP_{cal}) in the following KNN equation:

$$BIOknn_i = \frac{\sum_{k=1}^K W_k BIOssp_k}{\sum_{k=1}^K W_k} \text{ for } k \neq i \quad \text{where } W_k = \frac{1}{d_k^j} \quad (1)$$

where $BIOknn_i$ is the KNN biomass estimate for the i^{th} Landsat pixel, k is the k^{th} out of K nearest spectral neighbours, $BIOssp_k$ is SSP biomass, W_k is a weighting coefficient, d_k^j is the spectral Euclidian distance to which is applied a j power value. Biomass estimated using KNN is therefore an average of biomass values from the SSP which are spectrally the nearest from the considered Landsat pixel, where $j = 0$ results in a simple average and $j = 1$ or $j = 2$ provides an average weighted to the 1st or 2nd power of the inverse of the spectral distance.

Various combinations of N, K and j were tested within the following ranges: (i) $K = 1$ to 20, (ii) $j = 0, 1$ or 2, (iii) $N = 100, 300$ and 600. Each combination was tested with or without a coniferous forest mask to

remove non-coniferous pixels, considering that the method was being developed for coniferous forests dominated by black spruce. In all cases, we removed SSP with high surrounding spectral variability to account for co-registration errors between the QB biomass map and the L7 image. This was accomplished using a threshold applied on the L7 near infrared band 3x3 variance. For each combination of input parameters, Eq. 1 was applied to all Landsat 30m pixels within the extent of the QB coverage, providing a biomass grid layer where each 30m cell had its biomass estimate given by the KNN method, *BIOknn*.

RMSE and bias were calculated using: (i) a 30% GSP_{val} set to estimate the combined error of scaling from the plot to the regional level, and (ii) the full SSP dataset to estimate the scaling error from application of the local biomass maps to the regional level. Finally, we selected the combination of input parameters minimizing RMSE and bias to produce a final biomass map. *BIOknn* statistics (mean and standard deviation) were reported for five coniferous classes of different CC ranges.

4 RESULTS AND DISCUSSION

Concerning the optimal set of KNN input parameters, it was found that (i) the statistics were systematically stabilized for $K > 10$, (ii) the number of SSP_{val} was optimal around 600 but with a small improvement compared to 100 to 300, (iii) the power of the weighting coefficient j had no significant impact and (iv) the use of a coniferous stratification mask provided slightly poorer results than without using it. In the following, we report results using $N=600$, $K=15$, $j=1$ and no coniferous mask.

Fig. 2 reports *BIOknn* and *BIOssp* estimates compared to reference *BIOgsp* values for each test site. Good agreement was found between *BIOknn* and *BIOgsp* ($R^2=80\%$) with a slight under-estimation and increased variance at the higher end of the biomass range. Moreover, regional *BIOknn* estimates were only slightly more scattered than local *BIOssp* estimates.

RMSE and bias measures of *BIOknn* based on plot-level *BIOgsp* are reported in Table 2a, including those for *BIOssp* (Beaudoin et al., 2005) for comparison purpose. RMSE values were lowest for the WA site (10.1 t/ha) and highest for the CH site (18.9 t/ha) with an overall RMSE of 17.2 t/ha, which provided a relative error of about 20-30% depending on the site. Bias was less than 5 t/ha for all sites. Estimation errors for *BIOssp* are only slightly lower overall (15.3 t/ha), suggesting that the KNN method did not bring significant additional errors in the regional estimates compared to the local ones. Overall bias estimates are slightly higher for the local compared to the regional estimates (4.2 and 2.8 t/ha, respectively). Table 2b reports RMSE and bias for *BIOknn* based on the large set of *BIOssp* (all cells of the QB biomass map) to report the errors only due to the local to regional scaling-up. RSME values are comparable to those of the local scale maps validated with the GSP (13.8 and 13.3 t/ha, respectively). Bias tends toward 0 as the number of SSP samples increases and is consistent with observations reported by Fazakas et al. (1999).

Mean and standard-deviation of mapped *BIOknn* values within each coniferous cover class exhibited the expected biomass increase with increasing CC (Fig. 3). The 60-100 % CC class in Chibougamau showed a higher estimated biomass range compared to other sites because of denser stands more present in this southern site. Over-estimation was found in (i) some bogs or along lake shores but represent a small fraction of the test sites and (ii) on steep slopes opposite to the sun.

Overall, the scaling up method produced good estimates of biomass over the three test sites with very low biases and relative errors of the order of 20-30% depending on test-site. A source of error was due to

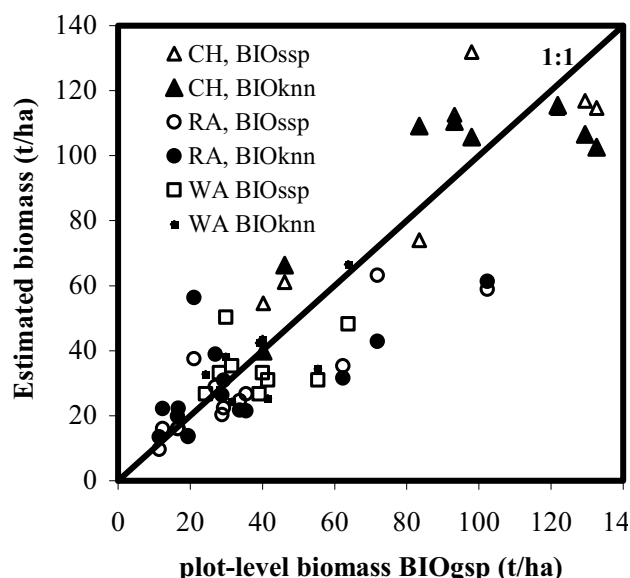


Figure 2: Estimated biomass *BIOssp* (open symbols) and *BIOknn* (close symbols) compared to GSP biomass *BIOgsp* for each test-site (CH, RA, WA)

co-registration errors between the QB and L7 images, as the outlier *BIOknn* estimates were generally those for GSP located in sparse coniferous forests with highly variable understory (lichen, shrubs). Therefore, heterogeneous pixels should be excluded from the SSP training set. In addition, terrain radiometric normalization could be tested to account for slope effects. Significant acquisition lag between the QB and L7 acquisition may have caused some problems due to land cover changes (fires, logging, insect). Overall, it was found that application of the KNN method using SSP was unbiased and *BIOknn* estimates reflected random and systematic errors brought by the SF approach applied to the QB image (Beaudoin et al., 2005). Above all these variance sources, a critical aspect is the representativeness of small QB samples within the extent of the Landsat scene, which will need further investigation towards applying this scaling-up method across large areas of Canadian subarctic forests.

Table 2a: Errors for *BIOssp* using GSP_{val} datasets, including *BIOknn* for comparison

SITE	Nb GSP _{val}	<i>BIOknn</i>		<i>BIOssp</i>	
		RMSE (t/ha)	Bias (t/ha)	RMSE (t/ha)	Bias (t/ha)
CH	9	18.9	0.7	17.3	-2.5
RA	14	19.6	4.7	14.5	6.5
WA	9	10.1	2.0	14.4	7.3
ALL	32	17.2	2.8	15.3	4.2

Table 2b: Errors for *BIOknn* using SSP_{val} datasets

SITE	Nb SSP _{val}	<i>BIOknn</i>	
		RMSE (t/ha)	Bias (t/ha)
CH	95 749	13.3	0.6
RA	104 244	9.1	0.1
WA	62 969	19.6	0.2
ALL	262 962	13.75	0.31

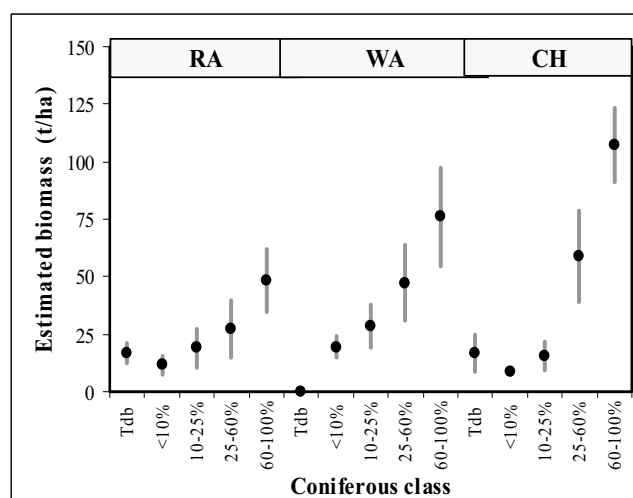


Figure 3: *BIOknn* (mean $\pm$ 1 std) for each test site within each coniferous class of various CC ranges (tdb = treed bog).

5 CONCLUSION

The application of Quickbird-based SSPs to implement the KNN method on Landsat images resulted in biomass estimates with overall RMSE and bias estimates of 17.2 and 2.8 t/ha respectively. These errors were similar to those of applying biomass equation based on tree shadow fraction of Quickbird image to produce local scale maps. Future development will expand the method across northern Canada based on a stratified sampling design, using a network of representative QB HRSI samples scattered within a normalized mosaic of Landsat images for all the subarctic forest.

ACKNOWLEDGEMENTS

We acknowledge support for this study provided by the Canadian Space Agency's Government Related Initiatives Program and the Earth Observation for Sustainable Development Project. The CFS Ecoleap Project supported the establishment of the test sites and GSP. We are grateful to Sandra Labrecque, Richard Fournier and Ron Hall for their helpful insights through many scientific discussions on biomass mapping in the Canadian context using various methods including KNN. We also thank Philippe Villemaire for his support in GIS and image processing. Finally we thank our field crew members: S. Dagnault, J.-P. Bérubé, P. Villemaire, J. Boudreau, G. Simard, N. Laflamme, J. Donnelly and G. Strickland, who collected ground truth data supporting this work under challenging conditions.

REFERENCES

- Beaubien, J., Cihlar, J., Simard, G. and Latifovic, R., 1999. Land cover from multiple thematic mapper scenes using a new enhancement-classification methodology, *Journal of Geophysical Research*, 104, 27909-27920.
- Beaudoin, A., Leboeuf, A., Guindon, L., Fournier, R.A., Luther, J. E., Lambert, M.-C. and Ung, H.-C. 2005. Local mapping of Canadian subarctic forest biomass using QuickBird high spatial resolution satellite images . In Proc. FORESTSAT 2005 workshop, 31 May – 1 June 2005, Borås, Sweden.
- Bernier P.Y., Fournier R.A., Ung C.H., Robitaille G., Larocque G.R., Lavigne M.B., Boutin R., Raulier F., Paré D., Beaubien J. and Delisle C. 2002 Linking ecophysiology and forest productivity: An overview of the ECOLEAP project. *The Forestry Chronicle*, vol.75, no. 3, 417-421.
- Fazakas, Z., Nilsson, M. and Olsson, H., 1999. Regional forest biomass and wood volume estimation using satellite data and ancillary data, *Agricultural and Forest Meteorology*, Vol 98-99, pp. 417-425.
- Fournier, R. A., Luther, J. E., Guindon, L., Lambert, M.-C., Piercey, D., Hall, R. and Wulder, M. 2003. Mapping aboveground tree biomass at the stand level from inventory information: test cases in Newfoundland and Quebec, *Canadian Journal of Forest Research*, 33: 1846-1863.
- Hall, F.G., Knapp, D.E. and Huemmrich, K.F., 1997. Physically based classification and satellite mapping of biophysical characteristics in the southern boreal forest. *Journal of Geophysical Research* 102(D24): 29567-29580.
- Labrecque, S., Fournier, R.A., Luther, J. E. and Piercey, D. E. A comparison of four methods to map forest biomass from Landsat-TM and inventory data in western Newfoundland, *Forest Ecology and Management* (Submitted)
- Luther, J. E. Fournier, R. A. Hall, R. J. Ung, C-H. Guindon, L. Piercey, D. E., Lambert, M.-C and Beaudoin, A. 2002. A strategy for mapping Canada's forest biomass with Landsat TM imagery. Page not available in IGARSS 2002, Proceedings: IEEE International Geoscience and Remote Sensing Symposium and the 24th Canadian Symposium on Remote Sensing (CD-ROM). June 24-28, 2002, Toronto, ON. IEEE, Piscataway, New Jersey. 4 p.
- Pilger, N. Peddle, D. R. and Luther, J. E. 2002. Estimation of forest cover type and structure from Landsat TM imagery using a canopy reflectance model for biomass mapping in western Newfoundland. Page not available in IGARSS 2002, Proceedings: IEEE International Geoscience and Remote Sensing Symposium and the 24th Canadian Symposium on Remote Sensing (CD-ROM). June 24-28, 2002, Toronto, ON. IEEE, Piscataway, New Jersey.
- Reese, H., Nilsson, M., Sandström, P., and Olsson, H., 2002. Applications using estimates of forest parameters derived from satellite and forest inventory data. *Comp. Electr. Agric.* 37, 37–55.
- Ung, C.-H., Beaudoin, A. Bernier, P.Y., Beaubien, J., Bérubé, J.-P., Dagnault, S., Gillis, M., Guindon, L., Lambert, M.-C. and Raulier, F. 2002. Estimating Sub-Arctic Forest Carbon Stock: Concepts, and Methods. Natural Resources Canada, Canadian Forest Service. Internal Report, 65 p.

REMOTE SENSING OF GROWTH DYNAMICS OF SITKA SPRUCE PLANTATION FORESTS IN UPLAND BRITAIN

D.N.M. Donoghue^a, P.J.Watt^a and K.B.McManus^b

^aDepartment of Geography, University of Durham, England: E-mail Danny.Donoghue@durham.ac.uk

^bCEH, Natural Environment Research Council, Monks Wood, Cambridgeshire, UK

ABSTRACT

Monitoring growth patterns over large forest estates using field survey methods is a time-consuming and costly exercise. This study paper evaluates remotely sensed data to help monitor changes in Sitka spruce (*Picea sitchensis*) plantations in Britain using satellite and airborne laser scanning imagery. The results demonstrate that low-cost satellite image data from the Landsat Enhanced Thematic Mapper (ETM+) sensor can be used to predict and map forest height and basal area characteristics with a good level of accuracy for young crops. Regression analysis of contemporaneous remote sensing data and ground based forest structural measurements is used to predict height and basal area. These models can be applied to any radiometrically normalised image of the same area to give a quantitative description of observed growth in young crops between successive images. This helps to identify and quantify areas of anomalous growth. High resolution airborne laser scanner imagery gives excellent estimation of forest characteristics but is very expensive. These data can, however, be used both to complement field measurements to improve predictions from Landsat data and to assess their quality. We conclude that retrieval of forest structure information is best achieved by the integration of satellite, airborne and ground-based measurements.

Keywords: Monitoring change, forest height estimates, Landsat, LiDAR,

1. INTRODUCTION

The ability to monitor change in woodland at regular intervals and in a cost-effective manner is becoming increasingly important for forest planners and environmental managers. Many studies have tried to quantify change in forest area, such as mapping the extent of clear-cutting, a few have tried to identify change on the basis of forest or ecological type, but very few have tried to relate change to a forest structural attribute. We describe a methodology for making rapid and low-cost assessment of Sitka spruce plantations which could be used to assist foresters check that plantations have established successfully and that growers have complied with the levels of stocking required by woodland grant schemes. Quantitative change models could also provide scientists and environmental managers with a wealth of with statistical data on forest structure that would otherwise be expensive and difficult to obtain.

Several scientific publications discuss the problems associated with deriving accurate information from multi-temporal remote sensed data (e.g. Hall et al. 1991, Olsson 1995). Examples of unwanted effects that may impact upon any change analysis include atmospheric effects, differences in illumination and observation angles, and drift in sensor radiometric quality over time. Furthermore, many of the image processing techniques used for change analysis assume that data is acquired by the same sensor, otherwise inter-sensor calibration issues must also be taken into account. Change analysis normally follows one of three paths, (1) data are corrected to physical units and atmospheric effects are removed, (2) data are radiometrically normalised to a relative radiometric scale using reference targets in an image whose reflectance is constant through time, and (3) classification of image into discrete land cover elements (say) which are then compared.

Correction of data to absolute physical units (1) is usually difficult to achieve except with active sensors such as Light Detection And Ranging (LiDAR) and interferometric Synthetic Aperture Radar (SAR) systems. With passive optical sensors it is difficult to collect data on atmospheric scattering and attenuation to allow absolute retrieval of surface reflectance. Image classification uses statistical information unique to each image to segment the data into discrete classes. Apart from problems of cloud or haze obscuring parts of an image, atmospheric or radiometric correction is not required for change analysis using classified images. However, the level of detail that can be extracted from a set of discrete classes is often very limited. Change analysis based on radiometrically normalised imaged (2) appears attractive because it does not require sensor calibration or atmospheric correction. The difficulty with this approach is the

subjectivity with which spectrally invariant targets are identified in multi-temporal images of the same area (Du et al. 2004).

While active systems such as LiDAR which make precise physical measurements offer considerable potential for forest change monitoring (Lefsky et al. 1999, Nilsson 1996), these are often impractical in terms of cost for large area survey. Furthermore, even these systems need to be validated against accurate ground based observations of forestry applications and they are expensive to acquire compared with optical satellite imagery. In this study we try to get the best of both worlds by evaluating radiometrically normalised low-cost satellite image data from the Landsat Thematic Mapper (TM) and Enhanced Thematic Mapper (ETM+) sensors with high-cost (and high quality) airborne LiDAR data and measurements from ground based sample plots.

The objective of this study is to evaluate the quality of quantitative predictions of forest change using low-cost satellite image data.

2. MATERIALS AND METHODS

2.1 STUDY AREA

The forest stands used in this study are located in Kielder Forest District, northern England and form part of the forest estate managed by the UK Forestry Commission. In these areas Sitka spruce is the dominant crop and in the Atlantic maritime climate of the UK it is a fast growing tree that is tolerant of acid and waterlogged soils. The topography consists of low undulating hills with an altitude range of 210 to 390 metres. Planting occurs on land with a mean altitude of 290 metres and a mean slope angle of 5.5°. The initial density of the plantation usually exceeds 2,500 trees per hectare (a spacing of 2x2 metres between trees and rows). These plantations are almost never thinned and so closure of the forest canopy normally occurs between 14 and 20 years after planting.

2.2 DATA

Landsat 5 TM and Landsat 7 ETM+ sensor data were used in this study because these sensors have short wave infrared (SWIR) spectral bands that correlate well with forest structural variables. The satellite data used is summarised in Table 1. The LiDAR data used in this study was acquired by the UK Environment Agency using an Optech ALTM 2033 sensor with a 1064 nm laser pulse. Part of Kielder forest was flown on 28th March 2003 at a resolution of approximately 2 points per m². Ground survey plots were selected from an analysis of the Forest Commission's sub-compartment database that contains information on species, planting date, soil type and expected yield among other variables. All data have been geo-referenced and integrated into a GIS (Geographic Information System). Ground survey plot and individual tree locations were determined using a Leica Series 300 differential GPS.

Table 1: Summary of imagery used

Sensor	Date	Resolution (m)	Quality
Landsat TM	05/7/2000	30	Cloud-free
Landsat ETM+	02/9/2002	30	Cloud-free
Landsat ETM+	01/3/2003	30	Cloud
LiDAR	28/03/2003	Density 2 hits/m ²	

3 METHODOLOGY

The LiDAR imagery is geometrically corrected as part of the post-processing procedure using differential GPS data. The geometric correction of the Landsat used 59 ground control points to project the image to the British National Grid using an affine transformation. The Root Mean Square Error was 0.86 pixels (approx. 24 m) for the Landsat ETM+. Previous studies using satellite imagery have shown that the Landsat TM and ETM+ sensor data acquired under good atmospheric conditions is of sufficiently high quality to differentiate forest stands of varying species, volume or health characteristics (Ardo 1992; Danson and Curran 1993; Dewulf et al. 1990; Puhr and Donoghue 2000). The raw digital image data used were Landsat TM level 1B and Landsat ETM+ Level 1R. This means that each channel in the multi-spectral imagery is calibrated into the same relative radiance units. However, the data has not been reduced to actual ground leaving radiance and no attempt has been made to correct for atmospheric attenuation of

signal. However, analysis and interpretation of spectral change through time does require that the data are radiometrically normalised.

Comparison of spectral response can be achieved by applying a linear radiometric normalisation function to a set of multi-temporal images. Using this technique, imagery acquired under different illumination conditions are normalised to a baseline image and quantitative data are converted into the same relative reflectance/radiance units (Hall et al. 1991, Heo and FitzHugh 2000). Factors that should be held constant (or quantified) between image data sets include sensor radiometry, topography, and interaction of radiation with ground surface (non-Lambertian reflectance). Factors that are corrected by the transformation are radiation differences due to small changes in solar zenith and azimuth angles and multiplicative atmospheric attenuation.

