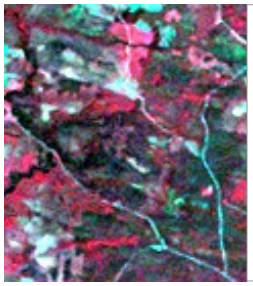
GROUND DATA PROCESSING & PRODUCTION OF THE LEVEL 1 HIGH RESOLUTION MAPS



VALERI 2005

ROVANIEMI site
(forest)

Philippe Rossello

June 2006

CONTENTS

1. Introduction

This report describes the production of the high resolution, level 1, biophysical variable maps for the Rovaniemi site in June 2005 (see campaign report for more details about the site and the ground measurement campaign: annex or http://www.avignon.inra.fr/valeri). Level 1 map corresponds to the map derived from the determination of a transfer function between reflectance values of the LANDSAT image acquired during (or around) the ground campaign, and biophysical variable measurements (LAI-2000 in this case).

The derived biophysical variable maps are:

- Leaf Area Index: LAI corresponds to effective LAI derived from the description of the gap fraction as a function of the view zenith angle;
- cover fraction (fCover) : it is the percentage of soil covered by vegetation between 0° and 7° view zenith angle.

The land cover is mainly composed of forests (spruces and pines). Note that the site is quite flat (for more information, see annex or campaign report: http://www.avignon.inra.fr/valeri).

The site coordinates are described in Table 1:

	GCS_KKJ24North, (units = meters)		Geographic Lat/Lon, WGS-84 (units = degrees)	
	Easting	Northing	Lat.	Lon.
Upper left corner	2558476.2361	7376528.4349	66.474275	25.312311
Lower right corner	2562046.2361	7372448.4349	66.437007	25.390363
Center	2560261.2361	7374488.4349	66.455646	25.351366

Table 1. Description of the site coordinates: they correspond to LANDSAT image coordinates.

2. Available data

2.1. LANDSAT Image

The LANDSAT 5 TM image was acquired the 19th June 2005 while the ground measurements were carried out from 13/06/2005 to 17/06/2005. The initial projection was UTM 35 North, WGS-84 (please, refer to the campaign report for more details: annex or http://www.avignon.inra.fr/valeri). The LANDSAT image was radiometrically and geometrically corrected by Eurimage (system corrected: product 1G). In order to reduce the residual error in the systematic 1G product[1], a rectification was performed from the SPOT image 2004 and the ground control points (GCP)[2]: -230 meters in Easting and +220 meters in Northing. No atmospheric correction was applied to the image since no atmospheric data were available. However, as the LANDSAT image is used to compute empirical relationships between reflectance and biophysical variable, we can assume that the effect of the atmosphere is the same over the whole 3.5 x 4 km site. Therefore, it will be taken into account everywhere in the same way.

Figure 1 shows the relationship between Red and near infrared (NIR) LANDSAT channels: the soil line is well marked and no saturated points are observed.

[1] http://www.eurimage.com/products/landsat.html
[2] Note that a very slight deviation remains from the topographic map.

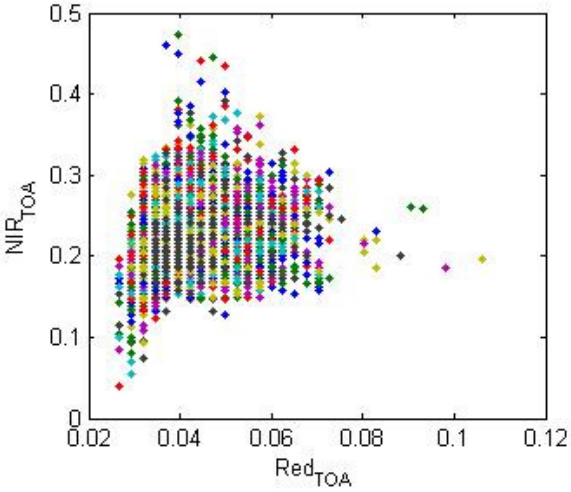


Figure 1. Red/NIR relationship[3] on the LANDSAT image for Rovaniemi, 2005.

2.2. LAI-2000 measurements

The biophysical variables (LAI, fCover) were estimated by LAI-2000 instrument. The measurements have been acquired below canopy. According to the sampling protocol, 48 measurements were taken for each ESU. However, in the VALERI context, we are interested in the whole leaf area index, therefore, the ESU biophysical variables that are used in the following were computed as:

- LAI = LAI_canopy + LAI_ground
- fCover is the percentage of soil covered by vegetation at 7° view zenith angle (ground level).

Figure 2 shows the distribution of the different measured variables over the sampled ESUs. LAI values vary from 0.34 to 2.86 and fCover values vary from 0.139 to 0.797. This range shows a heterogeneous site in terms of LAI.

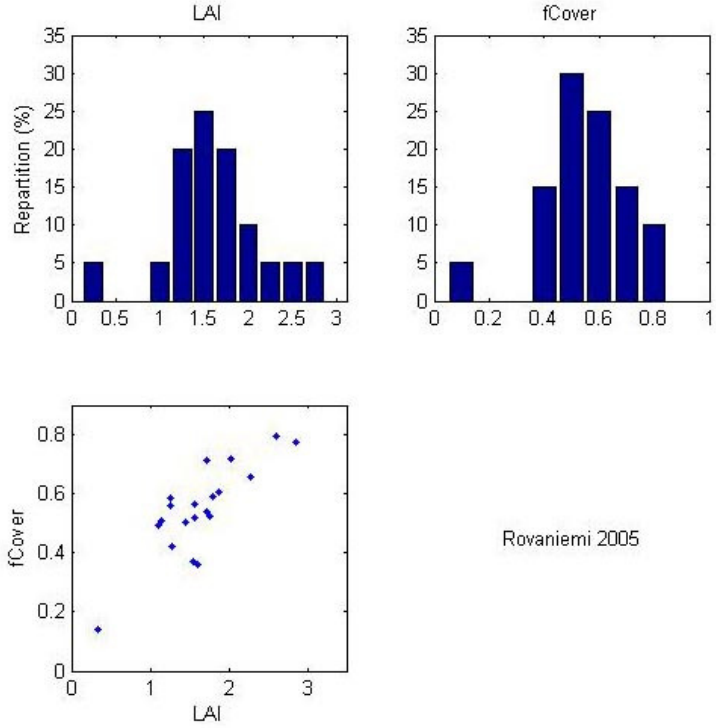


Figure 2. Distribution of the measured biophysical variables over the ESUs.

