VALERI summer 2003 : Järvselja site (boreal forest)

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

Philippe Rossello, Tiit Nilson

1 Introduction

This report describes the production of the high resolution, level 1, biophysical variable maps for the Järvselva site in summer 2003 (see campaign report for more details about the site and the ground measurement campaign : http://www.avignon.inra.fr/valeri). Level 1 map corresponds to the map derived from the application of a transfer function between reflectance values of the SPOT 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 site is mostly covered by a sub-boreal mixed forest of different age, including both conifers (Scots pine and Norway spruce) and deciduous (birch, aspen, alder). Agricultural fields are almost missing, however, a few unmanaged open areas are found (Figure 1). The site coordinates are described in Table 1:

	Lambert-Est-92 WGS84 (meters)	
	Lat.	Lon.
Upper left corner	6468740.0000	689160.0000
Lower right corner	6464740.0000	693000.0000

Table 1. Description of the site coordinates.

In fact, two series of measurements were conducted, a 'midsummer' series with full green leaves, and a 'winter' series, where the deciduous trees were leafless. The midsummer series was performed in two parts: one from 7 to 11 July 2003, and another from 12 to 13 August 2003.

2 Available data

2.1 Sampling strategy

The sampling strategy was attempting to represent as much as possible the range of variation of canopy types and conditions. In addition, some ESUs were organised within a cross pattern at the centre of the site to be able to get geostatistic estimates.

Figure 1 shows that the ESU locations are well spatially distributed over the site (3x3 km). The processing of the ground data has shown that:

- ESUs 43 et 44[1] (in black on Figure 1) were located on a small plot with a strong heterogeneity on the borders. Those two ESUs were eliminated.
- Considering that SPOT geo-location and GPS measurements are associated to errors, we found that processed LAI for ESUs 28, 30, 37, 38 and 39 did not correspond to the SPOT pixel in terms of reflectance as compared to the knowledge of the land use (data base kindly provided by Tiit Nilson): they have been shifted by 1 to 3 pixels, considering that the LAI measurements were performed with LAI-2000 (maximum zenith angle = 68°) and that the tree height is few tens of meters.

Finally 56 ESUs have been kept for the computation of the transfer function (Figure 1).

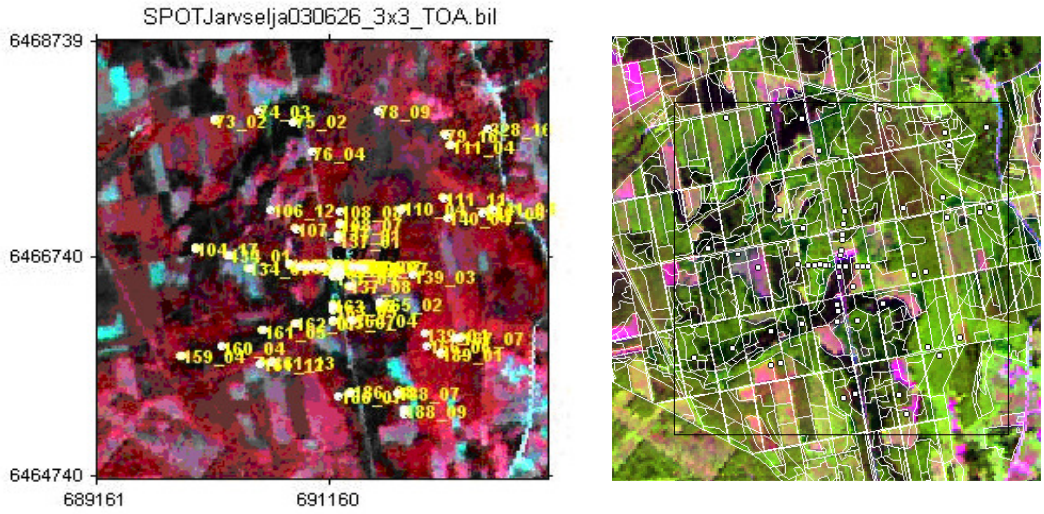

Figure 1. Distribution of the ESUs around the Järvselja site. ESUs in black were eliminated for the computation of the transfer function (left) and a false-color SPOT4[2] (XS2, XS3, XS4) image (26 June 2003), together with the superimposed stand borders and locations of LAI measurement sample plots (right). Coniferous forests appear in dark, deciduous forests in green tones, recent clear cuts and open areas in pink and yellowish.