The Landsat data were normalised using the method described by Hall et al. (1991). This technique uses relative rather than absolute radiometric adjustment by making the assumption that pixels with the same reflectance properties should have the same DN values. This implies that the only change in the spectral signature of these pixels between two different image dates is due to linear differences in atmospheric, solar irradiance and radiometric conditions. Normalising the images will account for these differences using parameters determined empirically from bright and dark target pixels that are assumed to be spectrally invariant over time. The transformed data appear as if they have been imaged under the same atmospheric and irradiance conditions. Figure 1 shows the radiometric relationship between the 3 Landsat scenes studied. The 2002 scene is used as the reference for normalisation and the graph shows that the regression based on the spectrally invariant targets for the 2000 and 2003 scenes match well to the ideal of a $y=x$ relationship. In theory, this technique is not absolutely robust since atmospheric scattering can have additive as well as multiplicative effects. However, the results show that atmospheric attenuation through additive (offset from $y=x$) or multiplicative (change in gradient) effects do not significantly affect these data. This normalisation method can be applied to any data set in an image time series where spectral bands overlap and so SPOT 4/5 or IRS LISS 3 data could be used in any future change analysis using the same procedure.

The second stage involved locating and recording forest structure information from field sample plots and relating these data to corresponding reflectance values from the Landsat images. The Forestry Commission's database was used to identify forest compartments. That were regular in shape, contained a minimum of 80 Landsat pixels and appeared homogenous on the aerial photographs and satellite imagery. Thirteen compartments were selected and each, a minimum of two 200 m² sample plots were established.

The mean DN number of a 3x 3 window around the plot location was recorded for each stand in each reflective TM band using the methodology of Puhr & Donoghue (2000) to reduce potential errors associated with identifying the exact location of the pixel on the image. As well as recording structural forest parameters, the type and proportion of understorey vegetation present in the sample plot was also noted. This information was used as a measure of the degree of forest canopy closure. Tree height was determined from the LiDAR data by subtracting the elevation information from the first and last laser pulses. Using TerraScan software laser pulses are classified into above ground and ground surface points. Tree height is determined by subtracting points classified as canopy-top from the derived ground surface model.

3.1 MODELLING

The sample plots were used to determine the relationship between forest structural variables and reflectance. Two observations were excluded from the dataset as they contained high levels of natural regeneration and appeared as outliers that weaken the predictive ability of the model. Single and multiple band regression models were tested, with simpler single band models preferred over the more complex models for two reasons. First, the amount of variation explained by the addition of other bands did not improve the fit of the models to the ground survey data as summarised by the R^2 values. Secondly, a simple model based on a single SWIR band yields a simple model that can be understood in a physical sense and can easily be transferred to other locations. Inspection of the scatter plots for Landsat TM and ETM+ SWIR data suggests that the relationship between reflectance and forest variables is non-linear. A number of non-linear regression models were applied to the data, and a model of the type selected below best describing the relationship:

$$y = ax^{-b} \quad (1)$$

Where y is mean tree height, x is the Landsat band 7 normalised DN value and a and b are empirically derived constants.

$$\text{Mean height} = 15709 \times \text{ETM7}^{-2.57} \quad (2)$$

The relationship between mean sample plot height and the LiDAR is linear (equation 3) and so a conventional least squares linear regression model is appropriate.

$$\text{Mean height} = 0.93 \times \text{LiDAR} + 0.95 \quad (3)$$

4 RESULTS

The regressions between ground reference data, Landsat and airborne LiDAR data are summarised in Table 2. Height is strongly related to radiance in all Landsat 7 ETM+ bands with the exception of the near infrared band. The relationship between stand diameter and radiance is negative and relatively weak. Landsat ETM+ band 3 ($R^2 = 0.57$) gave the highest R^2 values. These R^2 values indicate that stand diameter cannot be predicted with any certainty using models constructed from Landsat data. The relationship between basal area and radiance is strong. The band with the highest R^2 value is Landsat ETM+ band 3 ($R^2 = 0.76$). The sharpest decrease in radiance level occurs with basal areas of 20 m²/ha -1 or more. There is no observed relationship between tree density and the Landsat data. Height shows the strongest relationship with satellite radiance data in the shortwave infrared ($R^2 = 0.80$), see Figure 2. This band will also be affected less by additive atmospheric scattering than the other Landsat bands and so it was used to predict tree height. Since historical Landsat data have been radiometrically normalised it is possible to apply the prediction retrospectively and so produce a height change image.

Table 2: Summary of models used to estimate forest parameters from satellite and LiDAR

Sensor/ band		Height	Diameter	Basal Area	Age	Density
Coefficient of determination (R^2)						
LANDSAT ETM+ (02/09/2002)						
Band 1	Blue	0.67	0.15	0.28	0.79	<i>0.16</i>
Band 2	Green	0.84	0.49	0.64	0.87	<i>0.26</i>
Band 3	Red	0.86	0.57	0.76	0.89	<i>0.25</i>
Band 4	NIR	0.60	0.18	0.23	0.78	<i>0.24</i>
Band 5	SWIR	0.80	0.39	0.52	0.83	<i>0.28</i>
Band 7	SWIR	0.80	0.49	0.57	0.83	<i>0.31</i>
LiDAR	(28/03/2003)	0.98	0.53	0.60	0.80	<i>0.49</i>

R^2 values in italics denote P values (>0.05) all remaining P values (<0.05)

The LiDAR height model is very strongly related to mean height within the sample plot ($R^2=0.98$), and not surprisingly, is also correlated with diameter, basal area and age. Figure 3 shows the LiDAR height plotted against mean sample plot height values. Having derived a prediction of height from Landsat data it is interesting to compare this with the LiDAR height model data since we know that the LiDAR agrees very well with ground measurements. The $y=x$ relationship between these data is strong ($R^2=0.68$, see figure 4), but as expected a little weaker than the $y = ax^{-b}$ power function relationship with the ground sample data ($R^2=0.80$). The results appear to suggest that satellite data can provide robust mechanism for estimating and monitoring the height of Sitka spruce stands especially during the phase of growth leading up to canopy closure. While a qualitative image of change could be produced from the raw Landsat data, the advantage of using height prediction models to generate the composite is that it can be queried directly to obtain quantitative predictions meaningful to foresters. Despite the relatively low spatial resolution of the Landsat imagery the level of detail is sufficient to identify areas of anomalous behaviour quite easily. Obvious examples of applications include the monitoring of wind damage following severe storm events where it may be difficult and expensive to assess the degree of damage over very wide areas using airborne survey. Another obvious application areas is the monitoring of compliance with the conditions attached to woodland grant schemes where funding is given to plant crops to agreed areas and levels of stocking density. The use of satellite data would allow a regulator to make a preliminary assessment of every grant scheme area by overlaying the digital boundary data on a height change image. At present such schemes are only monitored by field visits to a random sample. The use of image data would allow field survey to

be targeted primarily at sites that appeared anomalous. Where forest management data are held in a GIS, the satellite predictions of variables such as average stand height can be compared with expected growth rates by overlaying the satellite predictions with GIS or web mapping software.

From a forest management perspective a rapid method of identifying anomalous growth in newly established plantations is useful for targeting limited field resources. By combining LiDAR survey with satellite observation it is clear that the need for accurate ground measurements, which are expensive to obtain, could be substantially reduced and it would be possible to generate an empirical prediction of height based largely on carefully targeted LiDAR survey.

5. CONCLUSIONS

Landsat TM and ETM+ imagery can be used to generate accurate predictions of stand height during the period of up to canopy closure (approx age 14-20 years) for Sitka spruce plantations in the UK. The results suggest that LiDAR may be used in place of field measurements to help drive stock assessment models over very large areas of forest at a very low cost. We conclude that retrieval of forest structure information is best achieved by the integration of satellite, airborne and ground-based measurements. Through such integration of data into a GIS, foresters have a powerful tool which is capable of providing a rapid overview of a large forest estate.

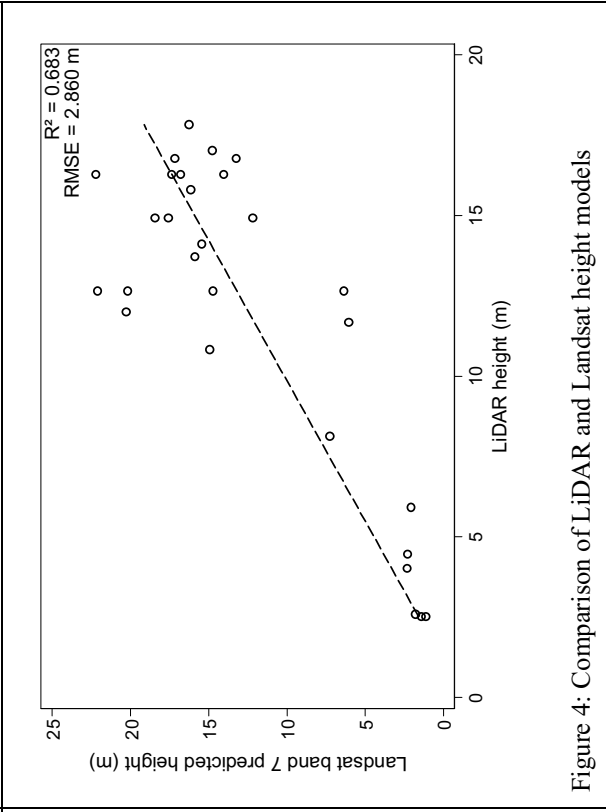
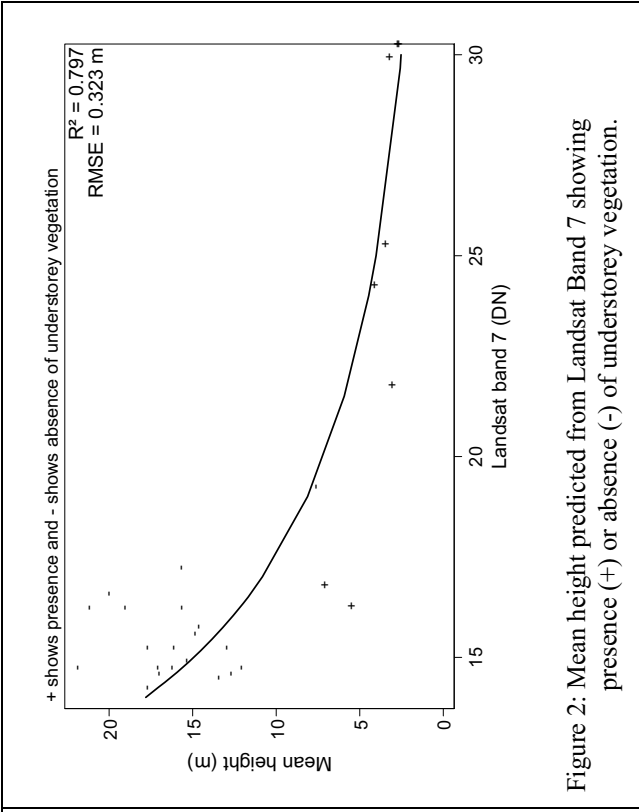
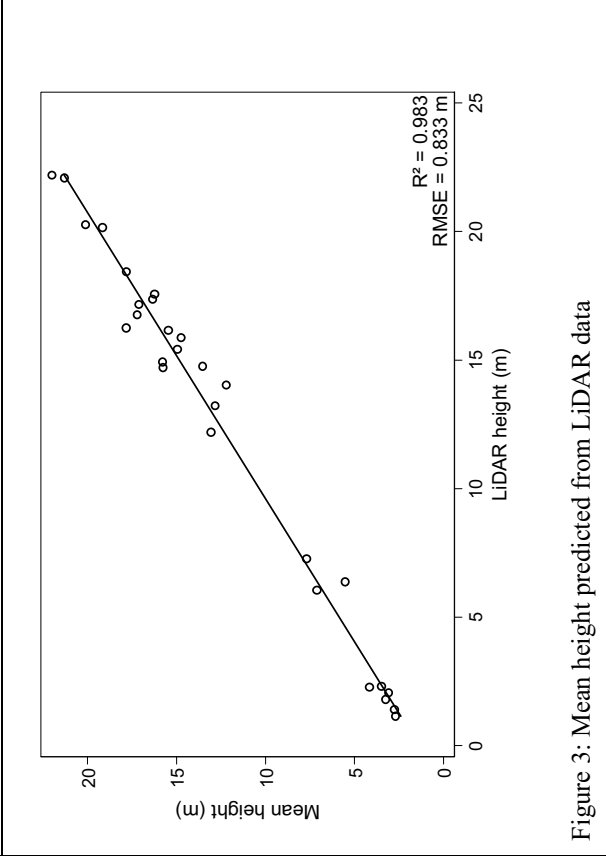
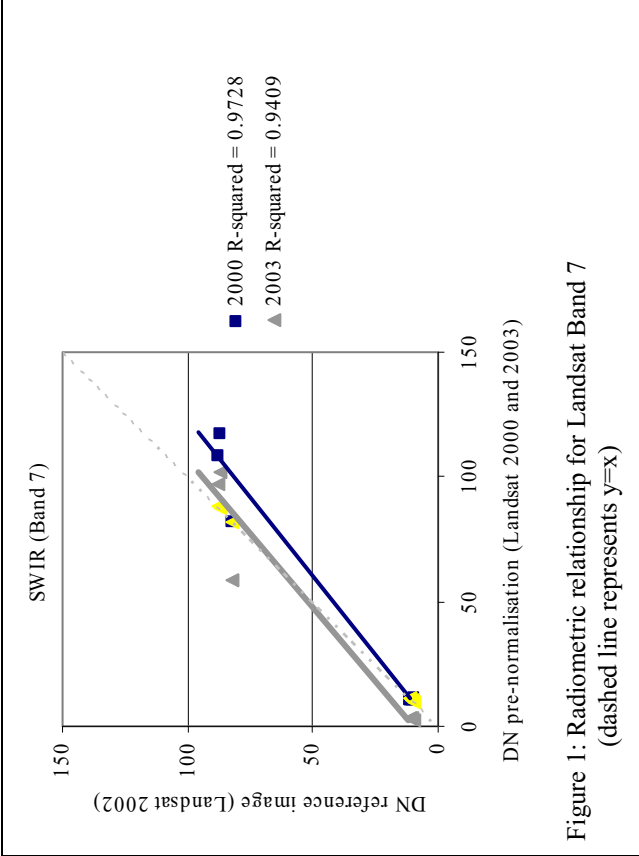
ACKNOWLEDGEMENTS

This research was carried out as part of an EU Life-Environment project ForestSAFE.

Project website: <http://www.geography.dur.ac.uk/ForestSAFE>

REFERENCES

- Ardö, J., 1992, Volume quantification of coniferous forest compartments using spectral radiance recorded by Landsat Thematic Mapper. *International Journal of Remote Sensing*, 9, 1779-1786.
- Danson, F.M., and Curran, P.J., 1993, Factors affecting the remotely sensed response of coniferous forest plantations. *Remote Sensing of Environment*, 43, 55-65.
- DeWulf, R.R., Goosens, R.E., De Roover B.P., and Borry, F.A., 1990, Extraction of forest stand parameters from SPOT-HRV data. *International Journal of Remote Sensing*, 11, 1571-1588.
- Du, Y, Teillet, P.M., and Cihlar, J. 2004, Radiometric normalization of multitemporal high-resolution satellite images with quality control for land cover change detection, *Remote Sensing of Environment*, 82, 123-134.
- Hall, F.G., Strebel, D.E., Nickeson, J.E., and Geotz, S.J., 1991, Radiometric rectification: toward a common radiometric response among multitemporal, multisensor images. *Remote Sensing of Environment*, 35, 11-27.
- Heo, J. and FitzHugh, T.W., 2000, A Standardized Radiometric Normalization Method for Change Detection Using Remotely Sensed Imagery, *Photogrammetric Engineering and Remote Sensing*.
- Lefsky, M.A., Cohen, W.B., Acker, S.A., Parker, G.G. Spies, T.A. and Harding, D., 1999, Lidar remote sensing of the canopy structure and biophysical properties of Douglas-fir western hemlock forests. *Remote Sensing of Environment*, 70, 339-361.
- Nilsson, M., 1996, Estimation of tree heights and stand volume using an airborne LIDAR system, *Remote Sensing of Environment*, 56, 1-7.
- Olsson, H., 1995, Reflectance calibration of thematic mapper data for forest change detection, *International Journal of Remote Sensing*, 16, 81-96.
- Puhr, C.B. and Donoghue, D.N.M., 2000, Remote sensing of upland conifer plantings using Landsat TM data: a case study from Galloway, south-west Scotland. *International Journal of Remote Sensing*, 21, 633-646.



DATA FUSION AND REFLECTANCE MODEL INVERSION FOR ESTIMATION OF FOREST PARAMETERS

Olle Hagner

Remote Sensing Laboratory
Department of Forest Resource Management and Geomatics
Swedish University of Agricultural Sciences
SE-901 83 Umeå, Sweden
Email: Olle.Hagner@resgeom.slu.se

ABSTRACT

Estimation of forest parameters from satellite imagery by use of the kNN method (and similar) requires large training datasets to achieve acceptable prediction accuracy. The practical limit adopted by SLU for operational use of the kNN-method for forest parameter mapping is a minimum of 500 independent samples within each satellite scene. This requirement severely limits the use of SPOT satellite imagery, especially in Northern Sweden due to the low spatial density of national forest inventory plots.

Another problem is that the prediction accuracy at the individual stand level is generally too low for operational forest management purposes. Although complementary sources of local information such as forest inventory data and stand maps might be available it is often difficult to utilise them in practice, mainly because not enough objective reference data is co-located with the local information.

Both of these problems can be addressed by a two-step approach using a forest reflectance model for absolute calibration. Alternatively, reflectance data derived from medium resolution satellite sensors (i.e. Terra MODIS and SPOT-Vegetation) can be as reference instead.

In a second step, an artificial neural network-based reflectance model inversion is used to estimate forest parameters. Data fusion capabilities can be achieved by the inclusion of forest inventory error models. An additional benefit of neural network-based estimation compared to the current kNN-method is a significant improvement in processing speed. The calibration and estimation methods are currently being implemented into an operational production line for country-wide forest parameter mapping.

Keywords: Data fusion, reflectance model inversion, forest parameter estimation, combined estimation, Landsat, SPOT, kNN, neural networks, National forest inventory.

1 INTRODUCTION

A problem with the method currently used at SLU for production of continuous forest variable maps from satellite imagery (i.e. the kNN method) is that the prediction accuracy at the individual forest stand level is too low for operational forest management (although adequate at the landscape level). The ability to also include other complementary sources of information such as existing forest inventory data and geographic features into the estimation process would be very useful. Another problem is the requirement of at least 500 plots within each scene to compute statistics. This requirement severely limits the use of SPOT satellite imagery and also in many cases Landsat scenes along the coast or mountain range have too few plots to be processed.

Neural networks have been found to be very useful for extracting forest parameters from spectral signatures and often perform significantly better than traditional techniques (Feychting et al. 1991, Kimes et al. 1998, Kimes et al. 1999).

The aim of this work was to develop a set of neural networks that are able to estimate an array of continuous forest variables from reflectance calibrated satellite imagery, latitude and elevation data. The networks should also be able to incorporate existing field inventory data from forest maps into the estimation process when available.

2 METHODS

2.1 REFLECTANCE AND FOREST DATA

A training dataset consisting of forest variables measured on 18350 national forest inventory (NFI) with co-registered reflectance signatures was compiled from seven reflectance calibrated Landsat-7 scenes. The scenes represent boreal forests in central and northern Sweden (Hagner and Olsson 2005).

2.2 SIMULATED OCULAR FIELD INVENTORY DATA

A set of simulated ocular field inventory data was compiled from the NFI-plot data. The error levels applied are shown in table 1 and correspond to the results of Ståhl (1992). A geographic site index was calculated from elevation and latitude according to formula 1.

$$\text{Site Index (dm, h100)} = 105.8 - 0.009429 * \text{elevation (meters)} - 0.00001187 * \text{Northing (RT90)} \quad (1)$$

The dataset was randomised into a training dataset consisting of 15000 plots and an evaluation dataset with 3000 plots. Also a mixed training set was generated where the simulated field inventory data was replaced with a “no data” value for 50 percent of the observations.

Table 1. Error levels used to simulate ocular field inventory data.

Variable	RMS (%)
Mean age (years, basal area weighted)	18
Mean height (meters, basal area weighted)	11
Basal area (m2/hectare)	19
Total stem volume (m3/hectare)	22
Relative species proportion (%)	8

2.3 NETWORK DESIGN AND TRAINING

Based on previous experience from similar studies a cascaded feed-forward network structure with 12 and 10 nodes arranged in two hidden layers was chosen (figure 1). Each node was connected to all nodes in the previous layers thus enabling the network to effectively deal with the parts of the problem that could be modeled with linear models. A separate network was created for each combination of input variables. The basic structure (except for the number of input nodes) was retained for all 10 networks. Each network was trained for 300 000 iterations using the back propagation paradigm to minimize the overall prediction error.