[3] The effect observed in the scatter plot (discretization on Red_{TOA} axis) is only due to the application of stretch procedures in digital LANDSAT image. The discretization step caused by original DNs of the LANDSAT image is finer on NIR_{TOA} axis than on Red_{TOA} axis.

2.3. Sampling strategy

2.3.1. Principles

The sampling strategy is defined in the campaign report: http://www.avignon.inra.fr/valeri. It was attempting to represent as much as possible the range of variation of canopy types and conditions. The sampling of each ESU is based on twelve measurements.

Figure 3 shows that the 20 ESUs are evenly distributed over the site (3.5 x 4 km). Note that the processing of the ground data has shown that all the ESUs have been kept for the computation of the transfer function.

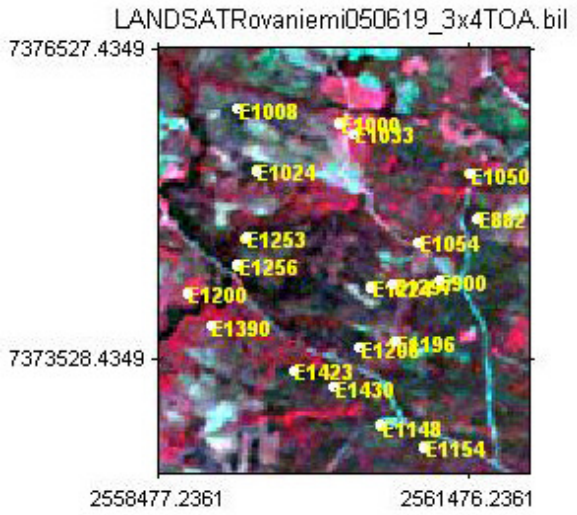


Figure 3. Distribution of the ESUs around the Rovaniemi site.

2.3.2. Evaluation based on NDVI values

The sampling strategy is evaluated using the LANDSAT image by comparing the NDVI distribution over the site with the NDVI distribution over the ESUs (Figure 4). As the number of pixels is drastically different for the ESU and whole site (WS = 10000 in case of a 3 x 3 km LANDSAT image at 30 m resolution), it is not statistically consistent to directly compare the two NDVI histograms. Therefore, the proposed technique consists in comparing the NDVI cumulative frequency of the two distributions by a Monte-Carlo procedure which aims at comparing the actual frequency to randomly shifted sampling patterns. It consists in:

1. computing the cumulative frequency of the N pixel NDVI that correspond to the exact ESU locations;
2. then, applying a unique random translation to the sampling design (modulo the size of the image);
3. computing the cumulative frequency of NDVI on the randomly shifted sampling design;
4. repeating steps 2 and 3, 199 times with 199 different random translation vectors.

This provides a total population of $N = 199 + 1$(actual) cumulative frequency on which a statistical test at acceptance probability $1 - \alpha = 95\%$ is applied: for a given NDVI level, if the actual ESU density function is between two limits defined by the $N\alpha/2 = 5$ highest and lowest values of the 200 cumulative frequencies, the hypothesis assuming that WS and ESU NDVI distributions are equivalent is accepted, otherwise it is rejected.

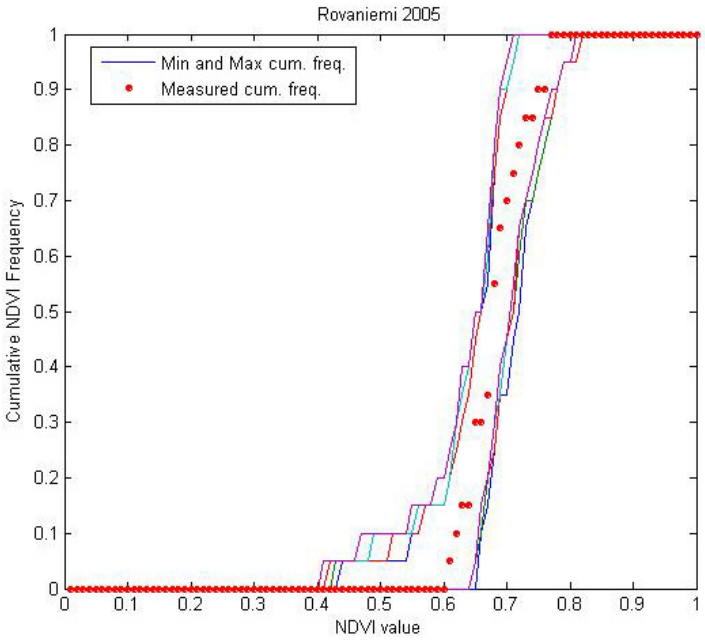


Figure 4. Comparison of the ESU NDVI distribution and the NDVI distribution over the whole image.

Figure 4 shows that the NDVI distribution of the 20 ESUs is very good over the whole site. Note that NDVIs lower than 0.61 have not been sampled although they are present in the image. They may correspond to roads, paths, but also sparse forest…

2.3.3. Evaluation based on classification

A non supervised classification based on the k_means method (Matlab statistics toolbox) was applied to the reflectance of the LANDSAT image to distinguish if different behaviours on the image for the biophysical variable-reflectance relationship exist.

A number of 4 classes was chosen (Figure 5). The distribution of the classes on the image and on the ESUs is not similar, except classes 1 and 2. The class 3 is indeed very under-represented (≈15% against 40%), while the class 4 appears to be very over-sampled (≈54% against 32%).