[1] Stand id: 137_03.
[2] From Järvselja campaign report.

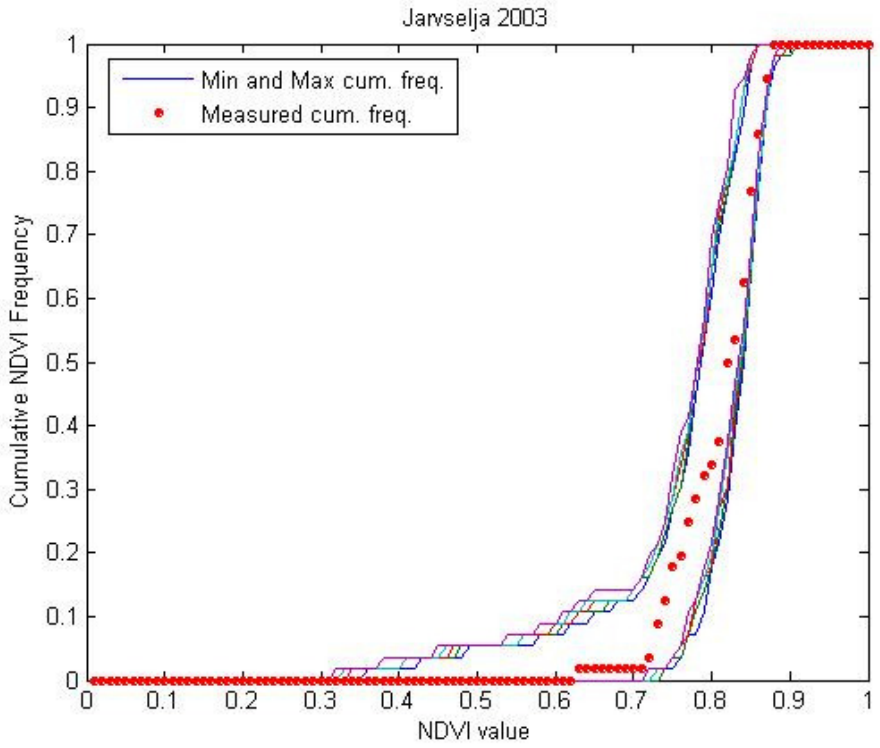

Figure 2. Comparison of the ESU NDVI distribution and the NDVI distribution over the whole image.

The sampling strategy is evaluated using the SPOT image by comparing the NDVI distribution over the site with the NDVI distribution over the ESUs (Figure 2). As the number of pixels is drastically different for the ESU and whole site (WS=22500 in case of a 3x3 km SPOT image), 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 2 shows that the NDVI distribution of the 56 ESUs is quite good over the whole site (comprised between the 5 highest and lowest cumulative frequencies) even if the cumulative frequency curve is close to the boundaries for high NDVI values. NDVIs lower than 0.62 have not been sampled although they are present in the image. Moreover, the site is quite homogeneous in terms of NDVI since the highest and lowest distributions are close.

2.2 SPOT image

The SPOT image was acquired the 26th June 2003 by HRVR1 on SPOT4. This image was geo-referenced and atmospherically corrected by the TARTU Observatory. The projection is Lambert-Est-92 (Lambert Conformal Conic 2 parallel). Please, refer to the campaign report for more details: http://www.avignon.inra.fr/valeri.

Figure 3 shows the relationship between RED and near infrared (NIR) SPOT channels. There is no saturation of the signal. Reflectance values in the NIR are quite high since the Järvselja site is mainly composed of forests.