The software Neural Works Professional-II from Neural Ware Inc. was used for network development, training and evaluation. The trained networks were automatically converted into C++ code and implemented into the ER-Mapper remote sensing system.

3 RESULTS

All networks converged rapidly and produced consistent accuracies for a given set of input data. No problems with over fitting were found when comparing the performance on the training and evaluation datasets.



Figure 1. Screenshot of the NeuralWorks software during network training. The network structure is shown in the lower part with 14 input nodes in the input layer, 12 and 10 nodes in the two hidden layers and 8 output nodes representing estimated forest parameters. In this example, input nodes representing age, basal area, stem volume and broadleaf proportion are disabled. Scatter plots show the correspondence between actual and estimated values for the evaluation data set.

3.1 PREDICTION ACCURACY

As expected the accuracy increased consistently as more input information was available. The network structure could adapt to all different combinations of input data (table 2). The network that was trained on the mixed input set produced accuracy levels equivalent to the specialized networks. Compared to multiple linear regression models the neural network produced more accurate results for species proportions estimated from reflectance data alone and equivalent results for all other variables and input combinations.

Table 2. Coefficient of correlation (R%) between actual and estimated forest variables based on different input datasets. Refl. = surface reflectance of Landsat bands 1-5,7 SI=Site index, H=height, PS = proportion Pine and proportion Spruce, BA = Basal area.

	Input data									
Predicted variables	SI	Refl	Refl +SI	Refl +SI +H	Refl +SI +H +PS	Refl +SI +H +PS +Age	SI +H +PS +Age	SI +H +PS +BA	Refl +SI +H +PS +Age +BA	Simulated ocular field inventory
Site Index	52	37	65	65	68	73	72	64	73	52
Age	6	56	58	74	77	94	94	74	94	94
Height	13	73	73	92	94	94	94	95	95	92
Basal area	4	72	71	82	83	82	80	93	93	91
Stem volume	10	68	67	85	85	85	83	91	92	92
Proportion Pine	6	63	64	63	96	96	96	96	96	95
Proportion Spruce	2	55	55	53	95	95	95	95	95	95
Proportion Broadleaf	5	67	67	67	95	95	95	95	95	94

4 DISCUSSION

Rules of thumb for network design suggest that the performance could be optimised by careful pruning the number of nodes in the hidden layers. Although the network structure in this study allows for more complexity than required the relatively large training sets might have prevented the problems with over fitting that has been reported in many studies.

A very important improvement is the ability to derive meaningful estimates even when no field data is available within the scene (provided that the satellite imagery can be reflectance calibrated). However, if available, field data can still be useful for calibration of seasonal effects. The networks also provide a convenient and flexible means to utilise existing field inventory data in the estimation. The results indicate that pine/spruce proportions and mean height are the best variables to complement the reflectance data with. This is good news since LIDAR-based canopy height might soon be available as a spin-off from the production of a new national elevation dataset. Both height and species proportions are also relatively easy to measure in the field with conventional field inventory methods or by interpretation of aerial photography. It seems to be possible to derive useful estimates of site index from reflectance, latitude and elevation data.

The computational efficiency of the network-based prediction is superior compared to the currently used kNN-method. An entire Landsat scene can be processed within a few minutes whereas the typical kNN batch job runs over night on a dedicated computer.

Work is currently in progress to also develop similar networks for processing of SPOT HRV-data.

REFERENCES

- Feychting, H., Hagner, O., Nilsson, M. 1991. Estimation of forest stand characteristics using neural networks and satellite remote sensing. In: Current Advances in the use of computers in forest research, Workshop of the IUFRO working party S4.11-03, Joensuu, Finland, february 14. pp 13-17. Helsinki: Mestsäntikimustuslaitos.
- Hagner, O., and Olsson, H. 2005. An Empiric Forest Reflectance Model for Automated Calibration of Landsat data. Manuscript for the ForestSAT 2005 conference.
- Kimes, D.S., Nelsso R.F., Manry, M.T., Fung A.K., 1998. Attributes of neural networks for extracting continuous vegetation parameters from optical and radar measurements. *Int. J. Remote Sensing* 19. pp 2639-2663
- Kimes, D.S., Nelson, R.F., Salas, W.A., Skole, D.L., 1999. Mapping secondary tropical forest and forest age from SPOT HRV data. *Int. J. Remote Sensing* 20(18) pp 3625-3740
- Ståhl, G. 1992. A study on the quality of compartmentwise forest data acquired by subjective inventory methods. Report 24. Department of biometry and forest management, Swedish University of Agricultural Sciences, Sweden, 128 p.

DATA CAPTURE FOR FOREST MANAGEMENT PLANNING USING SAMPLE PLOT IMPUTATION BASED ON LASER SCANNER AND SATELLITE IMAGE DATA

J. Wallerman^a and J. Holmgren^b

Swedish University of Agricultural Sciences
Remote Sensing Laboratory
SE 901 83 UMEÅ, Sweden

^ajorgen.wallerman@resgeom.slu.se

^bjohan.holmgren@resgeom.slu.se

ABSTRACT

The next generation of long-term forest management planning systems i.e. the Swedish Heureka system, requires single-tree data of each forest stand in the register. This article reports early results from a remote sensing aided data capture method aiming to provide a complete forest stand database with single tree information. The method applied is canonical correlation-based imputation of sample-plot forest data using laser scanning and medium-resolution satellite image data as the link (carrier data) between predicted point and the most similar surveyed plot. Laser scanning data contains detailed information of tree canopy height and structure, information expected very useful for imputation purposes. There may also be an advantage in combining this with medium resolution optical image data, to possibly enhance tree species discrimination. Performance of the method is evaluated on a large sample of 10 m radius field plots surveyed in south-western Sweden, optical image data (SPOT 5 HRG satellite sensor), and LIDAR data from the airborne TopEye system. The aim is to predict single-tree features of forest stands (0.5 – 20 ha), preserving natural multivariate variance and covariance. Imputation based on both remote sensing data sources produced estimation accuracy (root mean square error, in percent of true mean) of 20% for mean stand volume per hectare and 22% for stand mean stem density. Mean tree diameter, height and volume were estimated with 17, 12, and 31% accuracy, respectively. The addition of SPOT data to laser scanning data only slightly improved estimation of tree species proportions.

Keywords: Laser scanning, forest management, sample-plot imputation.

1 INTRODUCTION

The recent development in forest management planning systems has increased the need for detailed data available for each forest stand (for example, the HEUREKA long term planning application). New systems operate on single trees as the calculation unit, in contrast to most previous systems operating on summary characteristics of complete stands. Furthermore, it may not be necessary to provide estimates of all present trees in each stand, a representative list of trees is sufficient. It is important, though, that the list represent the natural multivariate dependence and variation present in the actual forest data. There are several novel remote sensing systems and methods potentially well suited for single-tree data capture, especially airborne laser scanning (Nilsson, 1997; Holmgren, 2003). This technique produces a dense spatial grid of highly accurate measurements of tree canopy height. Laser scanning data has shown a large potential, for example to estimate tree height (10-11% RMSE) and stem volume (22-26 % RMSE) of 10 m radius field plots (Holmgren *et al.*, 2003), automatic detection of trees, as well as estimation of mean stem volume of small stands (16% RMSE) using ordinary least-square regression. On the other hand, a set of dense measurements of canopy height do not necessarily contain much information of tree species, a very important issue in management planning. Thus, there may be an advantage to combine laser scanner data with optical image data in order to add information of tree species.

Previous studies propose imputation of available field sampled data, such as sample plots or surveyed stands (for example, Moeur and Stage, 1995; Ohmann and Gregory, 2002; Holmström, 2001; Temesgen *et al.*, 2003), as a method to estimate a list of trees for unsampled locations. Imputation is a non-parametric estimation method where, given carrier data (such as remote sensing data) of all observations in a dataset, primary data (such as a list of trees for a particular stand) missing for one observation is imputed from the most similar observation available (reference observations), where similarity is measured in the carrier data space. This approach automatically produces estimates with preserved multivariate dependency of the

features of each estimated observation. The over-all variation is not automatically preserved, though, it is determined by the information quality of the carrier data. To utilize several carrier data sources, with different information content of the primary data, it is recommended to apply a weighting of the secondary data sources to use the data efficiently (Moeur and Stage, 1995; Holmström, 2001; Ohmann and Gregory, 2002; Holmström and Fransson, 2003). Generally, a transformation of the carrier data space based on the linear relationship to the primary data is suggested, such as the canonical correlation transformation (Moeur and Stage, 1995; Temesgen *et al.*, 2003). This is, under some conditions, equivalent to a multivariate regression transformation (Moeur and Stage, 1995). Univariate regression transformation has been applied, using the kNN estimation method, to utilize the combination of aerial photo interpretations and stand register data (Holmström, 2001) as well as satellite image data (SPOT XS) and airborne radar data (CARABAS-II VHF SAR) (Holmström and Fransson, 2003). The former study showed a moderate benefit (10% lower RMSE of predictions of stem volume and age) of using the regression transformation compared to using Euclidean and Mahalanobis distance. The latter reports a clear benefit from using the regression transformation.

The aim of this work is to evaluate the performance of sample plot imputation based on airborne laser scanner data (TopEye), medium resolution satellite image data (SPOT XS), and the combination of both sensors, as a method to estimate single-tree, as well as over-all average characteristics, of forest stands.

2 MATERIAL AND METHODS

2.1 DATA

The data used in the study was collected at the Remningstorp estate in the south of Sweden (lat. 58°30'N, long. 13°40'E). The estate is privately owned and dominated by Scots pine (*Pinus sylvestris*), Norway spruce (*Picea abies*), and birch (*Betula* spp.).

Field data was collected by surveys of 10 m radius plots using the methods and models in the Forest Management Planning Package (Jonsson *et al.*, 1993). This includes callipering all trees with diameter at breast-height larger than 0.05 m on each plot, as well as recording additional measurements, such as height and age, on a sub-sample of trees. The height of each remaining tree was estimated using functions developed by Söderberg (1992) relating tree height to diameter as well as other, less influential, variables, and a local correction for systematic errors. The estimation functions show standard deviations of 11, 13, and 15 % for Scots pine, Norway spruce and birch, respectively (Söderberg, 1992). The position of each plot center was measured using averaging of differentially corrected GPS measurements, either using a local reference station in real-time producing sub-meter accuracy, or post-processing.

The field plots were measured in 1998 to 2003 using two different sampling allocations, not specifically designed for this study. First, a stratified random sample of forest stands was surveyed using a randomly allocated systematic grid of 8-10 plots for each stand. The stratification was made on the stand registry stem volume and aimed to sample fewer young stands than old. Secondly, a cluster of 4 by 4 adjacent plots was surveyed in the central parts of 16 stands. The state of each plot was forecasted to the year of 2003 in order to match the acquisition data of the remote sensing datasets. After removal of stands treated after the date of survey and plots with obvious position errors, 72 stands and 1087 plots were left.

Satellite image data was extracted at each plot center, from a SPOT 5 XS scene, acquired 10:05 PM on the 3rd of June 2003 and geometrically precision corrected. The green, red, near-infrared and mid-infrared bands were used.

Laser scanner data was acquired by the TopEye system at the 9th of August 2003 at a flight altitude of 130 m and resulted in 1.5 – 2.0 pulses / m². For each field plot, the laser measurements were classified in two classes, measurements of the ground level and measurements of the vegetation. For each measurement in the latter class, canopy height was determined using a local estimate of the ground level based on the first class of measurements and an empirical threshold to reduce the influence of shrubs and non-tree vegetation. For each plot, the 95th percentile, the variance and mean of the measurements classified as vegetation measurements were computed, as well as the proportion of vegetation measurements and all measurements.

2.2 IMPUTATION

Two measures of sample plot similarity, measured in carrier data space, were applied in parallel, the Mahalanobis distance (Krzanowski, 2000) and a canonical correlation approach (Moeur and Stage, 1995;

Krzanowski, 2000). Applying imputation based on the Mahalanobis distance conforms to rescaling the distances between plots to take in account the different scale (variance) of each carrier data variable as well as the covariance between the variables. In essence, this strategy results in weighting the information content in each variable equally (including less weight to each of two strongly correlated variables to represent the inherent common information content only once). The aim of imputation based on canonical correlation transformation is to weigh each carrier data variable optimally with respect to the primary variables. Imputation is then based on new carrier data features which explains the most of the field data. First, a canonical correlation analysis is performed using the n reference plots where p primary data variables, $\mathbf{X}$ ($n \times p$), and q carrier data variables, $\mathbf{Y}$ ($n \times q$), are available. The result is a set of pairwise new variables, canonical variates, where each pair consists of one variable being a linear transformation of $\mathbf{X}$, and one variable being a linear transformation of $\mathbf{Y}$. The transformation is made in order to maximize the correlation between the two variables in each canonical variate (Krzanowski, 2000). The canonical variates are usually ordered by decreasing magnitude of correlation. Then, secondly, imputation is made based on a distance measure constructed by the canonical transformation of $\mathbf{Y}$, by weighting each canonical variate dimension by the square of the corresponding correlation. This is, under some constraints, equivalent to measure distance in the space of $\mathbf{Y}$ transformed by multivariate regression of $\mathbf{Y}$ on $\mathbf{X}$ (Moeur and Stage, 1995). Furthermore, Moeur and Stage (1995) gained efficiency by selecting only those canonical variates which corresponds to significant canonical correlations (Moeur and Stage, 1995). Here, this was tested using a likelihood ratio test (Krzanowski, 2000).

Canonical correlation analysis was made by relating survey plot summary statistics (basal area, stem volume, number of trees, basal area proportion of pine, spruce and deciduous trees, mean and standard deviation of diameter, height and age) to the varying remote sensing measures of each data source combination.

2.3 EVALUATION

Imputation results were evaluated using cross-validation of stands, an approach proposed by Moeur and Stage (1995). Here, 30 plots were imputed to each stand using all plot data as reference data, excluding plots surveyed in the imputed stand. Errors were calculated by comparing imputed and field sampled tree lists of each stand. Especially, root mean-square error (RMSE) of mean estimations and the ratio of standard deviations (imputed/measured) were calculated. Furthermore, a measure of tree distribution similarity,

$$C = \left(\sum_{j=1}^{10} m_j \frac{n}{m} - n_j \right) / n \quad (1)$$

inspired by the χ^2 test, was also calculated, where m is the total number of imputed trees in the stand, n is the total number of measured trees in the stand, m_j is the number of imputed observations in class j , and n_j is the number of field sampled observations in class j . The classes were determined as the 10-percentile classes of the field sampled trees. Errors in distribution were determined for tree diameter, height, and volume. RMSE accuracy of estimations of volume per hectare, number of stems per hectare, and species composition (based on basal area proportion) were also calculated. Statistical computations were made using R, a language and environment for statistical computing (R Development Core Team, 2004). Especially, the *cancor* function was applied for canonical correlation calculations.

3 RESULTS AND DISCUSSION

All canonical variates were significant ($p < 0.05$) in all three canonical correlation analyses. Thus, all variates were used in the transformation of $\mathbf{Y}$. The correlation of the last canonical variates were low, though. For example, correlation ranged between 0.97 to 0.12 for the combination of SPOT XS and laser scanner data.

Imputation based on laser scanner data were, in terms of stand volume (m^3/ha), as accurate as subjective field estimations without support of ground measurements (Ståhl, 1992) (Table 1). For imputations based on laser data, only and in combination with SPOT data, the errors of distribution estimations were approximately equal for all variables, measured by the C statistic as well as the ratio of standard deviations (Table 2). Especially, the ratio of standard deviations for imputations based on SPOT data only show the height and volume distributions to be less accurately estimated than the diameter distribution. Generally, all imputations overestimated the variance of the distributions heavily (Table 2).

The canonical correlation (CC) approach performed better than imputations made using the Mahalanobis distance (MH) (Tables 1 and 2). Given laser scanning data, addition of SPOT data tended to reduce the performance of MH imputation, and increase the performance of CC imputation, although the differences were small. The addition of SPOT data did improve estimation of tree species composition (Pine, Spruce, and Deciduous trees), although only slightly, but the over-all accuracy was very low. Mean tree volume was estimated least accurate, compared to mean tree diameter and height (Table 2).

Table 1. Results from estimating tree species proportions and stem density in forest stands using remote sensing based sample-plot imputation. In parenthesis are RMSE in percent of field measured mean. Bias is defined as estimated – true value

Distance measure	Sensor combination		Tree species proportion RMSE			Density [st/ha]		Mean volume [m ³ /ha]	
	SPOT	laser	Pine	Spruce	Dec.	RMSE	Bias	RMSE	Bias
CC	x		0.43 (148)	0.47 (94)	0.40 (207)	409 (36)	5.5	98 (38)	-13
CC		x	0.36 (267)	0.43 (55)	0.27 (357)	284 (25)	-53	50 (20)	-8.3
CC	x	x	0.42 (149)	0.51 (95)	0.36 (235)	250 (22)	-45	51 (20)	-11
MH	x		0.44 (168)	0.48 (86)	0.36 (234)	408 (36)	2.0	100 (39)	-9.0
MH		x	0.36 (221)	0.42 (60)	0.30 (291)	266 (23)	-57	52.4 (21)	-8.6
MH	x	x	0.49 (165)	0.51 (96)	0.35 (236)	254 (22)	-50	59.9 (23)	-2.9

Table 2. Results of distribution estimation: The C statistic, RMSE of estimating mean (in percent of true mean in parenthesis) and the ratio of imputed and field measured standard deviation, R_{std}

Distance measure	Sensor combination		Diameter			Height			Volume		
	SPOT	laser	C	RMSE	R_{std}	C	RMSE	R_{std}	C	RMSE	R_{std}
CC	x		4.9	6.7 (34)	1.6	5.3	4.4 (28)	2.0	5.0	0.29 (69)	2.5
CC		x	4.1	3.4 (17)	1.3	4.2	1.9 (12)	1.6	4.1	0.13 (32)	1.3
CC	x	x	4.1	3.4 (17)	1.4	4.2	1.8 (12)	1.6	4.1	0.13 (31)	1.3
MH	x		5.0	6.9 (35)	1.7	5.4	4.5 (29)	2.1	5.0	0.30 (71)	2.8
MH		x	4.1	3.5 (18)	1.4	4.2	1.9 (12)	1.7	4.1	0.14 (35)	1.3
MH	x	x	4.2	3.8 (19)	1.5	4.4	2.2 (14)	1.8	4.2	0.16 (39)	1.9

There is clearly a potential in laser scanning data as data source for new systems of long-term forest management planning. Several issues needs to be resolved, though. The accuracy of tree species proportion estimations presented in this paper are not sufficient. This issue may be resolved by using image data from the Z/I DMC digital camera in combination with template matching, an approach which showed tree species classification accuracy in the range of 60-80% (Olofsson *et al.*, 2005). Furthermore, over-estimation of the tree list variance is a serious problem. Unfortunately, this is probably an effect of imputation error variance (similar to residual variance in OLS regression modeling), which cannot easily be removed by other than more informative carrier data or a more efficient imputation method. There may be a possibility to increase the performance of the canonical correlation method, though, by linearizing the multivariate relationships in the data.

ACKNOWLEDGEMENTS

This work was performed and financed within the HEUREKA research program at SLU (<http://heureka.slu.se>). The Remningstorp field data surveys were exclusively financed by the foundation Hildur and Sven Winquists Stiftelse. Special gratitude is directed to Assistant Professor Sören Holm who provided essential statistical consultation for the paper.