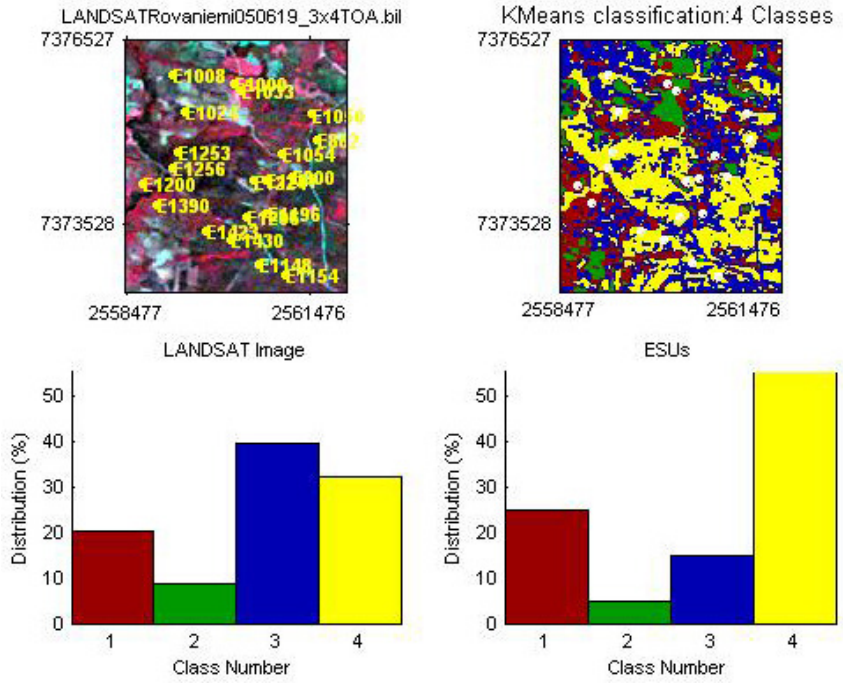


Figure 5. Classification of the LANDSAT image. Comparison of the class distribution between the LANDSAT image and sampled ESUs.

Figure 6 shows the different relationships observed between the biophysical variables and the corresponding NDVI on the ESUs, as a function of the LANDSAT classes determined from non supervised classification.

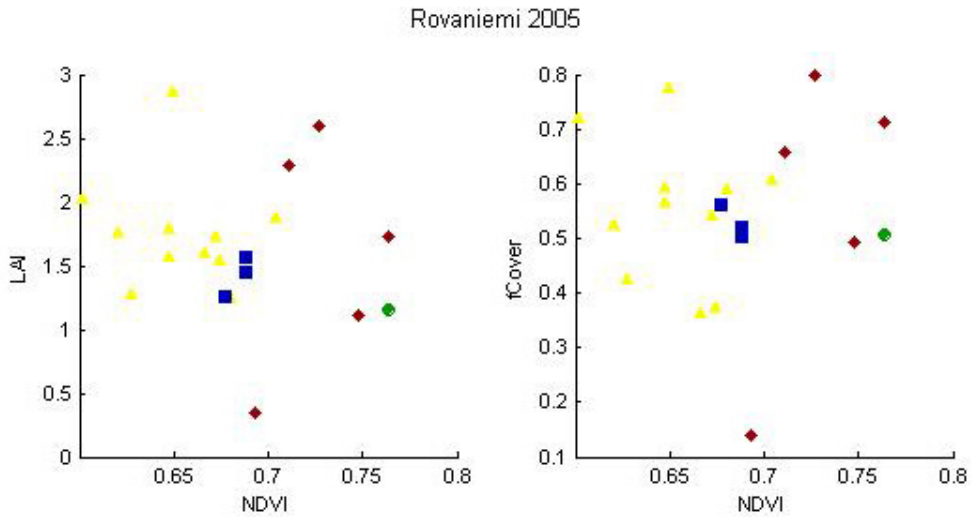


Figure 6. NDVI-biophysical variable relationships as a function of LANDSAT classes

There is no relation between NDVI and biophysical variables. However, even if the relationship between LAI and NDVI is not coherent, two different transfer functions will be generated in order to analyse the results. Note that as the number of ESUs belonging to the class 2 is insufficient, the average value of the biophysical variable measured will attribute (§3.1).

2.3.4. Using convex hulls

A test based on the convex hulls was also carried out to characterize the representativeness of ESUs. Whereas the evaluation based on NDVI values uses two bands (red and NIR), this test uses the four bands (green, red, NIR and SWIR in this case) of the LANDSAT image. A flag image, is computing over the reflectances (Figure 7). The result on convex-hulls can be interpreted as:

- pixels inside the 'strict convex-hull': a convex-hull is computed using all the LANDSAT reflectance corresponding to the ESUs belonging to the class. These pixels are well represented by the ground sampling and therefore, when applying a transfer function the degree of confidence in the results will be quite high, since the transfer function will be used as an interpolator;
- pixels inside the 'large convex-hull': a convex-hull is computed using all the reflectance combination (±5% in relative value) corresponding to the ESUs. For these pixels, the degree of confidence in the obtained results will be quite good, since the transfer function is used as an extrapolator (but not far from interpolator);
- pixels outside the two convex-hulls: this means that for these pixels, the transfer function will behave as an extrapolator which makes the results less reliable. However, having a priori information on the site may help to evaluate the extrapolation capacities of the transfer function.

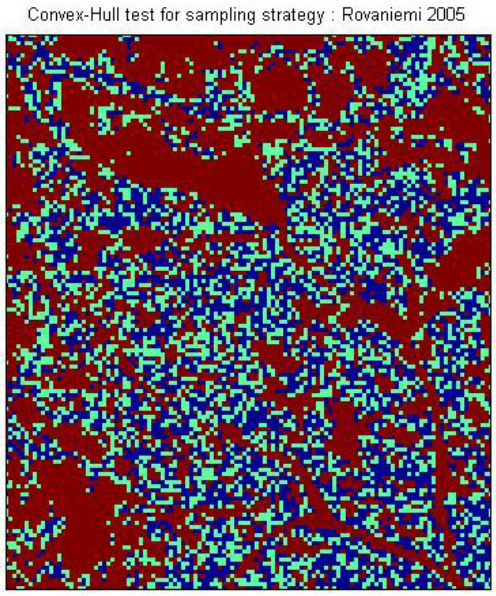


Figure 7. Evaluation of the sampling based on the convex hulls. The map is shown at the bottom: blue and light blue correspond to the pixels belonging to the 'strict' and 'large' convex hulls and red to the pixels for which the transfer function is extrapolating.