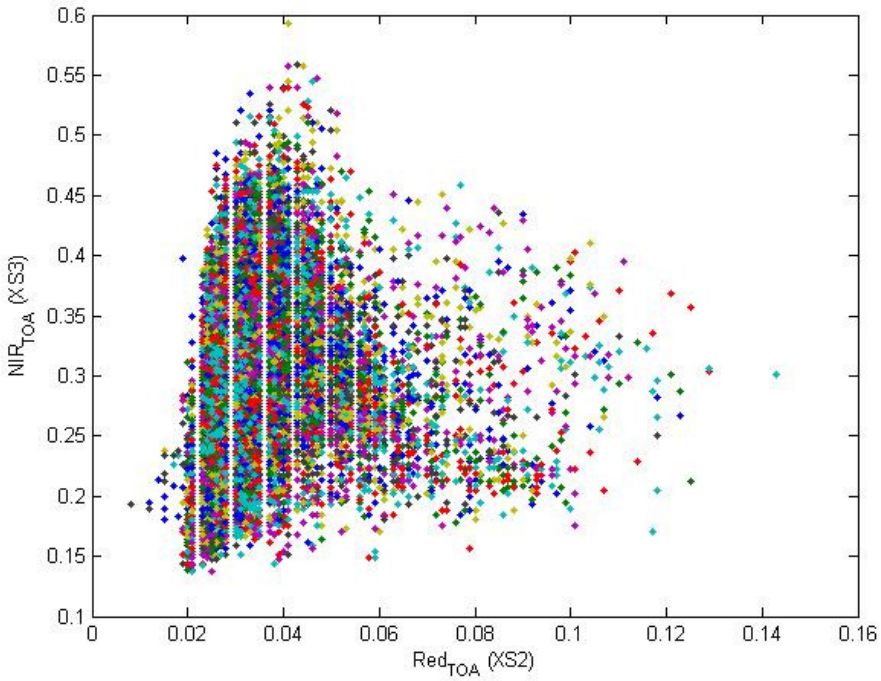

Figure 3. Red/NIR relationship[3] on the SPOT image for Järvselva, 2003.

A non supervised classification based on the k_means method (Matlab statistics toolbox) was applied to the reflectance of the SPOT image to distinguish if different behaviours on the image for the biophysical variable-reflectance relationship exist. A number of 5 classes were chosen (Figure 4). The distribution of the classes on the image and on the ESUs is rather different. Class 1 and class 3 appear to be over-sampled whereas classes 2, 4 and 5 are under-represented.

[3] The effect observed in the scatter plot (discretization on XS2 axis) is only due to the application of stretch procedures in digital SPOT image. The discretization step caused by original DNs of the SPOT image is finer on XS3 axis than on XS2 axis.

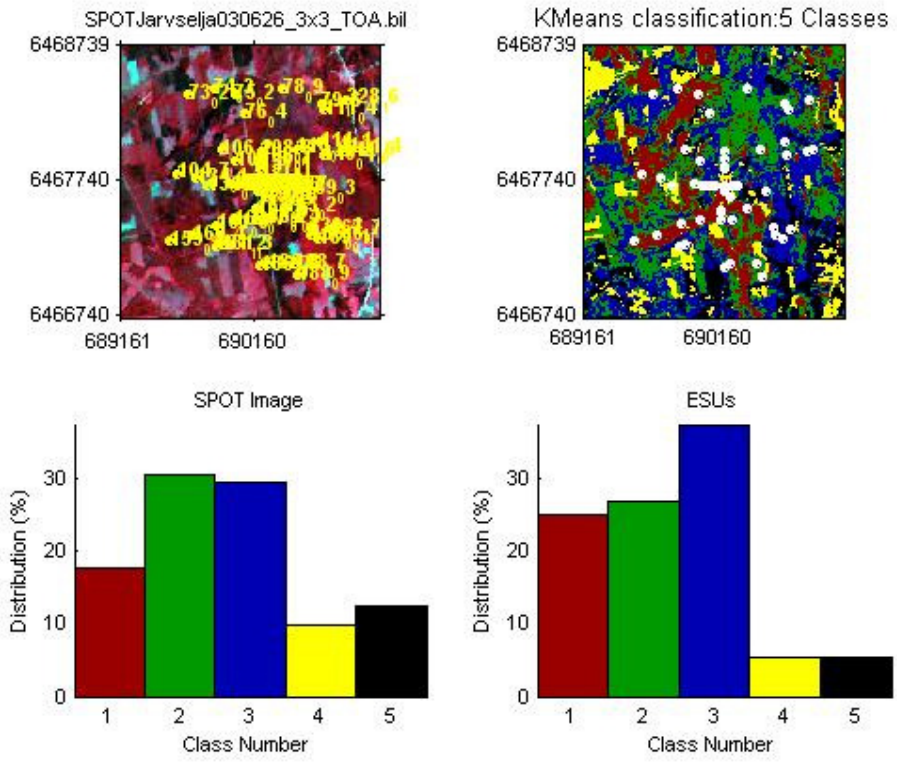

Figure 4. Classification of the SPOT image. Comparison of the class distribution between the SPOT image and sampled ESUs.

2.3 LAI-2000 measurements

The physical variables (LAI, fCover) were estimated by LAI-2000 instrument. Nilson's algorithm[4] was used to process data (Excel version of algorithm is available on the website: http://www.avignon.inra.fr/valeri).