REFERENCES

- Nilsson, M. 1997. Estimation of forest variables using satellite image data and airborne Lidar. Doctoral thesis. Acta Universitatis Agriculturae Sueciae, Silvestria 17. Swedish University of Agricultural Sciences, Department of Forest Resource Management and Geomatics. Umeå, Sweden.
- Holmgren, J. 2003. Estimation of forest variables using airborne laser scanning. Doctoral thesis. Acta Universitatis Agriculturae Sueciae, Silvestria 278. Swedish University of Agricultural Sciences, Department of Forest Resource Management and Geomatics. Umeå, Sweden.
- Holmgren, J., Nilsson, M. & Olsson, H. 2003. Estimation of tree height and stem volume on plots using airborne laser scanning. *Forest Science* 49 (3): 419-428.
- Moeur, M. & Stage, A.R. 1995. Most similar neighbor: an improved sampling inference procedure for natural resource planning. *Forest science* 41(2): 337-359.
- Ohmann, J.L., and Gregory, M.J. 2002. Predictive mapping of forest composition and structure with direct gradient analysis and nearest-neighbor imputation in coastal Oregon, U.S.A. *Canadian journal of forest research* 32: 725-741.
- Holmström, H. 2001. Data acquisition for forestry planning by remote sensing based sample plot imputation. Doctoral thesis. Acta Universitatis Agriculturae Sueciae, Silvestria 201. Swedish University of Agricultural Sciences, Department of Forest Resource Management and Geomatics. Umeå, Sweden.
- Temesgen, H., LeMay, V.M., Froese, K.L., and Marshall, P.L. 2003. Imputing tree-lists from aerial attributes for complex stands of south-eastern British Columbia. *Forest ecology and management* 177: 277-285.
- Holmström, H. & Fransson, J.E.S. 2003. Combining remotely sensed optical and radar data in kNN estimation of forest variables. *Forest Science* 49 (3): 409-418.
- Jonsson, B., Jacobsson, J. & Kallur, H. 1993. The forest management planning package. Theory and application. *Studia Forestalia Suecica* 189, 1-56.
- Söderberg, U. 1992. Functions for forest management, height, form height and bark thickness of individual trees. Department of Forest Survey, Swedish University of Agricultural Sciences, Report 52. Umeå, Sweden.
- Krzanowski, W.J. 2000. Principles of multivariate analysis. A user's perspective. Revised edition. Oxford University Press, New York, pp. 432-448.
- R Development Core Team, 2004. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org>.
- Ståhl, G. 1992. A study on the quality of compartmentwise forest data acquired by subjective inventory methods. Licentiate thesis. Report 24. Swedish University of Agricultural Sciences, Department of Biometry and Forest Management. Umeå, Sweden. (In Swedish with English summary).
- Olofsson, K., Wallerman, J., Holmgren, J. and Olsson, H. 2005. Tree species discrimination using Z/I DMC imagery and template matching of single trees. Manuscript.

Session 5b

EMPOWERED FOREST GIS – CROSS VALIDATION OF SRTM AND OPTICAL DATA

Piotr Wezyk¹, Roeland de Kok², Andrzej Swiader³

¹Lab. of GIS and RS, Faculty of Forestry, Agricultural University of Cracow, Poland

²landConsult; www.landconsult.de

³ProGea Consulting; www.progea.pl

ABSTRACT

The radar surface model ITED (Interferometric Terrain Elevation Data; SRTM) in combination with the differentiated forest mask can be used for determination of forest stand conversion as well as mayor changes of the stand structure. The data fusion of radar information, multi-temporal optical data as well as forest GIS information, allows the calibration of seamless continental coverage of RS data. The study demonstrates that the basic strategy developed in FOREMMS (IST 5FP) and it's continuity in GMES is focussed on extending forest information throughout different scale levels (after Häusler, et al, 1996).

Keywords: SRTM, Forest Stand Height, ETM+, Object Oriented, Texture.

1 INTRODUCTION

In this study the relationship of the SRTM (ITED) and local DSM (Digital Surface Model) derived from DEM and forest GIS data is used to correlate the height information of the forest development stage to radar height information. The development stage of forest has a complex relationship with the height of the forest stand in radar images (Kelldorfer, 2004). Especially in more mature stands, the variance of the canopy structure causes deviations in the height information of the SRTM model. The optical image information (texture, Landsat ETM+) of the forest stand evaluated for coherent groups of pixels (image objects) contains important information about forest cover (Burger, 1996, Wezyk et al, 2004) Although the texture information has been used to establish a forest mask from ETM+ imagery, the bio-physical factors underlying the typical response of forest cover towards optical data reveals a variety of forest parameters. In this study the standard deviation of the red light is evaluated for its link to the development stage of the forest (young, maturing, old). The cross calibration of radar and optical data works both ways. The age-differentiation from optical data allows for correction of radar height information and radar height information confirms stand maturity and highlights forest status-quo, forest conversion as well as stand changes. Besides the cross calibration of sensor data, the complete GIS analysis manifolds its potential (PowerGIS).

2 RESEARCH AREA

The area of Niepolomice Forest (http://argis.les.ar.krakow.pl/IWS_page/eng) near Cracow was selected because of large geodata resources collected during a variety of projects, including FOREMMS IST 5FP (Wezyk, 2000). Additionally a new Forest Management Plan and forest digital maps were completed for the Niepolomice Forest District area in the year 2002, according to a newly introduced standard in State Forests [Directive no. 21, 2004].

Forest areas (about 10 500 ha) of the main complex of old Niepolomice Forest (Fig.1) make dense forest stands with clearly marked division for three different tree stand types. Three northern forest parts (Kolo, Grobla, Grobelczyk – about 2400 ha) are remains of native oak-hornbeam forests (70% share of oaks) and differentiate definitely from the main coniferous complex (72% - pine tree stand) which is heavily transformed by anthropological influences. The Niepolomice Forest area was selected as a European Ecological Nature 2000 Network area.

3 INITIAL GIS ANALYSIS

Stand height is collected during forest inventory campaigns and registered in the SILP database (http://www.zilp.pl/z/test01/dzialalnosc/zasoby_silp/). The update has taken place during the timeframe in which the SRTM mission was accomplished. This allows for a comparative analysis.

3.1 PREPARATIONS

A forest GIS analysis was conducted, which resulted in thematic map of over 2200 forest sub compartments (polygons) containing stand-height. The value of the dominant species for “height” was taken from SILP. The ArcView 3.2 software was used to convert vector to raster files. As the result, the raster (GRID) file was created with objects (tree stands) height, also called THM (Tree stands Height Model). The effective timeframe of this information (2001) is covering the acquisition time of ITED (Flight in February 2002). By performing analysis on the raster layers, a layer with the sum of DEM and THM was created. The result layer was named Virtual Surface Model (VSM, Fig.2). It approaches object height for agricultural and forest areas, however without information about height in villages and the Niepolomice settlement.

3.2 THE DIFFERENCE IMAGE

After these preparations, the raster layers were loaded into an algorithm in ER Mapper, which subtracted DEM and VSM from SRTM. The resulting image of SRTM and VSM differences was classified into 10 classes (Fig.1).

In ArcView 3.2a (ESRI) vector SHP files were made, representing:

- border of forest complexes of Niepolomice Forest District,
- Main complex border (species composition: 79% coniferous; 21% broadleaved stands),
- Broadleaved complexes: Kolo, Grobla and Grobelczyk (see Fig. 1, 97,5% broadleaved, 2,5% coniferous),
 - forest areas outside State Forests administration and others remaining,
 - three classes of pine stands with different age class: 15, 55 and 150 years.

The polygons were imported to ERV format (native vector format of ER Mapper software) and were used for selecting so called “raster regions”, inside which statistics (minimum, maximum, mean, standard deviation, and median) were calculated. To estimate the SRTM model accuracy, a raster image containing three bands representing models: SRTM, DEM and VSM were prepared. This file was used to generate adequate profiles.

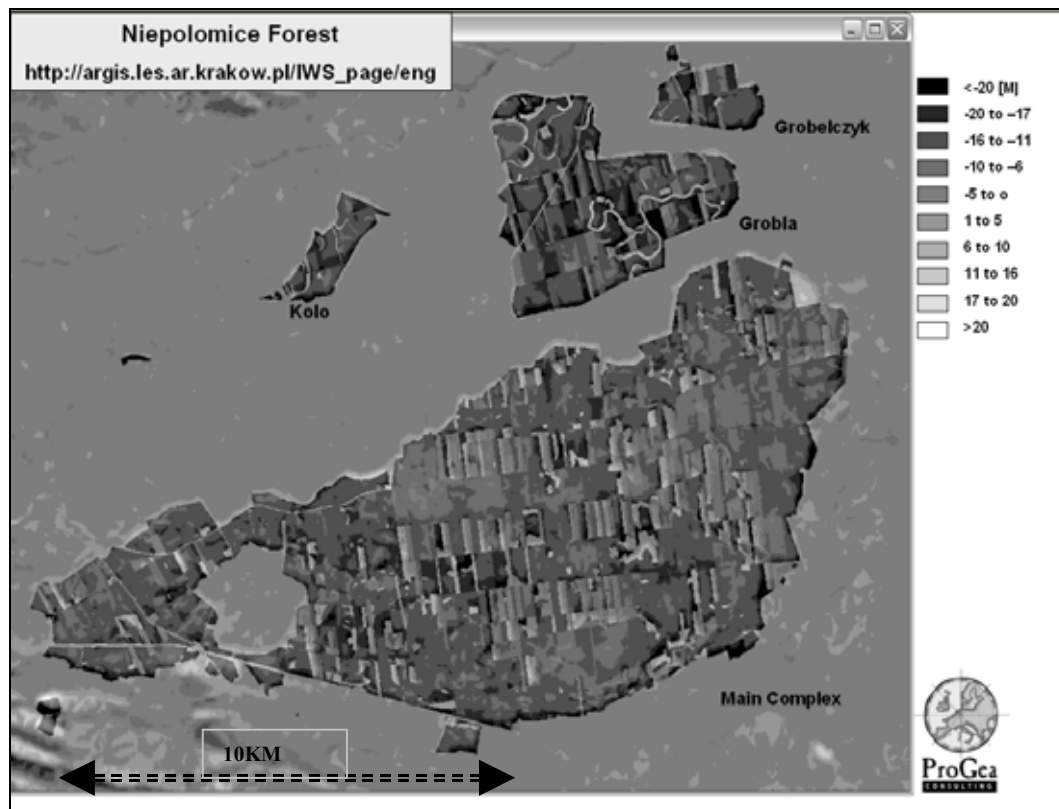


Figure 1. The difference image from SRTM-VSM., the basis for Profiles of Figure 2.

3.3 PROFILE ANALYSIS

For three selected classes, a profile (Fig 2.) was generated to demonstrate the details of Figure.1. The Height information of the VSM is a generalization of the DEM created by geodetic measurements in the fields and the dominant stand height derived from the official SILP data. The exact distance between measured points of the DEM and SILP are not known, but the representation of the total VSM can here be regarded as best practice available through terrestrial data acquisition. The SRTM with a 90 meter grid is less dense than the VSM information. For the SRTM 30 meter GRID or better Radar-Data, an improvement in the VSM is required. For the study purpose, the generalization of the VSM is still containing denser information than the 90 meter SRTM data and therefore comparable. The profiles are showing a preferential selection based upon visual analysis of the Figure 1. The Profiles highlight the tendency of radar height information to approach the correct height for young forest and deviates consequently with lower height values over maturing and old stands.

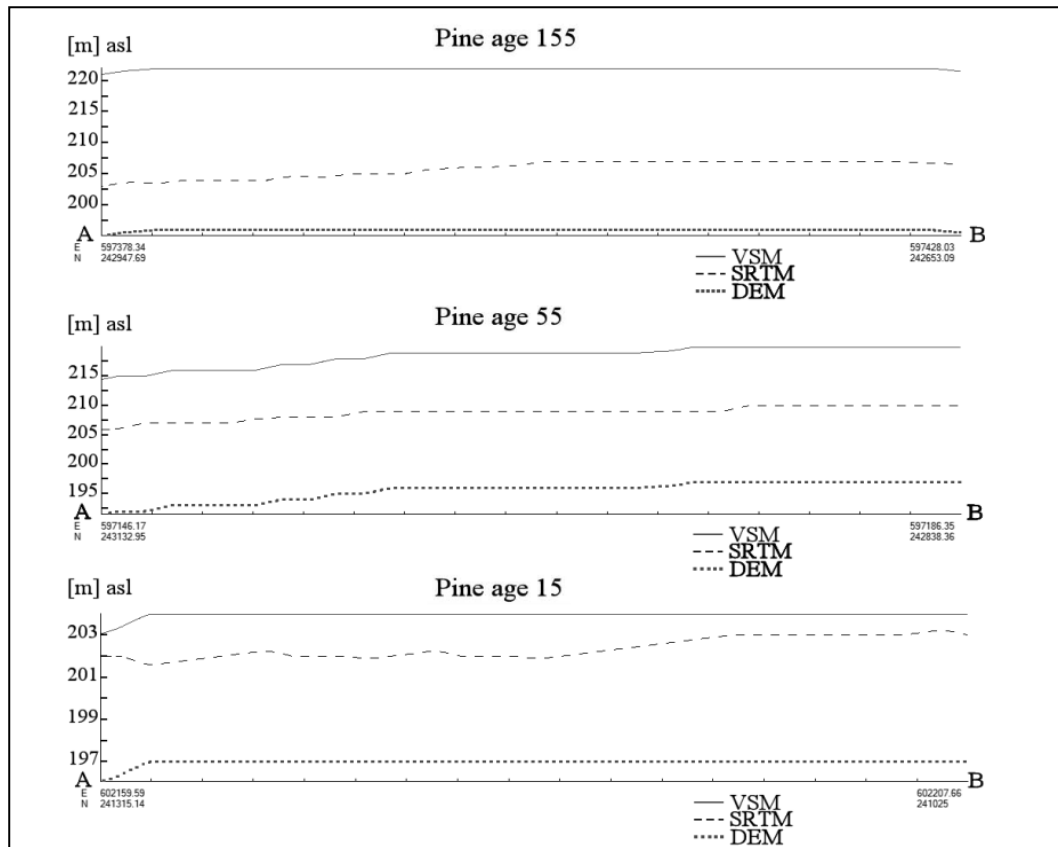


Figure 2. Selective profiles demonstrate the ‘tendency’ of the behaviour of radar height information over forest cover.

4 OBJECT ANALYSES

Two Landsat ETM scenes from the Niepolomice area were selected (7th May 2000, 7th August 2001) and segmentation was applied using the eCognition software (version 4.06). The segments cover homogeneous crown complexes inside each forest stand on average of 0.75 Ha. For around 12340 objects, the standard deviation of the red bands were selected. Because the standard deviation increases with increasing area in RS imagery, normalization over area was applied. The working hypothesis expects the standard deviation of the red band to be decreasing with increasing forest development age. As stand age and height are correlated, the standard deviation of the red band normalized for area was plotted against the average stand height. The 12340 forest segments are subsets of the 2200 forest stands with uniform stand treatment registered in the THM. The assignment of the stand height to all subsets introduces errors for those subsets which deviate from the dominant stand height of each plot. However, the majority of all segments can be considered as belonging to the dominant part of each stand and receiving correct height information from the THM. Due to the large population of segments, the deviating segments (‘gaps’ and juvenile plots inside mature stands) are tolerated and expected to have little effect on the final outcome of the linear regression but effect the visualization (Fig. 3A,B). A further check on the mean and standard deviation values for several height categories confirms that the median per height category is around the position of the line-equation. The figures 3A and 3B are showing the linear regression result for the ETM+ scene 2000 and 2001 for coniferous forest (7627 segments). Both figures show decreasing values with increasing height. For the May 2000 image the linear expression is $Y = -0,0018X + 0,081$ (coniferous) and $Y = -0,0015X + 0,0802$ (broadleaved) and for the august 2001 image $Y = -0,009X + 0,05$ (coniferous) and $Y = -0,009X + 0,021$ (broadleaved).

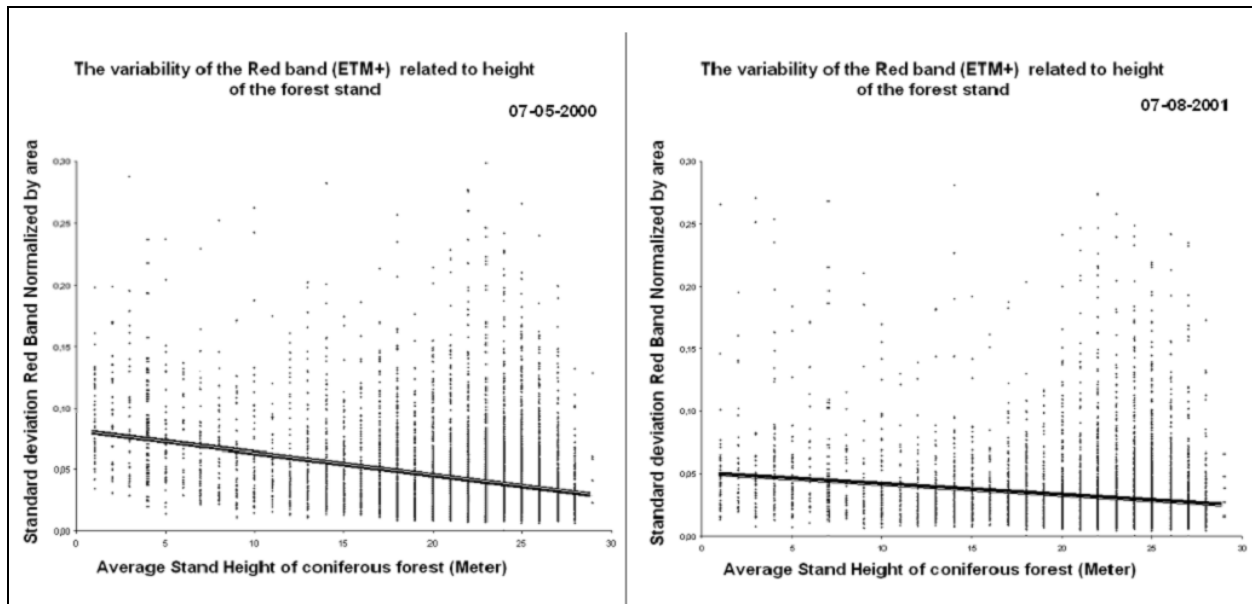


Figure 3. A and B. A selection of the linear regression for 7627 crown-cover plots in the coniferous forest, plotting normalized standard deviation of the red band versus stand height from 2 different ETM+ acquisition dates (3A,07-05-00 and 3B,07-07-01).

5 DISCUSSION

Although the 90 meter grid might seem to be too coarse for a proper registration of the height of forest stands, the initial visualization is surprisingly close to the artificial THM. The north-south oriented strips of young and old forest can be recognized in the radar image as well. The sequence of old-middle-young stands in a west to east order corresponds with the SRTM height model. The difference image of Figure 1 is showing extremes where in the field, real forest features can be found responsible. If strips cutting takes place over 100 meters, which is a ‘small’ strip, the radar returns a proper indication of height differentiation between neighbouring stand. Private forests contain a 0 height value because they are not part of the SILP related THM. Here the radar shows correct forest stands. The radar height information can thus function to confirm existence of forest over corresponding uncertainties in ETM+ classifications, if official forest GIS does not show any data. Further, a change from mature forest cover to deforestation should be noticed immediately by extreme decreasing height information. Important changes for forest are ‘catastrophic’ changes. Where the standing stock collapses due to storm, fire or cutting. These changes will appear in the extremes of the difference image.

The young stands will return a radar signal which gives a height value that is very close to the values measured in the field campaign. The deviation in older stands is a consequent lower height value for the radar image compared to the THM. Knowing the influence of the altering crown canopy, the compensation values for calibration of the surface model can be linked to the radar height model. The estimation of existing terrain-relief below forest canopy could be a difficult task. However if over young stands the height is returned correct, the younger stands can represent the terrain relief for the forest complex. The radar image itself will contain very little information about the development age of the forest stand. For this reason a link to optical EO information would give that opportunity.

Effective use of the red light by vegetation would result in a near zero albedo for active vegetation. The reason of dark appearance of forest in the red band is also due to shadow especially in Landsat data type. A red-band pixel over forest in Landsat ETM+ is always a mixed pixel. The pixel contains a combination of crowns and shadows, both reducing the red band albedo values. The amount of mixture and the effect of broadleaved and coniferous forest however are different towards the red light. When a simple texture measurement takes place, the differences between coniferous and broadleaved canopies starts to diminish. The standard deviation of a canopy mosaic part, in this case around .0.75 Ha average, shows the lowest values for very old forest for broadleaved as well as coniferous stands. This leads to the possibility to estimate the development age for both broadleaved and coniferous forest. By applying linear regression over 12340 parts of forest-plots in relationship with height, the values are showing a large spread. However at the extremes the information shows that very low stands height <5meter does not contain low standard deviation values for red light. This could be caused by the lack of the shadow component. Old closed

stands do almost not show high values for the standard deviation of the red light. In catastrophic events where old crown cover disappears, the standard deviation of the red is expected to increase. The study shows an initial test case with a promising strategy. However a more detailed study concept is required. For a better calibration of the relationship it is possible to confirm the height of each mosaic part (image object) in a field campaign. In this study, the complete forest stand of several hectares is taken for height information and assigned to each of the forest mosaic parts. Moreover a selection of ideal representative stands can be taken at forehand to improve the representative standard deviation measurements. Although both SRTM and ETM+ information might still be coarse for the task at hand, for offering detailed stand information, catastrophic changes have such an impact on the canopy signature in both sensor types that their combined use is recommended for further automatic analysis over large areas and even continental coverage.