This map shows that the pixels outside the two convex-hulls are numerous. They mainly correspond to bare soil (roads and paths), lowest NDVI values (including pine or spruce forest)… This is due to the fact that the distribution of reflectances corresponding to lowest NDVI values (< 0.61) is not well represented by the sampling and therefore, the range of these reflectance values is not taken into account to calculate the convex hulls.

3. Determination of the transfer function for the two biophysical variables: LAI, fCover

3.1. The transfer function considered

For each class determined in §2.3, the following transfer functions were tested:

- AVE: if the number of ESUs belonging to the class is too low. The transfer function consists only in attributing the average value of the biophysical variable measured on the class to each pixel of the LANDSAT image belonging to the class;

- REG: if the number of ESUs is sufficient, multiple robust regression between ESUs reflectance (or Simple Ratio) and the considered biophysical variable can be applied: we used the 'robustfit' function from the Matlab statistics toolbox. It uses an iteratively re-weighted least squares algorithm, with the weights at each iteration computed by applying the bisquare function to the residuals from the previous iteration. This algorithm provides lower weight to ESUs that do not fit well. The results are less sensitive to outliers in the data as compared with ordinary least squares regression. At the end of the processing, three errors are computed: classical root mean square error (RMSE), weighted RMSE (using the weights attributed to each ESU) and cross-validation RMSE (leave-one-out method).

The 'AVE' function is applied to the classes 2. For the classes 1, 3 and 4, the 'REG' function is tested using either the reflectance or the logarithm of the reflectance for any band combination as well as the simple ratio or NDVI. As the method has poor extrapolation capacities, a flag image, based on the convex hulls is computing over reflectances.

3.2. Results

3.2.1. Choice of the method

For all the ESUs that do not correspond to class 2, a unique transfer function is computed. Figure 8 shows the results obtained for all the possible band combinations using either the reflectance (ρ) or the logarithm of the

reflectance (log(ρ)): for LAI and fCover, the results using the reflectance are selected. Depending on the biophysical variable, note that the regression made on the logarithm of the reflectance sometimes provides close results. The Red*NIR ('+' or RN) combination is added to all the band combinations (except for NDVI and SR). Please read the document: "A method to improve the relation between the biophysical variables" (http://www.avignon.inra.fr/valeri/table_methods/new_linear.pdf).

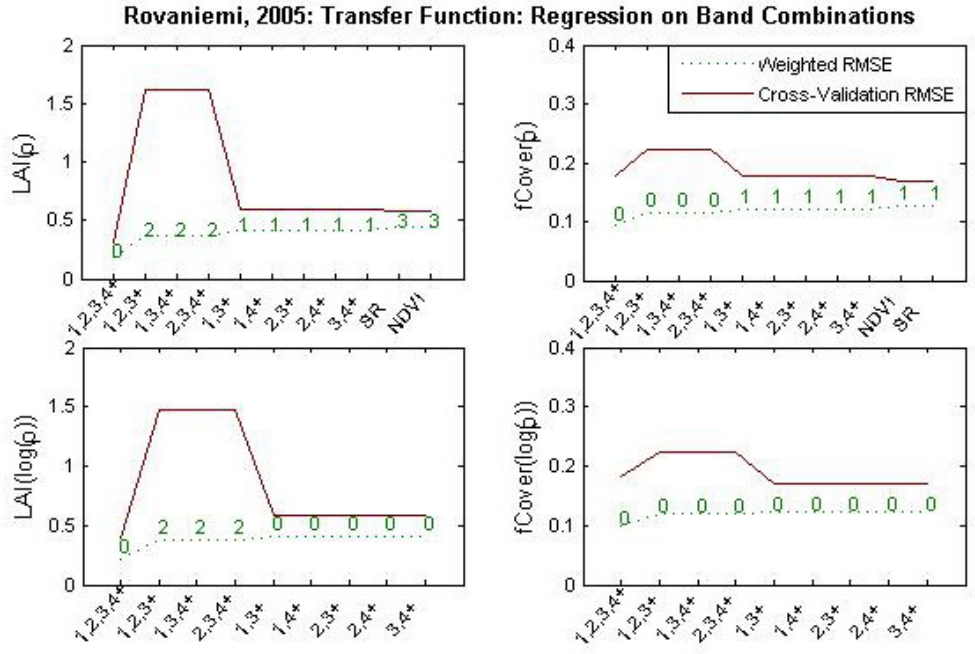


Figure 8. Transfer function: test of multiple regression applied on different band combinations. Band combinations are given in abscissa. The estimated biophysical variable is given in ordinate. Top graphs correspond to regression made on reflectance (ρ): the weighted root mean square error (RMSE) is presented in green along with the cross-validation RMSE in red. The numbers indicate the number of data used for the robust regression with a weight lower than 0.7 that could be considered as outliers. Bottom graphs correspond to regression made on the logarithm of the reflectance.

3.2.2. Choice of the band combination

For the LAI, the green, red, NIR, SWIR, RN band combination (Figure 9 an Figure 10) on reflectance was selected since it provides the lowest cross-validation RMSE value, the lowest weighted root mean square error value and zero weight lower than 0.7.