The measurements have been acquired at two heights: ground level and breast height (1.30m). The two levels allow the distinction between understorey and trees. According to the sampling protocol, 48 measurements were taken at the both level 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 5 shows the distribution of the different measured variables over the sampled ESUs. LAI values vary between 1 and 7, which shows a quite heterogeneous site in terms of LAI but most of the ESUs have LAI>4.

To build the relationships between biophysical variables and SPOT data, the reflectance of a given ESU was considered as the average reflectance over the central pixel + the 8 surrounding pixels. This takes into account the fact that the height of the trees are about 16m and consequently the fish-eye of the LAI-2000 instrument observes an area of $\pi x[16x\tan(68°)]^2 \cong 4900m^2$, *i.e.* close to the area of 9 SPOT pixels (=3600m²) when using a maximum view zenith angle of 68°.

[4] Nilson, T. (1999). Inversion of gap frequency data in forest stands. Agric. For. Meteorol., 98-99:437-448.

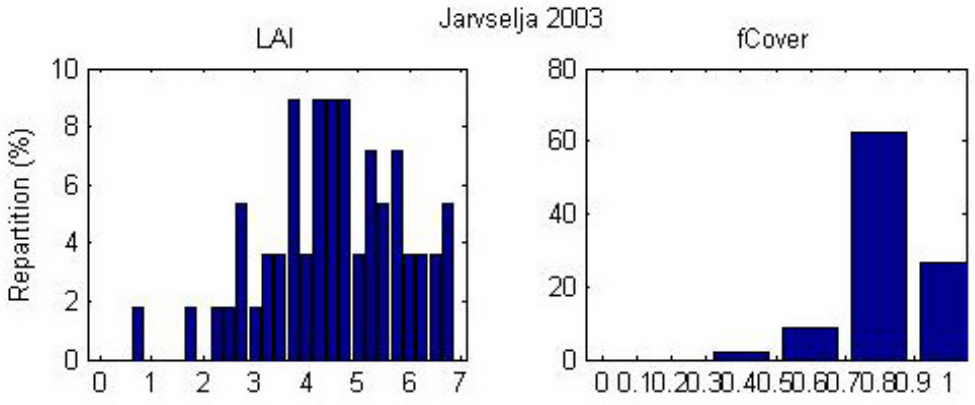

Figure 5. Distribution of the measured biophysical variables over the ESUs.

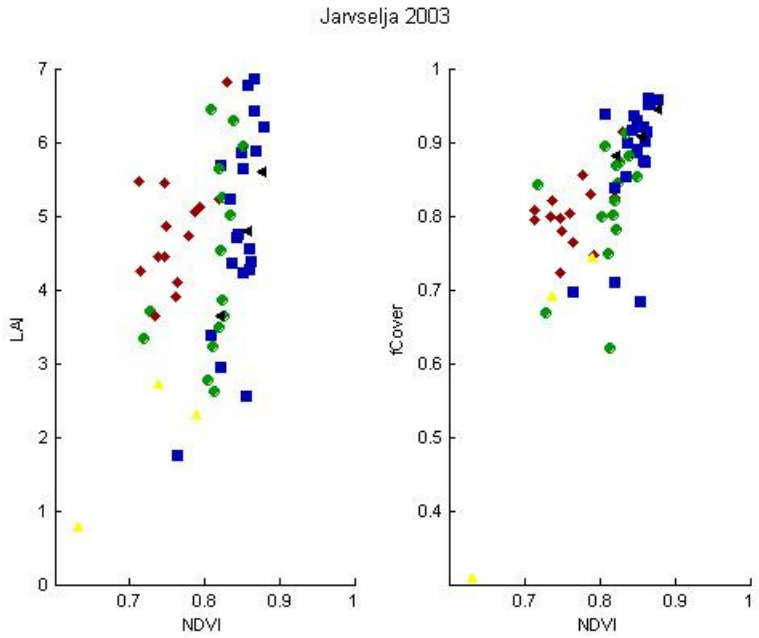

Figure 6. NDVI-Biophysical Variable relationships as a function of SPOT classes

Figure 6 shows the different relationships observed between the biophysical variables and corresponding NDVI or the ESUs, as a function of the SPOT classes determined in §2.2. No different behaviour between the classes can be observed; therefore a single transfer function per variable will be generated.