6 CONCLUDING REMARKS

A complete map on forest cover for Europe is available in the CORINE land cover map. The map is based upon RS characteristics of the early 90th. After more then 10 years of technical developments in satellite imagery and other RS fields, it is at the time to incorporate this knowledge in new forest cover mapping of Europe. Very cheap continental coverage of Radar and Multi Spectral imagery of high quality is a treasure grove in itself and should not be left unexploited. Besides the simple categories of Coniferous, Broadleaved and Mixed Forest, information about age and stand-height are crucial in establishing a correct overview. The quality of 'Old Growth' forest and its detection and monitoring should not be overlooked on the old continent where forest plantations are dominant. For the important land cover class 'forest'. It is not likely that age and correct stand height are impossible and impractical factors to be retrieved. The technique of the mapping process still requires a lot of attention and the presented technique is really developed on a shoe-string budget. It is not a simple task but it can be done. The combination of radar and Landsat data in this study are showing the direction which such mapping procedures have to take, to arrive at continental coverage results. Important as well is to use RS information on a continuous calibration base. RS imagery from Landsat type such as ASTER data as well as new Radar programs make it likely that continental coverage can be achieved on a very regular even seasonal interval. Thus the GIS information can be under constant automatic scrutiny for major changes in forest cover. This constant and direct updating of the land-cover GIS systems creates the conditions for 'Power-GIS' where automatic detection is empowered to initialize the automatic map-update.

REFERENCES

- Burger, H., Steinwender, J. (1996): Study of forest mask generation from satellite images using image segmentation algorithms. In: Kennedy P.J. (Ed): Application of remote sensing in European forest monitoring. Proceedings of the International Workshop. ISRS LI. BOKU, Vienna, 14-16th October '96, pp. 465-476.
- Häusler, T., Deshayes, M. (1996): Change detection & forest condition. In: Kennedy P.J. (Ed): Application of remote sensing in European forest monitoring. Proceedings of the International Workshop. ISRS LI. BOKU, Vienna, 14-16th October '96, pp.535-537.
- Josef M. Kellndorfer, 2004. Forest biophysical parameter retrieval from SAR data with the aid of object oriented image analysis techniques. PPT. Presentation on the Usermeeting of eCognition 2004. University of Michigan, USA. (<http://www.definiens-imaging.com/documents/um2004/pdf/kellndorfer.pdf>)
- Wezyk, P., 2000. System monitorowania i zarządzania środowiskiem leśnym "FOREMMS" – wkład do rozwoju polityki UE zrownoważonego wykorzystania zasobów naturalnych, 2-nd International Conference: Cadastre, Photogrammetry, Geoinformatics – modern technologies and development perspectives, October 17-19.2000, Lviv, Ukraine.
- Wezyk, P., R. de Kok, G. Zajackowski (2004): The role of statistical and structural texture analysis in VHR image analysis for forest applications - A case study on Quickbird data in Niepolomice Forest. In: Strobl, J., Blaschke T. & Griesebner, G. (Hrsg.): Angewandte Geoinformatik 2004. Beiträge zum 16. AGIT-Symposium Salzburg 2004, H. Wichmann Verlag, Heidelberg, S. 770-775.

MAPPING AND MONITORING FOREST STEM VOLUME USING CARABAS

Klas Folkesson^a, Johan E.S. Fransson^b, Anders Gustavsson^c,
Björn Hallberg^a, Mattias Magnusson^b, Gary Smith-Jonforsen^a,
Helena Sämgård^a, Lars M. H. Ulander^c, Fredrik Walter^d

^a Chalmers University of Technology, Department of Radio and Space Science,
SE-412 96 Göteborg, Sweden, email: klas@rss.chalmers.se

^b Swedish University of Agricultural Sciences (SLU), Umeå, Sweden

^c Swedish Defence Research Agency (FOI), Linköping, Sweden

^d Dianthus, Boden, Sweden

ABSTRACT

CARABAS is an airborne synthetic aperture radar (SAR) system. Due to the long wavelengths of the system (3–15 m), the radar signal is able to penetrate through the forest canopies and the main backscattering component is a ground-trunk dihedral bounce. The scattering strength is directly related to the stem volume and hence stem volume retrieval with CARABAS has higher accuracy than SAR systems using shorter wavelengths. Ground topography can reduce the accuracy but this effect is mitigated by multiple images and model based correction. This correction reduced the maximum retrieval error for a site from 307 m³ ha⁻¹ to 175 m³ ha⁻¹. Complementary optical satellite data can further improve the accuracy.

Keywords: CARABAS, synthetic aperture radar (SAR), stem volume retrieval, change detection.

1 INTRODUCTION

Remote sensing of forests is one of the oldest and most important civilian applications of remote sensing. Using remote sensing, fast and cost effective data for both short- and long-term planning of forest activities can be provided. The most important forest parameter is stem volume measured per hectare, i.e. the average volume per hectare of all trees excluding branches and stumps. Synthetic aperture radar (SAR) systems have shown good potential to measure stem volume (Fransson *et al.* 2000). A strong advantage of SAR is that it is independent of weather conditions and can be used even on cloudy days.

The wavelength of the SAR determines from which part of the trees the radar signal is reflected. For SAR systems using short wavelengths (i.e. high frequency) most of the signal is reflected from the canopy. Longer wavelengths are able to penetrate the canopy and thus have greater possibility of measuring the volume of the tree trunks and larger branches where most of the biomass is stored. However, for low-frequency SAR the ground topography will affect the radar signal and if not corrected for the stem volume might be underestimated.

Besides mapping of stem volume it is also interesting to monitor temporal changes. Time series of images acquired between 1997 and 2002 have been used to study what changes can be detected and how accurately they can be measured.

2 CARABAS

CARABAS-II ultra-wideband VHF SAR system is an airborne low-frequency SAR developed by the Swedish Research Agency (FOI) (Hellsten *et al.* 1996). CARABAS operates in the lower VHF-band (20–90 MHz) which means that the wavelengths are extremely long (3–15 m) compared to most remote sensing systems. The signal is thus able to penetrate through forest canopies and direct measurements of tree trunks are possible. Furthermore, the backscattered amplitude is proportional to the stem volume and has a large dynamic range. Hence, CARABAS has higher sensitivity to biomass compared to SAR systems using shorter wavelengths, even at high biomass levels where most other remote sensing methods are saturated. The spatial resolution of the system is about 2.5 m in both range and azimuth.

3 STEM VOLUME RETRIEVAL

A limiting factor for the accuracy of stem volume retrieval from CARABAS images is the ground topography. On sloping ground the dihedral scattering between ground and trunks is reduced and is aspect angle dependent. Since this is the main source of backscatter in CARABAS images of forest, the stem volume will be underestimated. A way to reduce this effect is to combine information from images acquired from different flight directions. The ground slope in range direction will vary between the images and this affects the measured backscatter. Using a backscatter model and automated image segmentation, the different backscatter values obtained from different flight directions can be combined and the accuracy of stem volume retrieval is improved.

In this paper, 10 CARABAS images at Tönnersjöheden forest research park were used to evaluate the backscatter model and segmentation. Accurate measurements of the ground topography were also available from the helicopter-borne TopEye nadir looking LIDAR system (Sternér, 1997). From the LIDAR measurements, a bare earth digital elevation model (DEM) with 0.25 m resolution was constructed.

3.1 BACKSCATTER MODEL

The basis of the model is the Rayleigh-Gans approximation for low-frequency scattering from thin cylinders, which can be used to show that the radar cross-section for the dihedral scattering is proportional to the square of the scatterer's volume. This can be used to relate the normalised backscattering coefficient, σ^0 , to the stem volume, and the number density of the trees (Smith and Ulander, 2000). A semi-empirical model for the backscattering from coniferous forests on sloping ground has been developed (Smith *et al.*, 2005). This model describes how terrain slope (relative to the radar look direction) decreases backscatter compared to the flat ground case.

3.2 SEGMENTATION OF SAR IMAGES

Image segmentation is a standard image processing technique where images are divided into regions which have strong correlation with objects in the image (Sonka *et al.*, 1998). The number of regions and their size and shape depends on the image data. Our goal is to find forested areas where the tree size and ground slope are constant. This corresponds to regions where the backscatter is homogenous in each image but the average backscatter amplitude may vary from image to image due to different look directions. Because the ground slope usually varies on a length scale that is smaller than the size of a forest stand, each stand will most often consist of several regions. The size of a region must also be large enough to provide an accurate measure of the backscatter without being too affected by speckle.

Since the goal is to find one region map from all available images the segmentation algorithm needs to be able to use information from several images simultaneously. The segmentation algorithm used is based on simulated annealing optimisation (Folkesson, 2005). The output from the segmentation is a region map and the average backscatter for each region and image. This is used as input data to the semi-empirical model of backscatter described above, to compensate for the topographical effects.

3.3 INVERSION OF BACKSCATTER MODEL

The backscatter measurements from different look directions and their theoretical variation given by the model are used to calculate the equivalent backscatter on flat ground. Figure 1 shows an example where 10 backscatter values for a region, one from each image, are marked with an x. The semi-empirical model is least-squares fitted to the measurements to find parameters related to stem volume, tree height, and background noise level in the images, and according for SAR viewing geometry and ground topography from the DEM. From these parameters the equivalent backscatter on flat ground is calculated (the line at -11.5 dB) and this value is used as the average backscatter within the region.

Some measurement points deviate significantly from the model, usually at higher backscatter amplitude due to artefacts and man-made structures in the images, or as in the example in Figure 1, by measurements close to the noise floor. To improve the backscatter estimation, these outliers are excluded as they do not represent measurements of forest or have very low accuracy. The least-squares fitting is then repeated with the outliers removed and if the new parameters provide a better fit they are used instead.

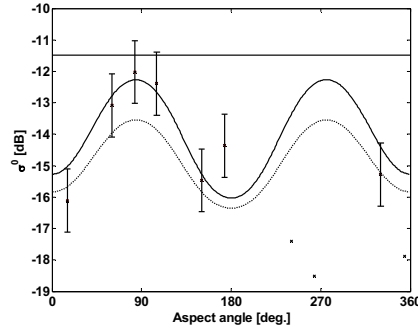


Figure 1. An example of backscatter least-squares fitting where the solid curve is the fitted curve when the three outliers have been excluded. The dotted curve is the fitted curve for all 10 measurement points. Error bars indicate the calibration accuracy of ± 1 dB. The three outliers are displayed without error bars. The straight line at -11.5 dB is the equivalent backscatter estimated after correction for the ground slope.

Figure 2 shows a comparison of two CARABAS images from perpendicular flight directions and the result after segmentation and topography correction of 10 images. Note the variations in backscatter between the two CARABAS images within areas that look much smoother in the combined image.

Figure 3 shows ground measurements of stem volume versus the average backscatter for 29 stands. In graph (a) the backscatter is measured from a single CARABAS image (shown in Figure 2 (a)) and graph (b) shows the calculated backscatter values from the combined image in Figure 2 (c). The deviation from a straight line is much smaller in (b) and the RMSE is $73 \text{ m}^3 \text{ ha}^{-1}$ compared to $124 \text{ m}^3 \text{ ha}^{-1}$ for the single image. The maximum estimation error is $175 \text{ m}^3 \text{ ha}^{-1}$ compared to $307 \text{ m}^3 \text{ ha}^{-1}$ for the single image.

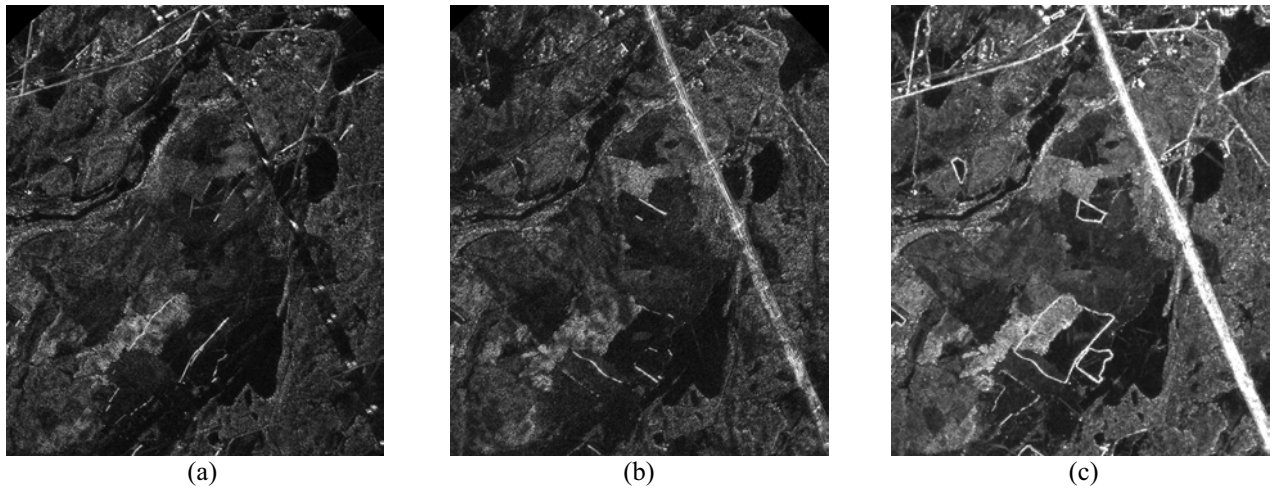


Figure 2. CARABAS images of Tönnersjöheden forest research park in south-western Sweden. Image size is 3×2 km. The flight headings are 47° (a) and 137° (b). Shown in (c) is result after segmentation and inversion of the backscatter model of totally 10 coregistered images.

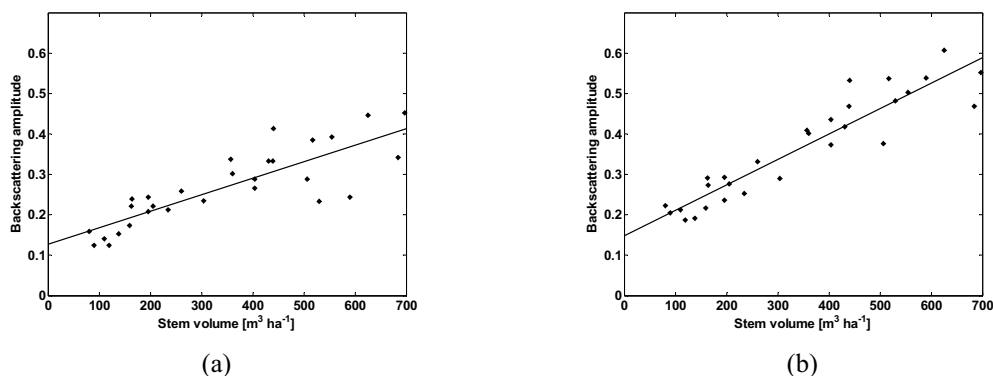


Figure 3. Ground measurements of stem volume versus CARABAS backscatter amplitude for a single image (a) and the combination of 10 images (b).



Figure 4. An RGB image made from three CARABAS images of Tönnersjöheden forest research park. The red channel was acquired in 1997, the green channel in 1999 and the blue channel in 2002. A stand map is overlaid the image.

4 CHANGE DETECTION

Temporal changes to the forest are also of great interest. Studies to detect changes have been performed using time series of CARABAS images from 1997, 1999 and 2002 (Sämgård, 2004). Clear cuts are easily detected as indicated by red and yellow areas in the RGB image in Figure 4. Red areas had high backscatter in 1997 compared to 1999 and 2002 and were thus clear cut in the period 1997–1999. Yellow areas had high backscatter in 1997 and 1999 compared to 2002 and were clear cut between 1999 and 2002.

Growth and thinning were not able to be detected by changes in backscatter. However, thinning changes the pattern in the images and this may be a way to detect it. Recent work to map storm damages supports this (Ulander *et al.* 2005). Areas that were unchanged through the time period were stable in backscatter even if the weather conditions, e.g. season and wetness, were different for the images.

5 COMBINING CARABAS WITH OTHER SENSORS

For standwise stem volume estimations below $100 \text{ m}^3 \text{ ha}^{-1}$, CARABAS has low accuracy due to low signal-to-noise ratio. By combining CARABAS data with data from another sensor that is more sensitive for low stem volumes, the accuracy of stem volume retrievals can be improved. CARABAS data have been combined with multi-spectral optical satellite data from SPOT-4 (Magnusson and Fransson, 2004a) and Landsat TM (Magnusson and Fransson, 2004b) at the test site Remningstorp in southern Sweden.

When combining CARABAS data with SPOT-4 or Landsat TM data, the improvements in accuracies were 15–23% compared to using only CARABAS data and 33–42% using only SPOT-4 or Landsat TM data for stem volumes in the range of $15\text{--}585 \text{ m}^3 \text{ ha}^{-1}$. For high stem volumes CARABAS gave the best result, while for lower stem volumes SPOT-4 or Landsat TM were more accurate as indicated for SPOT-4 in Figure 5. Hence, the combination of the two techniques provides significantly better results over the full range of stem volumes. The results imply that the combination of low-frequency SAR data and multi-spectral optical satellite data can be used for standwise stem volume estimation in forestry applications.

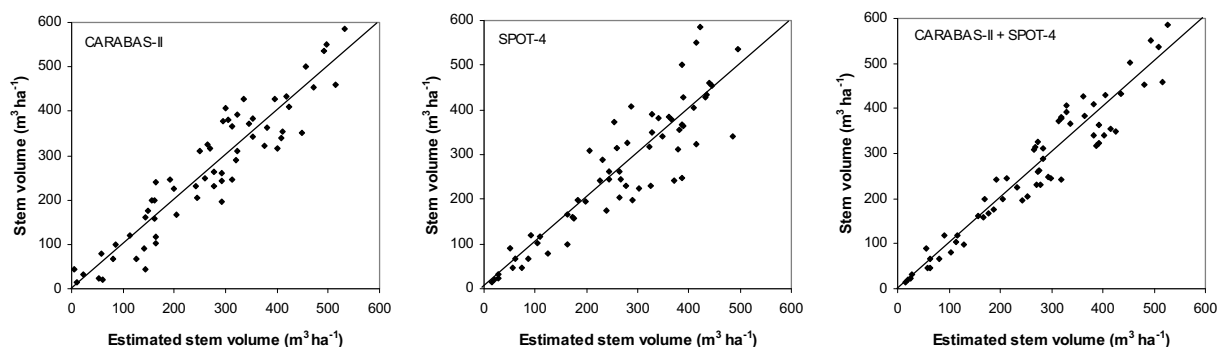


Figure 5. Stem volume from field measurements versus estimated stem volume from CARABAS data, SPOT-4 data, and the combination of CARABAS and SPOT-4 data, at stand level.

6 CARABAS PRODUCTION LINE

A pilot project aiming to develop a fully integrated system for retrieval of forest parameters from CARABAS images started in year 2001 (Walter *et al.*, 2003). The long-term goal has been to establish the VHF SAR technique on a commercial basis for stem volume mapping applicable in forestry management planning. The main effort during the project has been to assemble and to rewrite the analysis routines developed in earlier research projects into a single homogenous software environment. The system is developed as a fully integrated part of the commercial ESRI ArcGIS family of geographic information systems. At present, the system forms a complete production line from SAR data to derived forest parameters in the end-user data format.

The main output from the production line is standwise stem volume estimations. The estimations can be made within predefined forest stands or within stand like regions resulting from a segmentation of optical satellite images (e.g. SPOT-4 or Landsat TM). Recently, standwise stem volumes have been estimated within 1300 square-kilometres in southern Sweden and are now being used and evaluated by the Swedish forest company Holmen Skog.

7 CONCLUSIONS

The penetration of VHF SAR (CARABAS) through forest canopies has repeatedly shown good results for mapping forest stem volume – particularly for large coniferous trees. The backscatter mechanisms are well understood and the effect of ground slopes can be mitigated using images from multiple flight directions as described here using image segmentation combined with a backscatter model. The stability of the backscatter allows the simple detection of large changes such as clear-cutting, whereas other changes (e.g. thinning and storm felling) may require the use of image texture. The combination of CARABAS with optical satellite images has been shown to improve stem volume retrieval from both sources due to the complementary nature of the measurements.