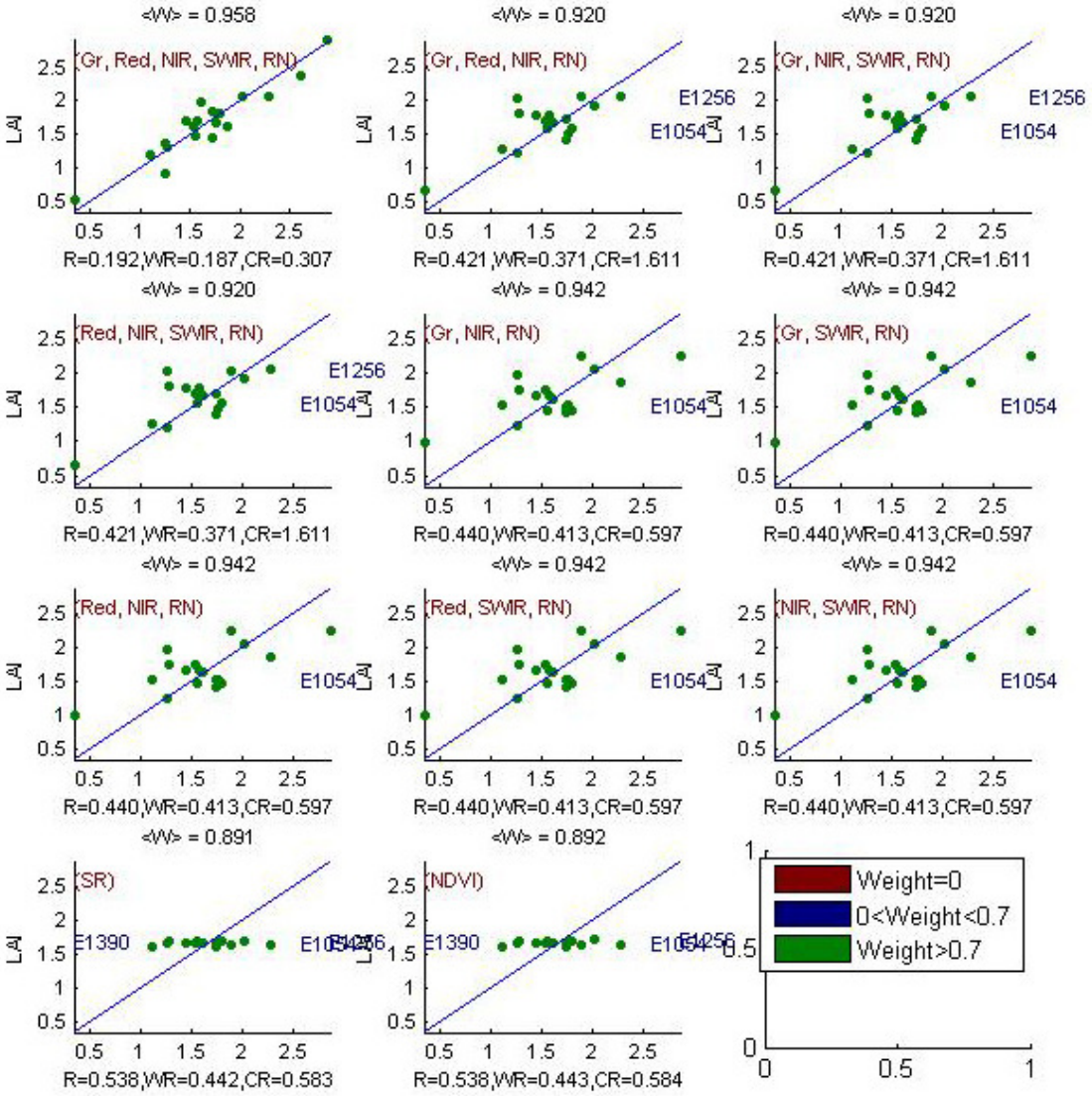


Figure 9. Leaf Area Index: results for regression on reflectance using different band combinations. R is the root mean square error computed between LAI and estimated LAI. WR is the weighted root mean square error and CR is the cross validation root mean square error.

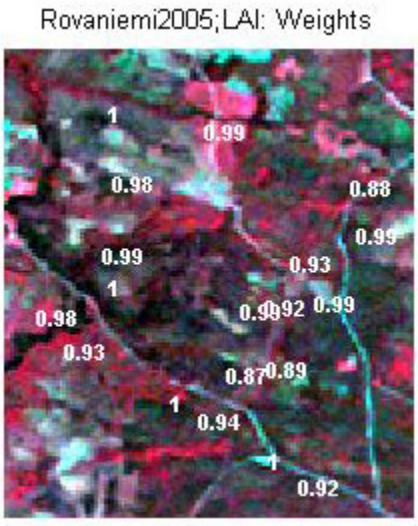


Figure 10. Weights associated to each ESU for the determination of LAI transfer function.

For the fCover, the green, red, NIR, SWIR, RN band combination (Figure 11 and Figure 12) on reflectance was selected since it provides the lowest weighted root mean square error value, zero weight lower than 0.7 and a cross-validation RMSE value among the lowest.