REFERENCES

- Fransson, J.E.S., Walter, F. and Ulander, L.M.H. 2000. Estimation of forest parameters using CARABAS-II VHF SAR data. *IEEE Trans. Geosci. Remote Sensing*, Vol. 38, no. 2, pp. 720–727, Mar. 2000.
- Hellsten, H., Ulander, L.M.H., Gustavsson, A. and Larsson, B. 1996. Development of VHF CARABAS II SAR. *In proc. Radar Sensor Technology*, Orlando, USA, SPIE Vol. 2747, pp.48–60.
- Sterner, H. 1997. Helicopter aerial laser ranging. *In proc. 3rd EARSeL Workshop on LIDAR Remote Sensing of Land and Sea*, Tallin, Estonia, 1997, pp. 113–118.
- Smith, G. and Ulander, L.M.H. 2000. A model relating VHF-band backscatter to stem volume of coniferous boreal forest. *IEEE Trans. Geosci. Remote Sensing*, Vol. 38, no. 2, pp. 728–740, Mar. 2000.
- Smith, G., Ulander, L.M.H., and Luo, X. 2005. Low frequency scattering of coniferous forests on sloping terrain. *Submitted to IEEE Trans. Geosci. Remote Sensing*.
- Sonka, M., Hlavac V. and Boyle, R. 1999. *Image processing, analysis, and machine vision*. Brooks/Cole Publishing Company, Pacific Grove.
- Folkesson, K. 2005. Segmentation of Radar Images for Remote Sensing of Forests. *Thesis for the degree of licentiate of engineering*, Department of Radio and Space Science, Chalmers University of Technology, Göteborg, Sweden.
- Sämgård, H. 2004. Change Detection in VHF SAR Images of Forest from 1997, 1999 and 2002. *Thesis for the degree of master of science*, Department of Radio and Space Science, Chalmers University of Technology, Göteborg, Sweden.
- Ulander, L.M.H., Gustavsson, A., Smith, G., Eriksson, L., Folkesson, K., Hallberg, B. Fransson, J.E.S., Joyce, S., Magnusson, M., Olsson, H., Persson, A. and Walter, F. 2005. Evaluation of using space- and airborne radar for mapping of wind-thrown forests after the January 2005 hurricane in southern Sweden. *In proc. ForestSAT2005*, Borås, Sweden.
- Magnusson, M. and Fransson, J.E.S. 2004a. Combining airborne CARABAS-II VHF SAR data and optical SPOT-4 satellite data for estimation of forest stem volume. *Canadian Journal of Remote Sensing*, Vol. 30, No. 4, pp. 661–670.
- Magnusson, M. and Fransson, J.E.S. 2004b. Combining CARABAS-II VHF SAR and Landsat TM satellite data for estimation of forest stem volume. *In proc. 2004 IEEE International Geoscience and Remote Sensing Symposium, IGARSS '04*, Anchorage, USA, Vol. 4, pp. 2327–2331.
- Walter, F., Fransson, J.E.S., Gustavsson, A., Larsson, B., Smith, G., Ulander, L.M.H., and Östman, I. 2003. A Production Line for Forest Stem Volume Measurements from VHF SAR Data. *In proc. 2003 IEEE International Geoscience and Remote Sensing Symposium, IGARSS '03*, Toulouse, France, Vol. 4, pp. 2544–2546.

FOREST CHANGE DETECTION WITH SPACEBORNE L-BAND INTERFEROMETRIC SYNTHETIC APERTURE RADAR

L.E.B. Eriksson^a, A. Wiesmann^b and C. Schmullius^c

^aChalmers Univ. of Technology, Dept. of Radio and Space Science, SE-41296 Göteborg, Sweden
E-mail: leif.eriksson@rss.chalmers.se

^bGamma Remote Sensing AG, Worbstrasse 225, CH-3073 Gümlingen, Switzerland

^cFriedrich Schiller University, Institute of Geography, D-07743 Jena, Germany

ABSTRACT

This paper presents an evaluation of the possibilities to detect changes in the forest cover using spaceborne L-band interferometric synthetic aperture radar (InSAR) data from the Japanese Earth Resources Satellite (JERS-1). A test area in central Siberia was chosen for the study. Two simple methods for change detection were tested and a change detection map was created. Visual interpretation and comparison with ground data showed that in general clear-cuts should be possible to detect. However, the available forest inventory data did not include enough information on when clear-cuts, fires and other types of changes occurred to allow a full accuracy assessment. A second change detection map was created by combining L-band coherence from JERS-1 with C-band coherence from the European Remote sensing Satellites (ERS-1/2).

Keywords: Change detection, clear-cuts, boreal forest, interferometric synthetic aperture radar (InSAR), repeat-pass coherence, L-band, JERS-1.

1 INTRODUCTION

The possibility to detect deforestation with JERS-1 repeat-pass coherence has been mentioned in several studies. These studies have been done for tropical forest in the Amazon (Luckman et al., 2000; Rosen et al., 1999) and on Sumatra (Suga and Takeuchi, 2000; Takeuchi and Oguro, 2003) and all show good results. Other studies have indicated that it is possible to detect areas damaged by forest fire (Takeuchi and Yamada, 2002). For boreal forest it has been shown that it is possible to separate sparse and dense forest (Askne et al., 2003; Eriksson et al., 2003), but no results for change detection have been published.

2 BACKGROUND

Based on the availability of ground data, satellite data and meteorological data Chunksy North on the Central Siberian Plateau was selected as test area. Chunksy North is 387 km² large and contains 1226 forest stands with an average size of 32 ha. Most stands are natural stands with mixed forest. The main tree species are pine, birch, larch and spruce. The area was covered by a large number of image pairs from JERS-1 and ERS-1/2. Four of these were selected in order to cover a time period long enough for change detection. Baselines and weather conditions are given in Table 1. A thorough description of the test area, available forest inventory, satellite and meteorological data is given in (Eriksson, 2004).

Table 1. Baselines and weather conditions for the selected JERS-1 and ERS-1/2 acquisitions over Chunksy North.

Satellites	Acquisition dates	Perp. spatial baseline [m]	Temperature [°C]	Wind speed [m/s]	Snow depth [cm]	Comments
JERS-1	1993-12-29 1994-02-11	550	-13 -13	3 4	27 45	Snowfall Light snowfall
JERS-1	1996-01-16 1996-02-29	319	-23 -5	0 3	27 31	Light snowfall
ERS-1/2	1996-01-14 1996-01-15	64	-19 -24	2 0	27 27	Snowfall Light snowfall
ERS-1/2	1997-10-05 1997-10-06	239	+6 +5	1 1	0 0	Rain between acquisitions

3 DATA ANALYSIS

3.1 MULTI-TEMPORAL COMPARISON

The availability of JERS-1 pairs from two different years, preferably acquired during the same season, make it possible to search for changes in the forest cover by looking at differences in the coherence images. If a forest mask with stand borders is available it is possible to make stand based change detection. In Fig. 1 the coherence for two images is plotted for all stands larger than 5 ha in Chunskey North. The correlation coefficient of 0.89 shows that there is a high interannual consistency between the coherence values. Plots that deviate from the 1:1 line indicate that these stands have been affected by some type of change between February 1994 and January 1996. In general, the coherence is low for dense forest and high for sparse forest and open areas (Eriksson et al., 2003). For a small group of stands the coherence increased dramatically from around 0.2 in 1994 to between 0.45 and 0.67 in 1996. In the forest inventory from 1998, six of these stands are classified as clear-cuts, five as unclosed natural forest and one as natural forest. However, due to their regular shapes and strong contrast compared to surrounding dense forest, new clear-cuts are often easy to identify visually. This is clearly illustrated by the two coherence images in Fig.2. An assumption is that the five unclosed natural stands were cut in 1994 and were then left for natural regrowth. The single natural stand is most likely a clear-cut that was never reported. In four cases the coherence decreased significantly between 1994 and 1996. These stands are labelled burned forest in the inventory. Several stands that are marked as clear-cuts or burned forest in the 1998 inventory data appear on the 1:1 line in Fig. 1. This indicates that they occurred before December 1993 if they have high coherence or after January 1996 if they have low coherence. Unfortunately, the available forest inventory data do not provide any information about when clear-cutting, fires or other disturbances occurred.

If no stand borders are available another simple method can be used. Two difference images, Diff1 and Diff2, were created. One where the coherence image from the earlier acquisitions is subtracted from the coherence image from the later acquisitions (Diff1) and one where the later coherence image was subtracted from the earlier (Diff2). For the histograms of each difference image lower and upper threshold values were selected, as indicated in Fig. 3, and the images were stretched between these values. The rescaled version of Diff1 then indicate where a strong increase in coherence occurred and the rescaled version of Diff2 where the coherence decreased. Clear-cuts should be captured by Diff1. The value of the lower threshold will decide how strong differences will be filtered out and the distance between the lower and the upper thresholds will determine the dynamic range of the change image. If a simple 1/0 mask is wanted, the lower and the upper thresholds should be given the same value. A change detection map can be created by combining the new Diff1 and Diff2 with one of the original coherence images into a RGB-composite. An example of such a change map is given in Fig. 4.

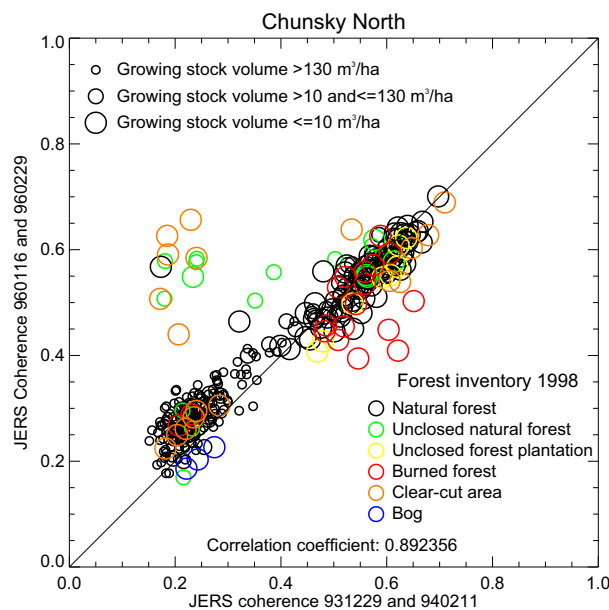


Figure 1 JERS repeat-pass coherence from two different dates plotted against each other. Only stands larger than 5 ha have been included. The sizes of the circles are determined by the stem volumes in the forest inventory from 1998.

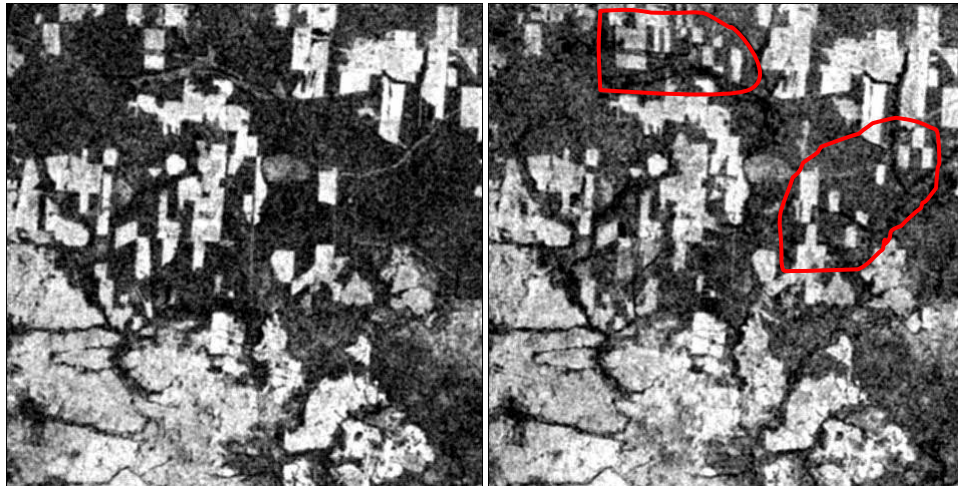


Figure 2. The image to the left is a section of the coherence image from 1993/1994 and the one to the right shows the same region two years later. The red lines mark areas with extensive change.

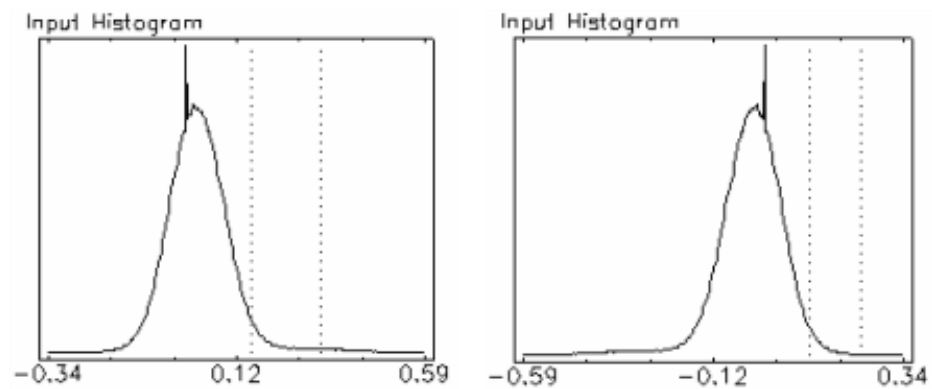


Figure 3. The histograms of the differences between the coherence images. The dotted lines indicate the lower and upper threshold values for the stretching that was used to create the change detection masks.

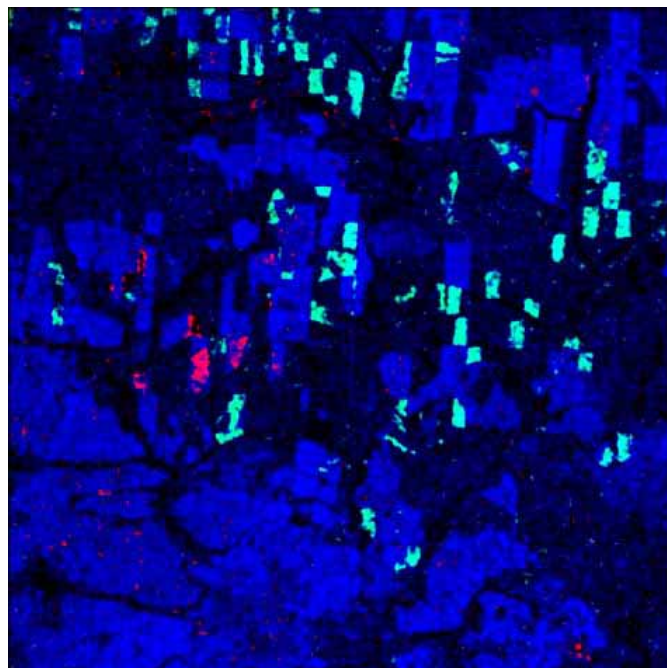


Figure 4 Change map of a section of the Chunsy area. Light green indicate areas with a strong increase in coherence and red areas with a decrease. Areas with no major change are blue. Open areas and sparse forest are bright blue and dense forest is dark blue.

3.2 COMBINING JERS-1 AND ERS-1/2

If a longer time series of JERS-1 coherence images had been available it would have been possible to create multiannual change maps that do not only give the location of the change but also show changes for more than one time period. For Chunskey this was not possible, but in addition to the JERS-1 pairs from 1993/1994 and 1996, two ERS-1/2 tandem pairs were available, the first from the same period as the last JERS-1 pair, and the second from October 1997. By combining these four pairs it was possible to create a change detection map where the identified changes could be divided in two groups depending on if they occurred in the period February 1994 to January 1996 or between January 1996 and October 1997. One version of this map is shown in Fig. 5. It was created with a methodology similar to the one described in the previous section, but instead of using Diff1 and Diff2 for the red and green colours, Diff1 from the JERS-1 difference image and Diff1 from the ERS-1/2 difference image were used. In this way only the changes where the coherence increases, e.g. clear-cuts, are included.

One source of uncertainty in this multi-sensor approach could be if the ERS-tandem coherence and the JERS-1 coherence react differently on some type of change. So far the only difference that has been identified was for bogs. One advantage of using the slope of the histograms of the difference images is that the method becomes independent of the overall coherence level of the coherence images. It is therefore no problem to use two coherence images with as different coherence levels as the ERS-1/2 pairs from January 1996 and October 1997. A more serious problem is that changes that are caused by natural seasonal differences in the coherence for some land cover types will show up as changes in the difference images. To avoid this type of misclassification it is recommended to use coherence images from the same season.

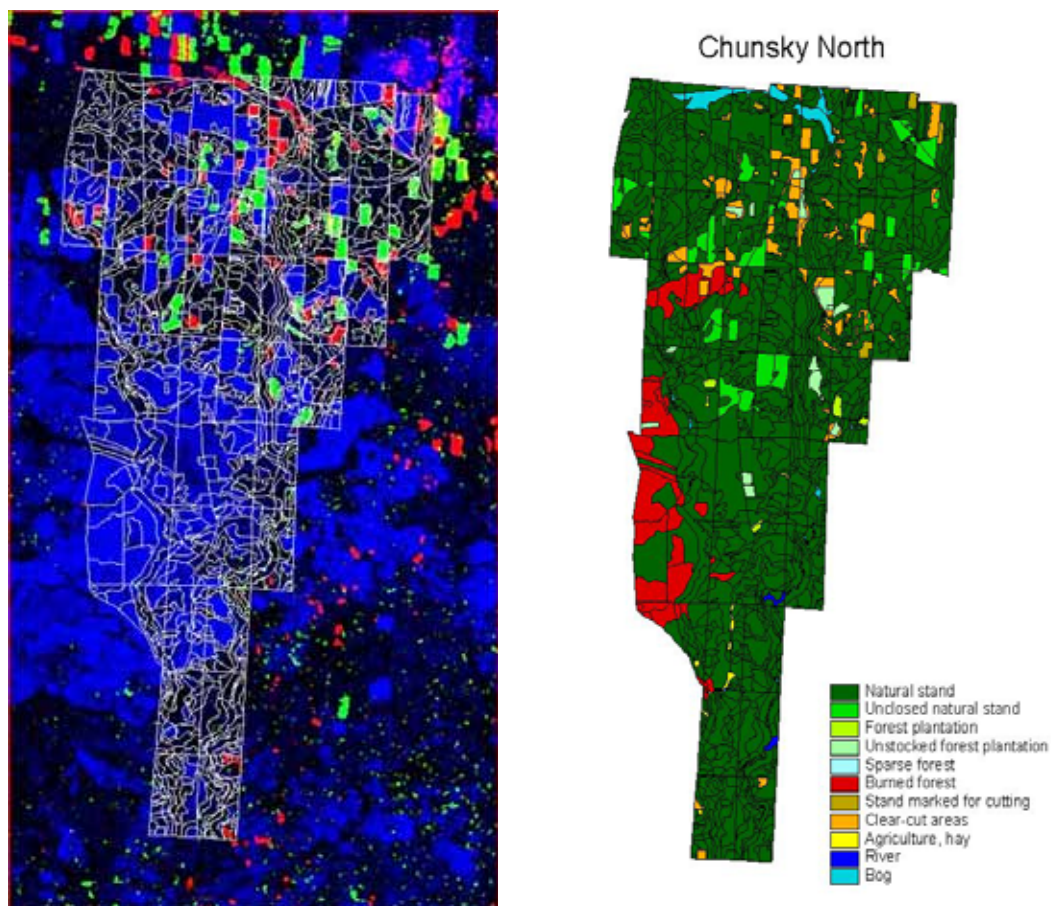


Figure 5 Change detection map for Chunskey (left figure) and land cover classes according to the forest inventory database from 1998 (right figure). The change detection map is an RGB-composite with the following bands: **Blue**: JERS-1 coherence from December 1993/February 1994. **Green**: Difference between JERS-1 coherence from January/February 1996 and JERS-1 coherence from December 1993/February 1994. **Red**: Difference between ERS-1/2 coherence from October 1997 and ERS-1/2 coherence from January 1996. **Interpretation**: **Blue**=Open areas or sparse forest before December 1993. **Green**=Changes that occurred between February 1994 and January 1996. **Red**=Changes that occurred between January 1996 and October 1997.

4 DISCUSSION

Both JERS-1 pairs that have been used in this study were acquired under frozen winter conditions. Only in boreal forest can we expect to find such conditions and compared to the continental climate in Siberia, the more coastal climate in Northern Europe reduces the probability to find image pairs acquired under similar frozen conditions. However, the vast majority of the boreal forests in the world are located far enough from the coasts to provide a climate suitable for winter acquisitions. Temporally the acquisition period is limited to a few months per year. With a repeat-cycle as long as 44 days the number of possible acquisitions is reduced so that only two, or in extreme cases three, coherence images can be formed annually. For clear-cut monitoring one annual coverage should be enough.

5. CONCLUSIONS

The presented results show that it is possible to detect large changes in boreal forest cover using satelliteborne L-band repeat-pass coherence acquired under frozen conditions. A test area in central Siberia was chosen for the study, based on the good availability of JERS-1 repeat-pass acquisitions suitable for coherence estimation. In addition, coherence from two ERS-1/2 tandem pairs was included in the study to extend the temporal coverage and allow change detection over a longer period of time.

Two different methods for change detection were tested, one that can be used if forest stand borders are known and one that is based on thresholds for difference images of an area up to the size of a satellite scene. The second method was used to create a change detection map. To demonstrate the flexibility of this method, a second change detection map was created by combining L-band coherence from JERS-1 with C-band coherence from ERS-1/2. Unfortunately the available forest inventory data did not include enough information on when clear-cuts, fires and other types of changes occurred to allow a full accuracy assessment. However, visual interpretation and comparison with the ground data showed that in general clear-cuts should be possible to detect. Forest fires, selective logging, thinning, regrowth and other types of changes where the differences are less dramatic will require improved systems and methods where changes in the growing stock volumes are quantified.

ACKNOWLEDGMENTS

The JERS-1 SAR data were acquired and provided to FSU by JAXA EORC within the framework of the Global Boreal Forest Mapping (GBFM) project. ERS-1/2 data were made available through ESA's project AO3.120 SIBERIA. Meteorological data were supplied by the DWD. Supplementary weather data were downloaded from <http://www.wunderground.com/>.