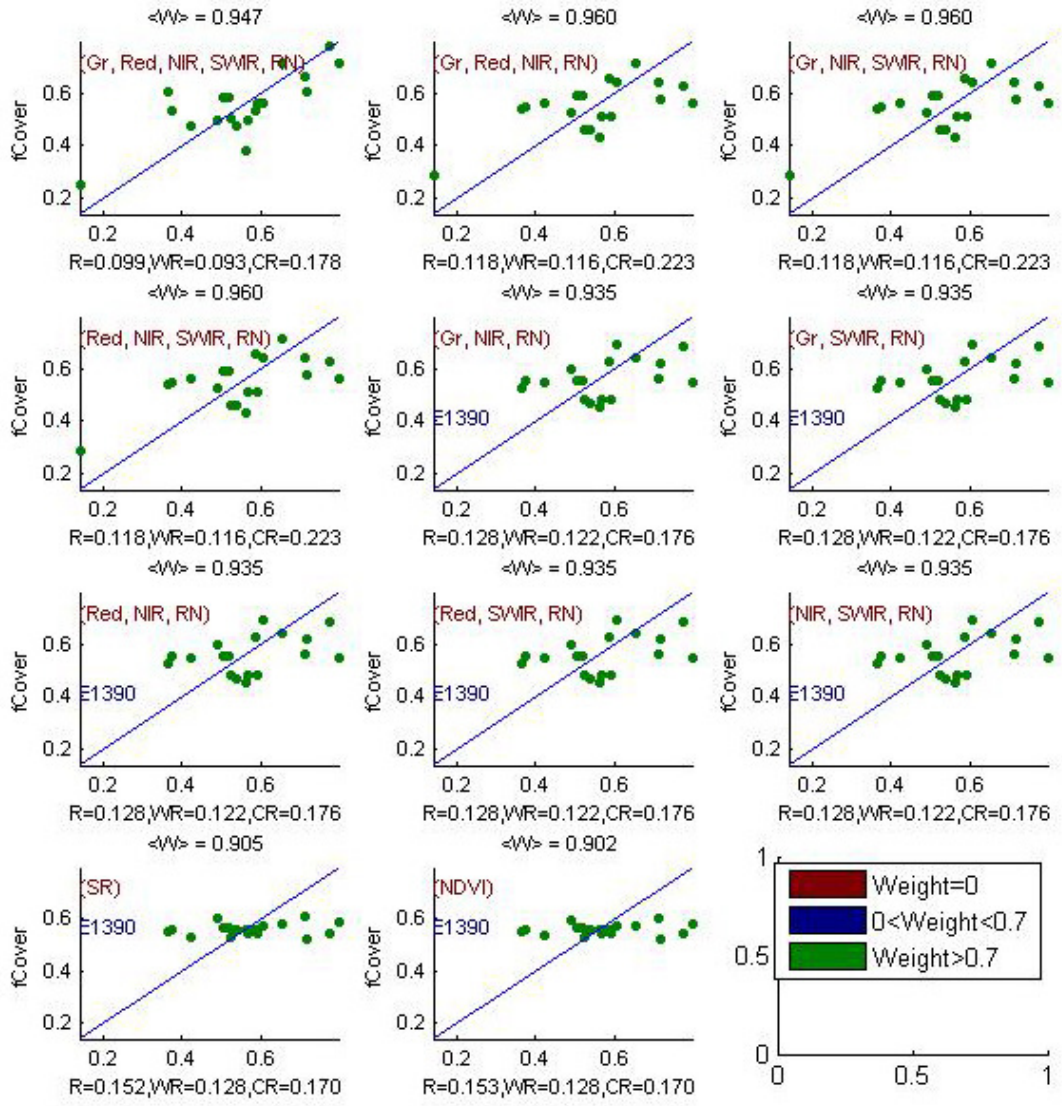


Figure 11. fCover: results for regression on reflectance using different band combinations. R is the root mean square error computed between fCover and estimated fCover. WR is the weighted root mean square error and CR is the cross validation root mean square error.

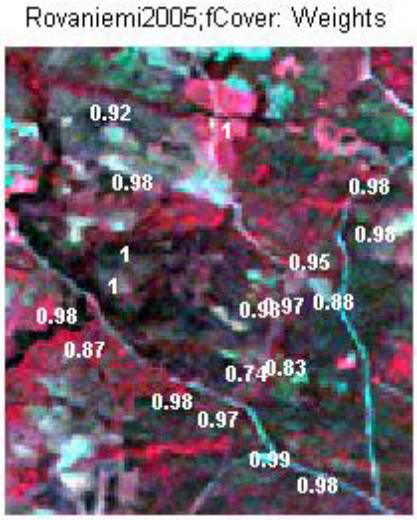


Figure 12. Weights associated to each ESU for the determination of fCover transfer function.

Following, the results of the transfer function (Table 2):

Variable	Band Combination	RMSE	Weighted RMSE	Cross-valid RMSE
LAI	-2.1161 - 44.803(green) + 221.2(red) + 43.352(NIR) - 46.707(SWIR) - 866.04(RN)	0.192	0.187	0.307
fCover	-0.85857 - 20.448(green) + 68.067(red) + 16.256(NIR) - 8.5253(SWIR) - 334.46(RN)	0.099	0.093	0.178

RN = Red*NIR

Table 2. Transfer function applied to the whole site for LAI and fCover, and corresponding errors

3.3. Applying the transfer function to the Rovaniemi LANDSAT image extraction

Figure 13 presents the biophysical variable maps obtained with the transfer function described in Table 2 for the classes 1, 3 and 4. The average value of the ESUs belonging to class 2 is applied to the pixels belonging to this same class (orange in the flag images). The maps obtained for the two variables are consistent, showing similar patterns: low LAI values where low fCover are observed and conversely…

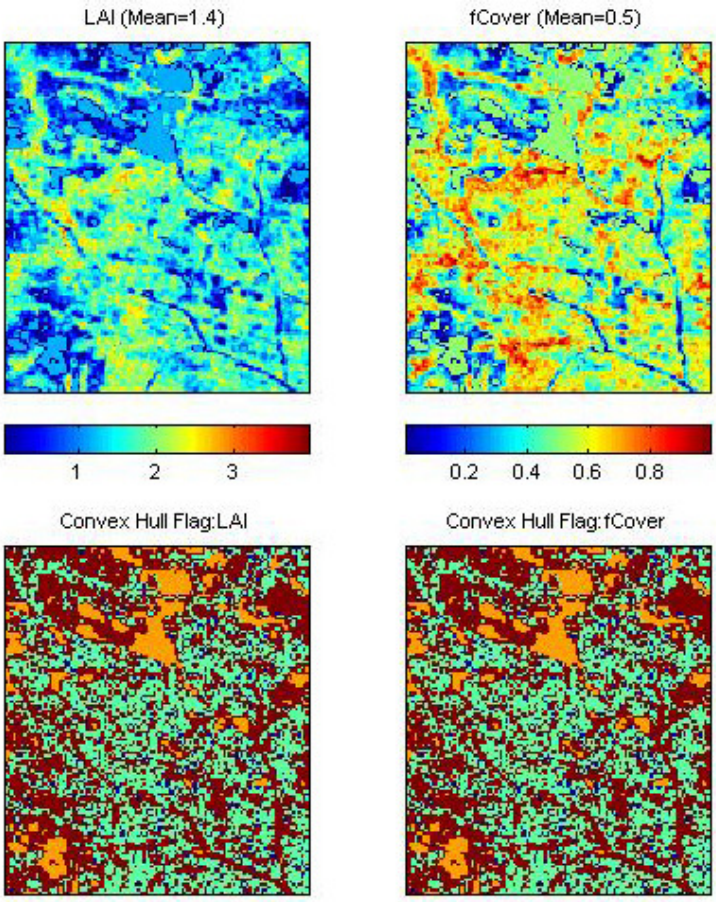


Figure 13. High resolution biophysical variable maps applied on the Rovaniemi site (top). Associated Flags are shown at the bottom: blue and light blue correspond to the pixels belonging to the 'strict' and 'large' convex hulls, red to the pixels for which the transfer function is extrapolating and orange to the pixels for which the 'AVE' transfer function is applied.

The flag maps are comparable between the different biophysical variables. Note that the pixels outside the two convex-hulls are numerous. They mainly correspond to bare soil and lowest NDVI values (2.3.4).

4. Conclusion

The 'REG' method is applied to the classes 1, 3 and 4 by using 20 ESUs, whereas the 'AVE' method is applied to the class 2. The representativeness of the land cover of the different ESUs is not very good since the bare soil and the low NDVI values are not well represented. However, the evaluation based on NDVI values 2.3.2) is very satisfactory. The results of the robust regression are good and the maps obtained for the biophysical variables are consistent. The flag associated to each map show that the extrapolation of the transfer function is mainly bounded to bare soil and lowest NDVI pixels (including forest pixels). For the two variables, the regression coefficients are computed by relating the variable itself to reflectance.

The biophysical variable maps are available in GCS_KKJ24North projection coordinates at 30m resolution.

5. Acknowledgements

We want to thank: **Sanna Ervasti** (University of Helsinki) and **Pekka Voipio** (Finnish Forest Research Institute) for the organisation and participation to the campaign, but also **Miina Rautiainen** (University of Helsinki) for the compilation of campaign report and **Terhikki Manninen** (Finnish Meteorological Institute) for his advices about the LANDSAT image rectification.

ANNEX

Ground measurement acquisition report for the VALERI site **Rovaniemi**

sampled from 13.6 to 17.6.2005

Organization: Finnish Forest Research Institute,
University of Helsinki
email: pekka.voipio@metla.fi, pauline.stenberg@helsinki.fi,
miina.rautiainen@helsinki.fi

Date of report 3.1.2006

People participating in the field experiment:

Name	Organization
Pekka Voipio	Finnish Forest Research Institute
Sanna Ervasti	University of Helsinki
Miina Rautiainen*	University of Helsinki

(*compilation of report only)

Site coordinates

	Lat-Long WGS84 (Deg min.00)		**Other projection*** GCS_KKJ_24_N	
	Lat.	Long.	Easting	Northing
Upper left corner	66.471580	25.321713	2559080	7376190
Lower right corner	66.437682	25.364904	2561087	7372452

*The other projection used is GCS_KKJ, 24 North. All the characteristics are provided in the following table:

Geodesic Map Datum D-KKJ		**Map Projection**	
Associated Ellipsoïd	Int. 1924	Latitude of origin	0.000000
Semi-major axe	6378388.000	Longitude of origin	24.000000
Semi-minor axe	6356911.9461279465	Parallels 1st	
1/flattening	297.000000	2nd	
Eccentricity	0.08199189	Xo: false easting	2500000.0000
		Yo: false northing	0.0000
		Scale factor	1.0000

Ground control points

# Name	**Easting (m)**	**Northing (m)**	**Comments on the vegetation status, condition of acquisitions, etc...**
GCP1	2550233	7366851	Lake Louejarvi, SW-river outlet
GCP2	2556626	7382321	Northern lake, the biggest of four ponds, diam. appr. 600 m, SW-corner that makes a 90 degree turn
GCP3	2564735	7370879	Island in river Kemijoki, SW-corner
GCP4	2561851	7367353	Road by the riverside, the middle point of the bridge crossing a side river.

*This is extracted from the Excel file GPSSIteNameYear.xls

GPS system used: Trimble Geo XT
Typical uncertainty of GPS position: 1 m with afterwards differential correction.

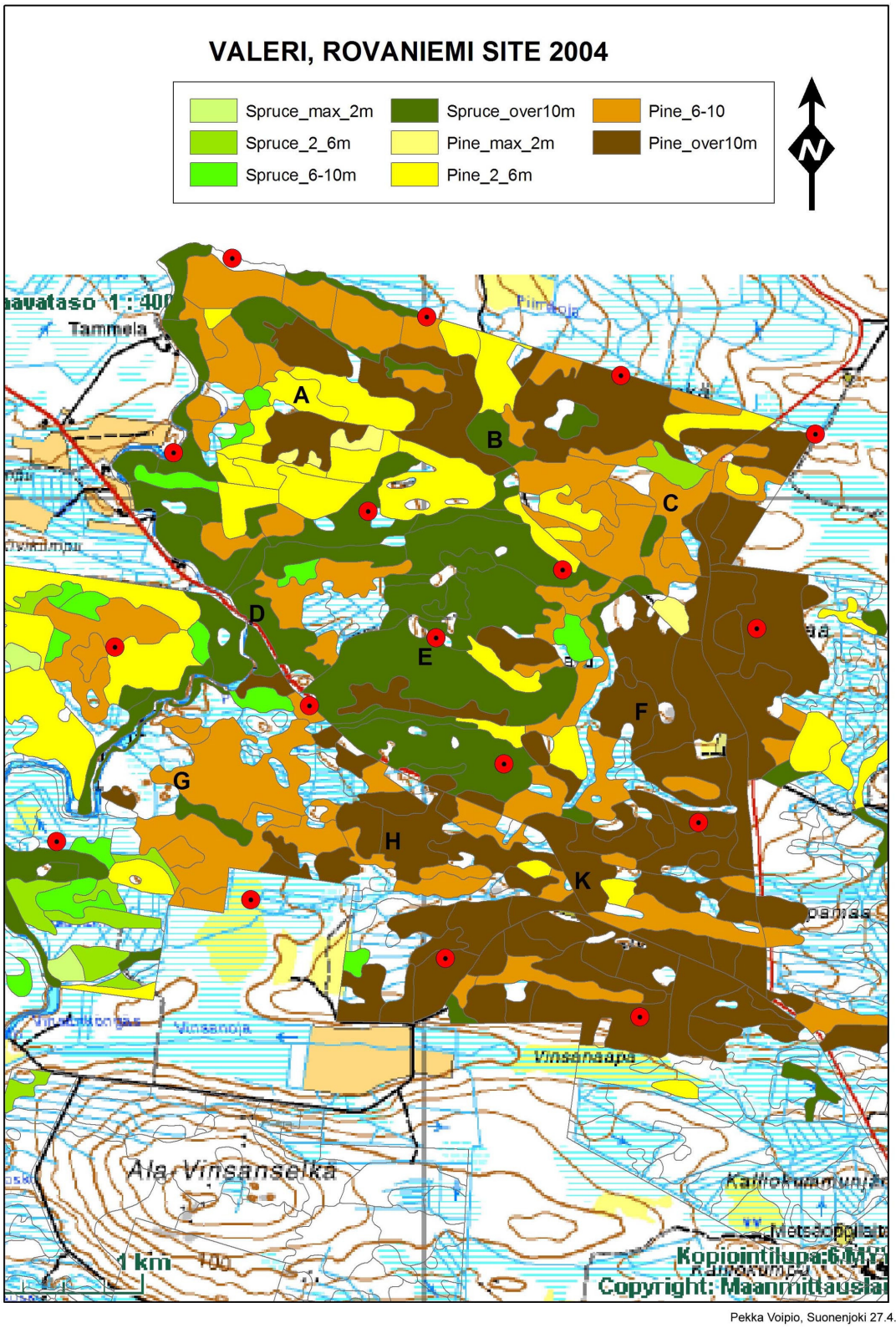
Description of the site and land cover

Category according to IGBP classification

Needle-leaved evergreen forest.