The preparation of the ground database was done within the project SIBERIA (Contract No. ENV4-CT97-0743-SIBERIA), which was financed through the 4th Framework Programme of the European Commission.

REFERENCES

- Askne, J., Santoro, M., Smith, G. and Fransson, J.E.S., 2003. Multitemporal repeat-pass SAR interferometry of boreal forests. *IEEE Transactions on Geoscience and Remote Sensing*, 41(7): 1540-1550.
- Eriksson, L.E.B., 2004. Satellite-borne L-band interferometric coherence for forestry applications in the boreal zone. Ph. D. thesis, Friedrich-Schiller-University, Jena, Germany.
- Eriksson, L.E.B., Santoro, M., Wiesmann, A. and Schmullius, C., 2003. Multi-temporal JERS repeat-pass coherence for growing stock volume estimation of Siberian forest. *IEEE Transactions on Geoscience and Remote Sensing*, 41(7): 1561-1570.
- Luckman, A., Baker, J. and Wegmüller, U., 2000. Repeat-pass interferometric coherence measurements of disturbed tropical forest from JERS and ERS satellites. *Remote Sensing of Environment*, 73: 350-360.
- Rosen, P.A., Hensley, S., Peltzer, G., Rignot, E. and Werner, C., 1999. JERS-1 synthetic aperture radar interferometry applications: mapping of rain forest environments and crustal deformation studies. In: M. Shimada (Editor), *JERS-1 Science Program '99 PI Reports*. EORC, National Space Development Agency of Japan, pp. 179-184.
- Suga, Y. and Takeuchi, S., 2000. Application of JERS-1 InSAR for monitoring deforestation of tropical rain forest, *IGARSS 2000*. IEEE Publications, Piscataway, NJ, Honolulu, 24-28 July 2000, pp. 432-434.
- Takeuchi, S. and Oguro, Y., 2003. A comparative study of coherence patterns in C-band and L-band interferometric SAR from tropical rain forest areas. *Advances in Space Research*, 32(11): 2305-2310.
- Takeuchi, S. and Yamada, S., 2002. Monitoring of forest fire damage by using JERS-1 InSAR, *IGARSS 2002*. IEEE Publications, Piscataway, NJ, Toronto, Canada.

COMMERCIAL AIRBORNE X-BAND SYNTHETIC APERTURE RADAR FOR FOREST HEIGHT INVENTORY

E.D. Wallington*, I.H. Woodhouse and Izzawati

Edinburgh Earth Observatory, School of GeoSciences (Geography), The University of Edinburgh,
Drummond Street, Edinburgh, EH8 9XP, Scotland, UK. *email: e.d.wallington@ed.ac.uk

ABSTRACT

This study assesses the capabilities of commercially available airborne short wavelength (X-band) SAR interferometry data for retrieving forest stand height using data from the Intermap Technologies Star 3i system over three test sites in the UK. Forest stand height is an important indicator of standing biomass for management purposes as well as for the assessment of carbon storage. Tree height retrieval is achieved by subtracting the ground surface (UK Ordnance Survey DEM, OSDEM) from the canopy surface (Intermap Digital Surface Model, DSM), and the results are then compared to known stand heights. The DSM-OSDEM procedure produces results in line with expectation – i.e. the DSM underestimates the canopy height by 10-20%. Relationships between radar retrieved height and forest parameters such as stocking density and age, and radar dependant properties such as slope and edge effects are presented as possible explanations for variations across the collected data. Supporting work using a simple coherent interferometric scattering model was also used to characterise and explain the effects on tree height retrieval due to variations in slope, number density, stand height and forest edges. Also introduced are forest height maps produced from the SAR height retrieval methodology. These maps can be used for management purposes as a visual aid, as well as being incorporated into existing GIS systems to aid decision making at the forest, stand and, due to the high resolution, sub-stand level. This new data set therefore potentially allows a rapid and timely management tool for use in cost-effective sustainable forest management and conservation initiatives.

Keywords: SAR interferometry, height, forest inventory, monitoring, DSM, edge effects, PRIS model.

1 INTRODUCTION

At present, UK forest cover (11.6%, 2,731,000 ha) falls significantly below the European average of 46% (FC, 2004a), but represents a sizeable spatial resource. In accordance with the UK Forestry Standard, management of both existing and expanding forests must be sustainable to address the increase of timber and other woodland resources, as well as social and economic impacts (FC, 2004b). Furthermore, sustainable forest management (EC, 2003) directly contributes to the sequestration of atmospheric carbon, thus reducing levels of and the negative effects of greenhouse gases. Governments are committed to achieving emission reduction targets set by the Kyoto Protocol. Crucially, the Protocol calls for provision to consider afforestation, reforestation, and deforestation (ARD) when meeting national commitments (IPCC, 2000). Therefore, techniques for mapping, quantifying, managing and monitoring forest expanse have critical relevance to achievement of a low carbon economy, as well as a sustainable forest resource (Hickey *et al.*, 2005). Remote sensing effectively addresses these issues and has already been identified by the Intergovernmental Panel on Climate Change (IPCC) as the desired technology to measure and monitor forested landscapes (IPCC, 2000).

The use of remote sensing as a means of data collection, quantification and assessment has been actively pursued for forest management applications at a range of scales from individual trees, stand, forest, and landscape, to national, international and global (Franklin, 2001; Donoghue, 2002). Remote sensing can provide useful information about the vertical and horizontal structure of the forest at a range of scales. Image processing analysis further enhances the diverse data at a variety of scales that can be readily integrated into a GIS-based management system.

This study assesses the capabilities of commercially available airborne short wavelength (X-band) Synthetic Aperture Radar (SAR) interferometry data for retrieving forest stand height using data from the Intermap Technologies Star-3i system (Intermap, 2005) over three test sites in the UK. Such data is rapidly being acquired for a large number of countries, for example Nextmap Britain has produced UK-

wide digital elevation data, and so is becoming increasingly available to users. The data has been extensively assessed for height accuracy over bare and sparsely vegetated ground (Dowman *et al.*, 2003; Wallington *et al.*, 2004); in this study we assess its performance in combination with ground elevation data to estimate canopy height. Forest stand height is an important indicator of standing biomass for management purposes as well as for the assessment of carbon storage. Height is also an important ecological parameter in its own right, and an important input parameter for line-of-site analysis. This new data set therefore potentially allows a rapid and timely management tool for use in cost-effective sustainable forest management and conservation initiatives.

2 STUDY SITES AND DATA

2.1 STUDY SITES

Three study sites were considered: Coed Y Brenin Forest District in North Wales (N52:49:12 W3:53:27 lat/long), Kielder Forest District in Northumberland, North England (N55:11:44 W2:32:11 lat/long), and Glen Affric in the Scottish Highlands (N57:17:00 W4:54:49 lat/long). 59 plantation forest stands were considered, and consisted of a number of species including Sitka spruce (*Picea sitchensis*), Norway spruce (*Picea abies*), Japanese larch (*Larix kaempferi*), Western hemlock (*Tsuga heterophylla*) and Scots pine (*Pinus sylvestris*).

2.2 DATA SETS

2.2.1 InSAR digital surface model

The InSAR Digital Surface Model (DSM) was supplied by Intermap Technologies, and was geo-referenced to OSGB36 prior to delivery. The DSM is produced from the first return of the signal which represents the first surface the signal came into contact with, whether it is the ground or vegetation canopy. The DSM has a pixel size of 5m and a RMSE of between 0.5 – 1.0m (95% = 1.0 - 2.0m) dependent on flying height. These accuracies are quoted for moderately sloped, unobstructed terrain (Intermap, 2003).

2.2.2 Ordnance Survey DEM

The Ordnance Survey (OS) Profile 10m Digital Elevation Model (OSDEM) was used as a ground reference surface. The OSDEM has a pixel size of 10m (re-sampled to 5m to match DSM; this was purely to ease comparison and not an attempt to improve resolution). This DEM is assumed to be a representation of the true ground surface for the purposes of this study.

2.2.3 Complimentary reference data

Aerial photographs were supplied by Forest Enterprise, and covered all of the Coed Y Brenin test area, and partial areas of the Kielder test area. These photographs were used for assessing ground cover type, and relative (visual) tree height assessment. The Forestry Commission Sub Compartment Database (SCDB) was used to gather information on stand details, for example species and age. Traditional forest inventory techniques (Philip, 1994; Husch *et al.*, 2003; Hamilton, 1998;) were used to establish the top height of the sample stands. This height was then used as the true top height of the stand. In total, 59 plots were measured, distributed between Coed Y Brenin (23), Kielder (11) and Glen Affric (25).

3 FOREST HEIGHT RETRIEVAL

Tree height and stand height can be described in a number of ways. In this study, stand height is expressed as the traditional Top Height (H_{100}) measurement, which is defined as either the average height of the 100 trees/ha with largest diameter at breast height (DBH), or the average height of the 100 tallest trees/ha (Philip, 1994):

$H_{100} = \sum_{i=1}^{100} \frac{H_i}{N}$	(1)
--	-----

where H_i is total height and N is the sample size. Similarly, the retrieved top height (H_{r100}) is defined as the average of the highest 100 retrieved heights/ha:

$$H_{r100} = \sum_0^{100} \frac{H_r}{N} \quad (2)$$

where H_r is the retrieved height. Top height is the average of the total heights of the specified number of trees; total height is defined as:

$$\text{Height}_{\text{Tree}} = \text{Height}_{\text{Canopy}} - \text{Height}_{\text{Ground}} \quad (3)$$

The specific realisation of equation 3 used to retrieve tree height using SAR in this study was achieved by subtracting the OSDEM from the DSM, as defined by:

$$H_{\text{Tree}} = H_{\text{DSM}} - H_{\text{OSDEM}} \quad (4)$$

In each of the chosen stands, height values from the DSM and OSDEM were retrieved from a 50 x 50m plot; smaller plots were used if the stand was too small to accommodate a 50 x 50m. Equation 4 was implemented to retrieve the height per pixel within the plot, and equation 2 was used to estimate the retrieved top height. This top height was then compared to the measured top height per plot.

4 HEIGHT RETRIEVAL RESULTS AND DISCUSSION

Subtraction of the OSDEM from the DSM to retrieve a top height produced a strong Pearson's correlation of 0.79 (Figure 1, left). As expected, the retrieved height is an underestimation due to signal penetration through the canopy. At the shorter wavelengths (X-band), penetration is limited to the upper canopy, and as such the resultant height of the scattering phase center is predominately dominated by scattering from the smaller scatterers in the canopy. The resulting average underestimation is around 30-35%.

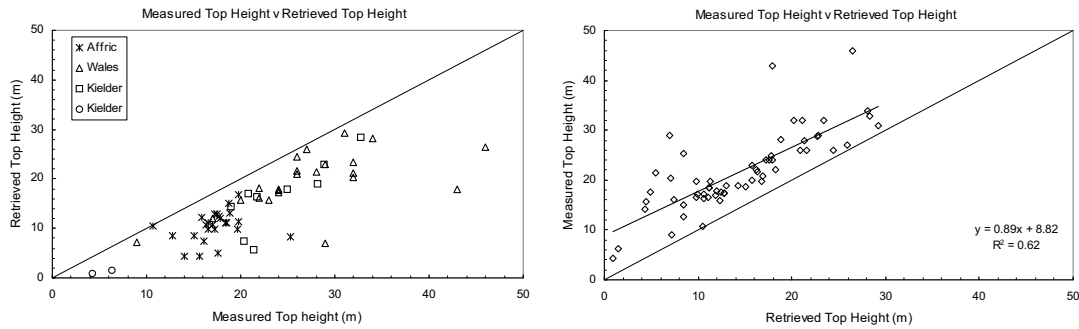


Figure 1. Left: Measured Top height plotted against Retrieved Top height. Average error of 7.23m (34.7%), observations = 59, Pearson's $R = 0.79$. Note: Kielder (○) data were collected independently of Kielder (□). **Right:** Regression analysis of Retrieved Top height plotted against Measured Top, $R^2 = 0.62$, predicts tree height to within ± 4.7 m RMSE, for a 20m high tree, this represents a prediction error of 23.5%.

Inversion of the retrieved heights through regression allows the prediction of actual top height based on the retrieved top heights (Figure 1, right). This results in a prediction accuracy of ± 4.7 m RMSE; for a 20m high tree, this represents a prediction error of 23.5%, which is an improved accuracy when compared to the raw retrieved heights. There are however a number of outlying points which are evident and can be better visualised by looking at the residuals of the predicted heights (Figure 2, left). Accounting for the 7 outliers with residuals greater than 6m (Figure 2, right), the RMSE is reduced to ± 2.5 m ($R^2 = 0.84$); for a 20m high tree, this represents a prediction error of 12.5%, a further improvement on the raw height retrieval. When looking at individual species, for example Sitka spruce (*Picea sitchensis*), regression analysis predicts actual height with an R^2 of 0.84 and 0.97 in Coed Y Brenin and Kielder respectively.

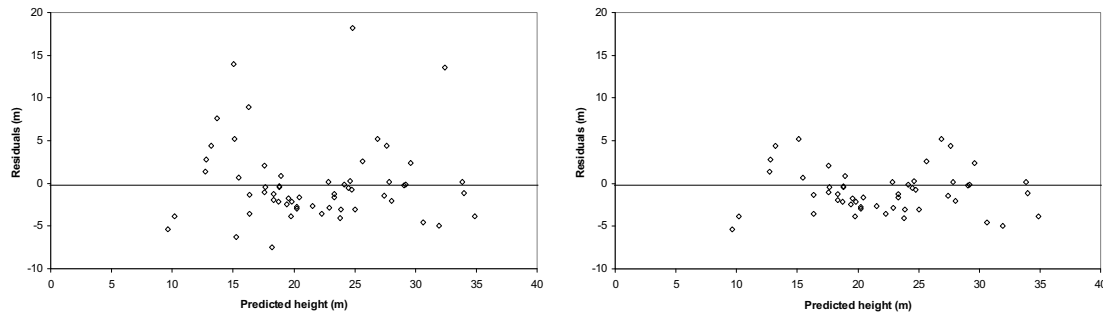


Figure 2. Predicted Top height residuals (left) and with residuals > 6m (7 plots) accounted for (right).

5 MODELLING RETRIEVED HEIGHT

A newly developed Polarimetric Radar Interferometry Simulator (PRIS) (Izzawati *et al.*, 2004; Izzawati *et al.*, Submitted; Woodhouse *et al.*, Submitted.) was used to examine relationships between retrieved height and forest parameters such as stocking density and age, and radar dependent properties such as slope and edge effects are assessed and quantified by way of a simple interferometric model. Underestimation is expected due to the variation in the location of the scattering phase centre, whose values depend on the relative scattering contribution from the ground surface and forest trees. At X-band, short penetration depth means that most scattering comes from the ground surface and near top of the trees. Increasing contribution from the ground surface near the edge of the plantation results in severe height underestimation. Better height retrieval can be obtained at the areas far from the edge as the main scattering contribution comes mainly from the top of the tree crowns (Woodhouse *et al.*; submitted). The extent and amount of height underestimation appears to be affected by viewing angle, tree height and density. Further height retrieval differences will be species dependent, whereby the canopy characteristics will vary, with some canopies being conical and dense (e.g. Sitka spruce) and other being more elliptical and open (e.g. Scots pine), resulting in different penetration depths.

6 CONCLUSION

In this paper, the use of commercial airborne X-band SAR data has been demonstrated and analysed for forest height inventory. Subtracting a ground surface (OSDEM) from a top of canopy surface (DSM) provides an estimate of tree height. The retrieved tree heights were converted into retrieved top height and compared to measured top heights. The findings are summarised as:

- Top height retrieval when compared to actual top heights gave an average error of 7.23m (34.7%).
- Inversion of the height retrieval predicted measured heights based on retrieved heights to within ± 4.7 m RMSE, for a 20m high tree, this represents a prediction error of 23.5%.
- Accounting for outliers with a residual error of >6m improved height prediction to ± 2.5 m RMSE, for a 20m high tree, this represents a prediction error of 12.5%.
- When comparing individual species (e.g. Sitka spruce), regression could predict measure height based on retrieved height with a 97% correlation.
- Reasons for height under prediction were investigated. Canopy characteristics (shape and density) as well as radar viewing angle, tree height and density were seen as potential factors. The influence of edge effects was also examined, and results indicate that better height retrieval is achieved far from edges.

These results themselves are interesting, but the data becomes more operationally useable when viewed as forest height maps (Figure 3, left). These maps can be used for management purposes as a visual aid, as well as being incorporated into existing GIS systems to aid decision making at the forest, stand and, due to the high resolution, sub-stand level (Figure 3, right). This new data set therefore potentially allows a rapid and timely management tool for use in cost-effective sustainable forest management and conservation initiatives.

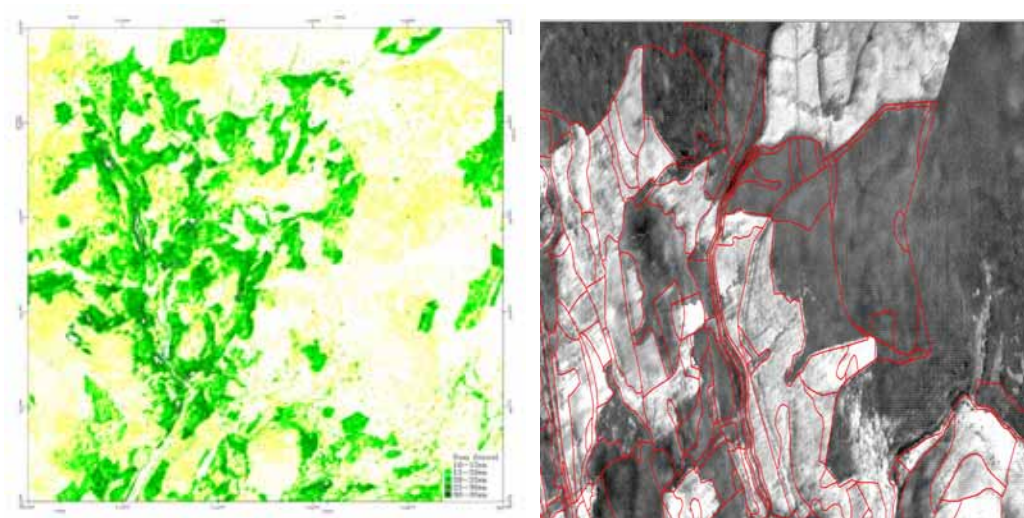


Figure 3. Left: Forest height map of Coed Y Brenin, N. Wales, UK. Darker shades of green represent taller trees. **Right:** Forest height map with Forestry Commission stand boundaries overlaid (red lines). Brighter areas represent taller trees. Stand boundaries show a good fit, as well as added value of stand internal variation.

ACKNOWLEDGMENTS

The authors would like to thank Intermap Technologies for kindly supplying the SAR data for the test sites. Thanks are also due to Stuart Snape (UK Forest Enterprise) and Juan Suárez (UK Forest Research) for supplying ground reference data for Coed Y Brenin and Kielder respectively, and to Daniel Donoghue and Pete Watt of Durham University for supplementary data for Kielder. This study was funded by the NERC and the UK Forestry Commission.

REFERENCES

- Donoghue, D.N.M. 1992. Remote sensing: environmental change. *Progress in Physical Geography*, Vol. 26, No. 1, pp. 144-151.
- Dowman, I., Balan, P., Renner, K. and Fischer, P. 2003. An Evaluation of Nextmap Terrain Data in the Context of UK National Datasets. A report for Getmapping. University College London. UK.
- EC. 2003. Sustainable Forestry and the European Union. European Commission.
- FC. 2004a. Forestry Fact & Figures 2004. Forestry Commission, Edinburgh, UK.
- FC. 2004b. The UK Forestry Standard. Forestry Commission, Edinburgh, UK.
- Franklin, S.E. 2001. Remote Sensing for Sustainable Forest Management. CRC Press LLC, Florida, USA.
- Hamilton, G.J. 1998. Forest Mensuration. Forestry Commission Booklet 39. Fifth impression. Her Majesty's Stationery Office, UK.
- Hickey, G.M., Innes, J.L., Kozak, R.A., Bull, G.Q. and Vertinsky, I. 2005. Monitoring and information reporting for sustainable forest management: An international multiple case study analysis. *Forest Ecology and Management*. Vol. 209, pp.237-259.
- Husch, B.; Beers, T.W. and Kershaw, J.A. 2003. Forest Mensuration. 4th Edition. John Wiley & Sons, Inc. New Jersey.
- Intermap. 2003. Product Handbook and Quick Start Guide. Intermap Technologies Corp.
- Intermap. 2005. <http://www.intermap.com>. Intermap Technologies. Last accessed 4th April 2005.
- IPCC. 2004. Special Report: Land Use, Land Use Change and Forestry. Intergovernmental Panel on Climate Change.
- Izzawati, Wallington, E.D. and Woodhouse, I.H. 2004. Forest height retrieval from commercial X-band SAR products in the UK. In *Proceedings of 4th International Symposium on Retrieval of Bio- and Geophysical Parameters from SAR Data for Land Applications*, 16th-19th November 2004, Innsbruck, Austria.
- Izzawati, Wallington, E.D. and Woodhouse, I.H. Submitted. Forest height retrieval from commercial X-band SAR products. *IEEE Transactions on Geoscience and Remote Sensing*.
- Philip, M.S. 1994. Measuring Trees and Forests. Second Edition. CAB International, Oxon, UK.
- Wallington, E.D., Izzawati and Woodhouse, I.H. 2004. Forest Height Estimation from X-band SAR. In *Proceedings of IGARSS 2004*, 20th-24th September 2004, Anchorage, Alaska.
- Woodhouse, I.H., Izzawati, Wallington, E.D. and Turner, D. Submitted. Edge effects on tree height retrieval using X-band interferometry. *IEEE Transactions on Geoscience and Remote Sensing*.