Comments on the land cover

Tree species and size distribution over the 3 x 3 km test site:



Pekka Voipio, Suonenjoki 27.4. 2004

Topography

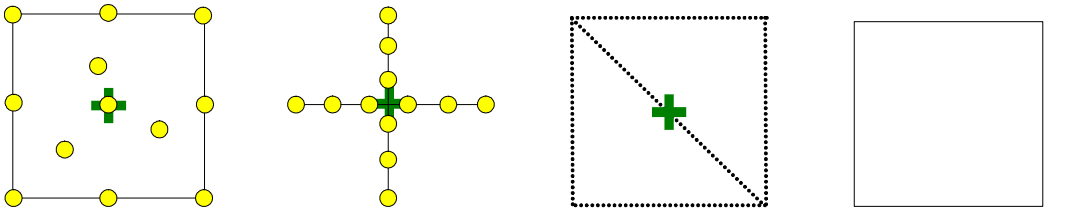
Flat.

Spatial Sampling scheme

Sensors used for sampling the ESUs

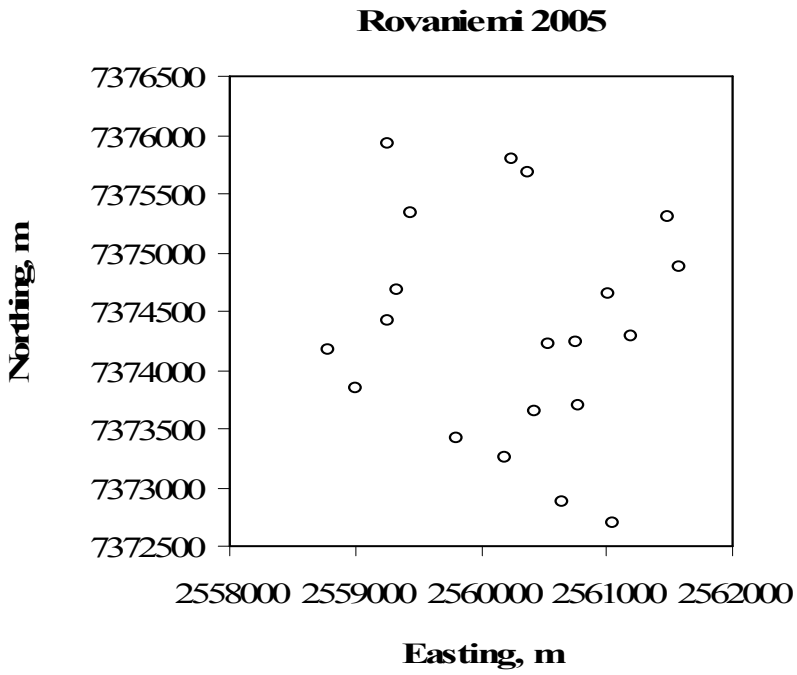
	Method	Comments
x	LAI-2000	Below canopy (height ca 1m)
☐	TRAC	
☐	Ceptometer	
☐	Direct measurements	
x	Other	Basic stand inventory

Sampling strategy for the ESU



a ☐ **b** x **c** ☐ **d** ☐ (specify)

Distribution of the Elementary sampling units



The high spatial resolution image

Satellite

Satellite used: Landsat 5 TM
Level of processing: 1G
Projection type: UTM, WGS-84
Acquisition date: 19/06/2005

List of the ESUs in 2005

Plot#	Easting(m)	Northing(m)	Vegetation
882	2561574	7374885	Scots pine (*Pinus sylvestris* L.) age 55 years, h 10 m
900	2561195	7374284	Scots pine, thinned stand, age 70 y, h 14 m
1000	2560240	7375790	Mixed seedling stand, pine 45 %, spruce (*Picea abies*) 45 %, birch (*Betula pendula*) 10 %
1008	2559256	7375920	Mixed Spruce 80%, Pine 10%, Birch 10 %
1024	2559442	7375343	Scots pine, age 50y, h 6 m
1033	2560371	7375686	Pine-birch seedling stand, age 70 y, h 15 m
1050	2561493	7375297	Scots pine, age 65 y, h 16 m
1054	2561005	7374652	Scots pine, age 20 y, h 7 m
1148	2560641	7372883	Scots pine, age 55 y, h 13 m
1154	2561056	7372699	Scots pine, age 60 y, h 12 m
1196	2560771	7373698	Pine, age 55 y, 6 m, understory small dwarf birches (*Betula nana* L.)
1200	2558777	7374165	Scots pine, 20 y, 3.5 m
1206	2560435	7373642	Mixed, pine 45 %, spruce 45 %, birch 10 %, age 150 y, h 15
1224	2560546	7374218	Mixed, pine 45 %, spruce 45 %, birch 10 %, age 160 y, h 16 m; forest path crossing the area
1253	2559341	7374684	Mixed, spruce 60 5, pine 20 %, birch 20 %
1256	2559265	7374420	Spruce 95 %, birch 5 %, age 160 y, 15 m; many dead trees
1297	2560760	7374242	Logdepole pine (*Pinus contorta*) planted in 1978, h 7 m; ditches
1390	2559009	7373848	Scots pine, age 110 y, h 7 m
1423	2559800	7373420	Scots pine, age 50 y, 12 m
1430	2560187	7373257	Scots pine, age 40, h 11 m

(y=year, h=height)

Photo gallery

Please see photos from year 2004.