MAPPING THE 3D DISTRIBUTION OF FOREST CANOPY BIOMASS WITH SAR

K. Morrison

Dept. of Aerospace, Power & Sensors, University of Cranfield, Shrivenham,
Swindon, SN6 8LA, UK. E-mail: k.morrison@cranfield.ac.uk

ABSTRACT

A programme of research is proposed which will examine, validate and develop polarimetric SAR for forest biomass measurement, which is potentially the key contribution of SAR for global studies of the carbon cycle.

Keywords: SAR, GB-SAR, biomass.

1 INTRODUCTION

Forests sequester carbon from the atmosphere during photosynthesis, release carbon through respiration and ecological disturbances and are valued for their capacity to store carbon over long time scales. To fully understand quantitative relationships in the terrestrial components of the global carbon cycle, forest biomass estimates are needed that are spatially comprehensive and repeatable. Over the past few years novel observing techniques have been developed exploiting polarimetric and interferometric capabilities that are now available from airborne and satellite SAR systems, of which PolInSAR is the most prominent such technique. It exploits the polarisation dependence of scattering mechanisms to estimate scattering phase centre heights, which can be extrapolated to tree height.

However, no comprehensive assessment of the operational potential and limitations of PolInSAR is yet available. Particular open questions relate to the conditions under which PolInSAR produces accurate results, with respect to structural canopy types, technical sensor specifications, and imaging conditions. In addition, current model-based approaches used to interpret the data need much stronger experimental confirmation. To address these deficiencies a ground-based 3D SAR tomography measurement campaign based around the GB-SAR System (Morrison, 2003) is proposed.

2 METHODOLOGY

GB-SAR is a mobile, fully polarimetric SAR imaging system (Morrison, 2003). A programme of field measurement is proposed in combination with analyses from coherent modelling, and airborne and satellite imagery. Tomography exploits multi-incidence-angle 2D imaging, and a fully polarimetric mode will provide truly 3D realisation of the polarimetric scattering pattern through a canopy, allowing examination of the effects of variations in vertical biomass distribution of forest stands (Morrison, 2001). It will make it completely clear how to interpret the measurements produced by polarimetric interferometric SARs, remove interpretation ambiguities, and introduce a whole new set of possibilities for SAR applications and inversion schemes.

ACKNOWLEDGMENTS

This work is supported in part by the Natural Environment Research Council under grant NER/M/2003/00091.

REFERENCES

- Morrison, K, Bennett, J.C., Cookmartin, G., Quegan, S. and Race, A., 2003. Development and Capabilities of the Ground-Based SAR (GB-SAR) Facility. In proceedings of Advanced SAR Workshop, Montreal.
- Morrison, K, Bennett, J.C., Cookmartin, G., McDonald, A.J., Race, A. and Quegan, S., 2003. Three-dimensional X-band imaging of a small conifer tree. *Int. J. Remote Sens.* 22:705-710.

Av Skogsstyrelsen publicerade Rapporter:

- 1988:1 Mallar för ståndortsbonitering; Lathund för 18 län i södra Sverige
- 1988:2 Grusanalys i fält
- 1990:1 Teknik vid skogsmarkskalkning
- 1991:1 Tätortsnära skogsbruk
- 1991:2 ÖSI; utvärdering av effekter mm
- 1991:3 Urboträffar; utvärdering
- 1991:4 Skogsskador i Sverige 1990
- 1991:5 Contortarapporten
- 1991:6 Participation in the design of a system to assess Environmental Consideration in forestry a Case study of the GREENERY project
- 1992:1 Allmän Skogs- och Miljöinventering, ÖSI och NISP
- 1992:2 Skogsskador i Sverige 1991
- 1992:3 Aktiva Natur- och Kulturvårdande åtgärder i skogsbruket
- 1992:4 Utvärdering av studiekampanjen Rikare Skog
- 1993:1 Skoglig geologi
- 1993:2 Organisationens Dolda Resurs
- 1993:3 Skogsskador i Sverige 1992
- 1993:5 Nyckelbiotoper i skogarna vid våra sydligaste fjäll
- 1993:6 Skogsmarkskalkning – *Resultat från en fyraårig försöksperiod samt förslag till åtgärdsprogram*
- 1993:7 Betespräglad äldre bondeskog – *från naturvårdssynpunkt*
- 1993:8 Seminarier om Naturhänsyn i gallring i januari 1993
- 1993:9 Förbättrad sysselsättningsstatistik i skogsbruket – *arbetsgruppens slutrapport*
- 1994:1 EG/EU och EES-avtalet ur skoglig synvinkel
- 1994:2 Hur upplever "grönt utbildade kvinnor" sin arbetssituation inom skogsvårdsorganisationen?
- 1994:3 Renewable Forests - Myth or Reality?
- 1994:4 Bjursåsprojektet - *underlag för landskapsekologisk planering i samband med skogsinventering*
- 1994:5 Historiska kartor - *underlag för natur- och kulturmiljövård i skogen*
- 1994:6 Skogsskador i Sverige 1993
- 1994:7 Skogsskador i Sverige – *nuläge och förslag till åtgärder*
- 1994:8 Häckfågelinventering i en åkerholme åren 1989-1993
- 1995:1 Planering av skogsbrukets hänsyn till vatten i ett avrinningsområde i Gävleborg
- 1995:2 SUMPSKOG – ekologi och skötsel
- 1995:3 Skogsbruk vid vatten
- 1995:4 Skogsskador i Sverige 1994
- 1995:5 Långsam alkalinisering av skogsmark
- 1995:6 Vad kan vi lära av KMV-kampanjen?
- 1995:7 GROT-uttaget. Pilotundersökning angående uttaget av trädrester på skogsmark
- 1996:1 Women in Forestry – What is their situation?
- 1996:2 Skogens kvinnor – Hur är läget?
- 1996:3 Landmollusker i jämtländska nyckelbiotoper
- 1996:4 Förslag till metod för bestämning av prestationstal m.m. vid självverksamhet i småskaligt skogsbruk.
- 1997:1 Sjövatten som indikator på markförsurning
- 1997:2 Naturvårdsutbildning (20 poäng) Hur gick det?
- 1997:3 IR-95 – Flygbildsbaserad inventering av skogsskador i sydvästra Sverige 1995
- 1997:5 Miljeu96 Rådgivning. Rapport från utvärdering av miljeurådgivningen
- 1997:6 Effekter av skogsbränsleuttag och askåterföring – *en litteraturstudie*
- 1997:7 Målgruppsanalys
- 1997:8 Effekter av tungmetallnedfall på skogslevande landsnäckor (*with English Summary: The impact on forest land snails by atmospheric deposition of heavy metals*)
- 1997:9 GIS-metodik för kartläggning av markförsurning – *En pilotstudie i Jönköpings län*
- 1998:1 Miljökonsekvensbeskrivning (MKB) av skogsbränsleuttag, asktillförsel och övrig näringskompensation
- 1998:2 Studier över skogsbruksåtgärdernas inverkan på snäckfaunans diversitet (*with English summary: Studies on the impact by forestry on the mollusc fauna in commercially used forests in Central Sweden*)
- 1998:3 Dalaskog - Pilotprojekt i landskapsanalys
- 1998:4 Användning av satellitdata – *hitta avverkad skog och uppskatta lövröjningsbehov*
- 1998:5 Baskatjoner och aciditet i svensk skogsmark - tillstånd och förändringar
- 1998:6 Övervakning av biologisk mångfald i det brukade skogslandskapet. *With a summary in English: Monitoring of biodiversity in managed forests.*
- 1998:7 Marksvampar i kalkbarrskogar och skogsbeten i Gotländska nyckelbiotoper
- 1998:8 Omgivande skog och skogsbrukets betydelse för fiskfaunan i små skogsbäckar
- 1999:1 Miljökonsekvensbeskrivning av Skogsstyrelsens förslag till åtgärdsprogram för kalkning och vitalisering
- 1999:2 Internationella konventioner och andra instrument som behandlar internationella skogsfrågor
- 1999:3 Mållklassificering i "Gröna skogsbruksplaner" - betydelsen för produktion och ekonomi
- 1999:4 Scenarier och Analyser i SKA 99 - Förutsättningar

2000:1	Samordnade åtgärder mot försurning av mark och vatten - Underlagsdokument till Nationell plan för kalkning av sjöar och vattendrag
2000:2	Skogliga Konsekvens-Analyser 1999 - Skogens möjligheter på 2000-talet
2000:3	Ministerkonferens om skydd av Europas skogar - Resolutioner och deklarationer
2000:4	Skogsbruket i den lokala ekonomin
2000:5	Aska från biobränsle
2000:6	Skogsskadeinventering av bok och ek i Sydsverige 1999
2001:1	Landmolluskfaunans ekologi i sump- och myrskogar i mellersta Norrland, med jämförelser beträffande förhållandena i södra Sverige
2001:2	Arealförluster från skogliga avrinningsområden i Västra Götaland
2001:3	The proposals for action submitted by the Intergovernmental Panel on Forests (IPF) and the Intergovernmental Forum on Forests (IFF) - in the Swedish context
2001:4	Resultat från Skogsstyrelsens ekenkät 2000
2001:5	Effekter av kalkning i utströmningsområden <i>med kalkkross 0 - 3 mm</i>
2001:6	Biobränslen i Söderhamn
2001:7	Entreprenörer i skogsbruket 1993-1998
2001:8A	Skogspolitisk historia
2001:8B	Skogspolitiken idag - en beskrivning av den politik och övriga faktorer som påverkar skogen och skogsbruket
2001:8C	Gröna planer
2001:8D	Föryngring av skog
2001:8E	Fornlämningar och kulturmiljöer i skogsmark
2001:8G	Framtidens skog
2001:8H	De skogliga aktörerna och skogspolitiken
2001:8I	Skogsbilvägar
2001:8J	Skogen sociala värden
2001:8K	Arbetsmarknadspolitiska åtgärder i skogen
2001:8L	Skogsvårdsorganisationens uppdragsverksamhet
2001:8M	Skogsbruk och rennäring
2001:8O	Skador på skog
2001:9	Projekterfarenheter av landskapsanalys i lokal samverkan – (LIFE 96 ENV S 367) Uthålligt skogsbruk byggt på landskapsanalys i lokal samverkan
2001:11A	Strategier för åtgärder mot markförsurning
2001:11B	Markförsurningsprocesser
2001:11C	Effekter på biologisk mångfald av markförsurning och motåtgärder
2001:11D	Urvalskriterier för bedömning av markförsurning
2001:11E	Effekter på kvävedynamiken av markförsurning och motåtgärder
2001:11F	Effekter på skogsproduktion av markförsurning och motåtgärder
2001:11G	Effekter på tungmetallens och cesiums rörlighet av markförsurning och motåtgärder
2001:12	Forest Condition of Beech and Oak in southern Sweden 1999
2002:1	Ekskador i Europa
2002:2	Gröna Huset, slutrapport
2002:3	Project experiences of landscape analysis with local participation – (LIFE 96 ENV S 367) Local participation in sustainable forest management based on landscape analysis
2002:4	Landskapsekologisk planering i Söderhamns kommun
2002:5	Miljöriktig vedeldning - Ett informationsprojekt i Söderhamn
2002:6	White backed woodpecker landscapes and new nature reserves
2002:7	ÄBIN Satellit
2002:8	Demonstration of Methods to monitor Sustainable Forestry, Final report Sweden
2002:9	Inventering av frötaktssbestånd av stälke, bergek och rödek under 2001 - Ekdöd, skötsel och naturvård
2002:10	A comparison between National Forest Programmes of some EU-member states
2002:11	Satellitbildsbaserade skattningar av skogliga variabler
2002:12	Skog & Miljö - Miljöbeskrivning av skogsmarken i Söderhamns kommun
2003:1	Övervakning av biologisk mångfald i skogen - En jämförelse av två metoder
2003:2	Fågelfaunan i olika skogsmiljöer - en studie på beståndsnivå
2003:3	Effektivare samråd mellan rennäring och skogsbruk -förbättrad dialog via ett utvecklat samrådsförfarande
2003:4	Projekt Nissadalen - En integrerad strategi för kalkning och askspridning i hela avrinningsområden
2003:5	Projekt Renbruksplan 2000-2002 Slutrapport, - ett planeringsverktyg för samebyarna
2003:6	Att mäta skogens biologiska mångfald - möjligheter och hinder för att följa upp skogspolitikens miljömål i Sverige
2003:7	Vilka botaniska naturvärden finns vid torplämningar i norra Uppland?
2003:8	Kalkgranskogar i Sverige och Norge – förslag till växtsociologisk klassificering
2003:9	Skogsägare på distans - Utvärdering av SVO:s riktade insatser för utbör
2003:10	The EU enlargement in 2004: analysis of the forestry situation and perspectives in relation to the present EU and Sweden
2004:1	Effektuppföljning skogsmarkskalkning tillväxt och trädvitalitet, 1990-2002
2004:2	Skogliga konsekvensanalyser 2003 - SKA 03
2004:3	Natur- och kulturinventeringen i Kronobergs län 1996 - 2001

2004:4	Naturlig föryngring av tall
2004:5	How Sweden meets the IPF requirements on nfp
2004:6	Synthesis of the model forest concept and its application to Vilhelmina model forest and Barents model forest network
2004:7	Vedlevande arters krav på substrat - sammanställning och analys av 3.600 arter
2004:8	EU-utvidgningen och skogsindustrin - En analys av skogsindustrins betydelse för de nya medlemsländernas ekonomier
2004:9	Nytt nummer se 2005:1
2004:10	Om virkesförrådets utveckling och dess påverkan på skogsbrukets lönsamhet under perioden 1980-2002
2004:11	Naturskydd och skogligt genbevarande
2004:12	När vi skogspolitikens mångfaldsmål på artnivå? - Åtgärdsförslag för uppföljning och metodutveckling
2005:1	Access to the forests for disabled people
2005:2	Tillgång till naturen för människor med funktionshinder
2005:3	Besöksstudier i naturområden - en handbok
2005:4	Visitor studies in natureareas - a manual
2005:5	Skogshistoria år från år 1177-2005
2005:6	Vägar till ett effektivare samarbete i den privata tätortsnära skogen
2005:7	Planering för rekreation - Grön skogsbruksplan i privatägd tätortsnära skog
2005:8a-8c	Report from Proceedings of ForestSAT 2005 in Borås May 31 - June 3

Order of Reports

National Board of Forestry
551 83 JÖNKÖPING
Tel: 036 – 15 55 92
vx: 036 – 15 56 00
fax: 036 – 19 06 22
mail: sksforlag.order@svo.se
www.svo.se/forlag

I Skogsstyrelsens författningssamling (SKSFS) publiceras myndighetens föreskrifter och allmänna råd. Föreskrifterna är av tvingande natur. De allmänna råden är generella rekommendationer som anger hur någon kan eller bör handla i visst hänseende.

I Skogsstyrelsens Meddelande-serie publiceras redogörelser, utredningar m.m. av officiell karaktär. Innehållet överensstämmer med myndighetens policy.

I Skogsstyrelsens Rapport-serie publiceras redogörelser och utredningar m.m. för vars innehåll författaren/författarna själva ansvarar.

Skogsstyrelsen publicerar dessutom fortlöpande: Foldrar, broschyrer, böcker m.m. inom skilda skogliga ämnesområden.

Skogsstyrelsen är också utgivare av tidningen Skogseko.

Av Skogsstyrelsen publicerade Meddelanden:

1991:2	Vägplan -90
1991:3	Skogsvårdsorganisationens uppdragsverksamhet – Efterfrågade tjänster på en öppen marknad
1991:4	Naturvårdshänsyn – Tagen hänsyn vid slutavverkning 1989–1991
1991:5	Ekologiska effekter av skogsbränsleuttag
1992:1	Svanahuvudsvägen
1992:2	Transportformer i väglöst land
1992:3	Utvärdering av samråden 1989-1990 /skogsbruk – rennärning
1993:2	Virkesbalanser 1992
1993:3	Uppföljning av 1991 års lövträdsplantering på åker
1993:4	Återväxttaxeringarna 1990-1992
1994:1	Plantinventering 89
1995:2	Gallringsundersökning 92
1995:3	Kontrolltaxering av nyckelbiotoper
1996:1	Skogsstyrelsens anslag för tillämpad skogsproduktionsforskning
1997:1	Naturskydd och naturhänsyn i skogen
1997:2	Skogsvårdsorganisationens årskonferens 1996
1998:1	Skogsvårdsorganisationens Utvärdering av Skogspolitiken
1998:2	Skogliga aktörer och den nya skogspolitiken
1998:3	Föryngringsavverkning och skogsbilvägar
1998:4	Miljöhänsyn vid föryngringsavverkning - Delresultat från Polytax
1998:5	Beståndsanläggning
1998:6	Naturskydd och miljöarbete
1998:7	Röjningsundersökning 1997
1998:8	Gallringsundersökning 1997
1998:9	Skadebilden beträffande fasta fornlämningar och övriga kulturmiljövården
1998:10	Produktionskonsekvenser av den nya skogspolitiken
1998:11	SMILE - Uppföljning av sumpskogsskötsel
1998:12	Sköter vi ädellövskogen? - Ett projekt inom SMILE
1998:13	Riksdagens skogspolitiska intentioner. Om mål som uppdrag till en myndighet
1998:14	Swedish forest policy in an international perspective. (Utfört av FAO)
1998:15	Produktion eller miljö. (En mediaundersökning utförd av Göteborgs universitet)
1998:16	De trädbevuxna impedimentens betydelse som livsmiljöer för skogslevande växt- och djurarter
1998:17	Verksamhet inom Skogsvårdsorganisationen som kan utnyttjas i den nationella miljöövervakningen
1998:18	Auswertung der schwedischen Forstpolitik 1997
1998:19	Skogsvårdsorganisationens årskonferens 1998
1999:1	Nyckelbiotopsinventeringen 1993-1998. Slutrapport
1999:2	Nyckelbiotopsinventering inom större skogsbolag. En jämförelse mellan SVOs och bolagens inventeringsmetodik
1999:3	Sveriges sumpskogar. Resultat av sumpskogsinventeringen 1990-1998
2001:1	Skogsvårdsorganisationens Årskonferens 2000
2001:2	Rekommendationer vid uttag av skogsbränsle och kompensationsgödsling
2001:3	Kontrollinventering av nyckelbiotoper år 2000
2001:4	Åtgärder mot markförurning och för ett uthålligt brukande av skogsmarken
2001:5	Miljöövervakning av Biologisk mångfald i Nyckelbiotoper
2001:6	Utvärdering av samråden 1998 Skogsbruk - rennärning
2002:1	Skogsvårdsorganisationens utvärdering av skogspolitikens effekter - SUS 2001
2002:2	Skog för naturvårdsändamål – uppföljning av områdesskydd, frivilliga avsättningar, samt miljöhänsyn vid föryngringsavverkning
2002:3	Recommendations for the extraction of forest fuel and compensation fertilising
2002:4	Action plan to counteract soil acidification and to promote sustainable use of forestland
2002:5	Blir er av
2002:6	Skogsmarksgödsling - effekter på skogshushållning, ekonomi, sysselsättning och miljön
2003:1	Skogsvårdsorganisationens Årskonferens 2002
2003:2	Konsekvenser av ett förbud mot permetrinbehandling av skogsplanter
2004:1	Kontinuitetsskogar - en förstudie
2004:2	Landskapsekologiska kärnområden - LEKO, Redovisning av ett projekt 1999-2003
2004:3	Skogens sociala värden
2004:4	Inventering av nyckelbiotoper - Resultat 2003

Reports 8a-8c from the Swedish National Board of Forestry are the proceedingsvolumes from the scientific part of the ForestSat conference about “Operational Tools in Forestry Using Remote Sensing Techniques”, held in Borås, May 31 – June 1, 2005.

The volumes contains 70 contributions from Europe, Canada and USA, which together gives a picture of the state of the art in research and practice in large area forest remote sensing in Europe and North America.