3 Determination of the transfer function for the 2 biophysical variables : LAI, fCover

3.1 The Transfer functions considered

For each class determined in §2.2, two types of transfer functions were tested:

- REG: If the number of ESUs is sufficient, multiple robust regression between ESUs reflectance (or Simple Ratio) and the considered biophysical variable can be considered: 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).

- LUT: If the number of ESUs is sufficient, Look-Up-Tables are also enviewed : a look-up table is build using ESUs reflectances and corresponding measured biophysical variable. For a given pixel, a cost function is computed as the sum square difference between the pixel reflectances and the ESU reflectances over the 4 bands, divided by the standard deviation computed on ESU reflectances. The result of the cost function is sorted in ascending order, and the biophysical variable estimated for the given pixel is computed as the mean value of the first *n* ESUs providing the lowest value of the cost function. Different values of *n* are considered to get the lowest cost function. This method is reliable only if the ESU NDVI distribution is quite comparable with the whole site NDVI distribution, which is the case for Järvselja.

Both regression and Look-Up-Tables are tested using either the reflectance or the logarithm of the reflectance for any band combination as well as the simple ratio. As both methods have poor extrapolation capacities, a flag image, based on the computation of convex hull over reflectances, is computed showing:

- Pixels inside the 'strict convex-hull': for each class, a convex-hull is computed using all the reflectance combination used for the transfer function, and corresponding to the ESUs belonging to the class. For those pixels, the transfer function is used as an interpolator, and the degree of confidence in the results obtained is quite high.
- Pixels inside the 'large convex-hull': for each class, a convex-hull is computed using all the reflectance combination (±5% in relative value) used for the transfer function, and corresponding to the ESUs belonging to the class. For those pixels, the transfer function is used as an extrapolator (but not far from interpolator), and the degree of confidence in the results obtained is quite good.
- Pixels outside the two convex-hulls: this means that for these pixels, the transfer function acted like 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.

3.2 Results on the Järvselva site

3.2.1 Choice of the method

For the 5 classes, a unique transfer function was computed. Figure 7 and Figure 8 show the results obtained for all the possible band combinations using either the reflectance or the logarithm of the reflectance:

- The REG method provides better results in terms of cross-validation RMSE for all the variables and is therefore selected as the transfer function instead of the LUT.
- For leaf area index, the results using the logarithm of the reflectance are slightly better (cross-validation RMSE) whereas the use of the reflectance itself provides better RMSE for fCover.

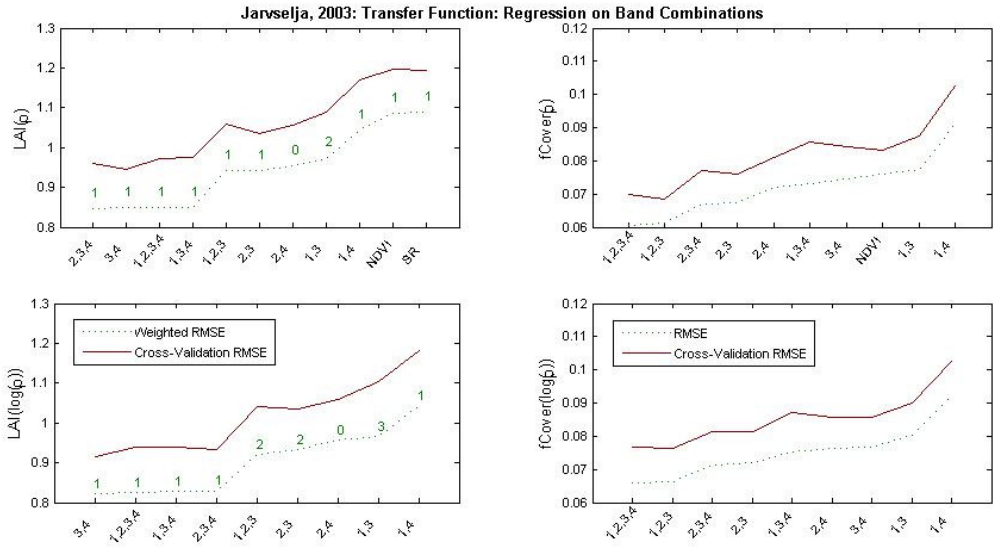

Figure 7. Transfer function: test of multiple regression applied on different band combinations. Band combinations are given in abscissa. The estimated biophysical variable is given in ordinate. Top graphs correspond to regression made on reflectance (ρ): the weighted root mean square error (RMSE) is 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. Bottom graphs correspond to regression made on the logarithm of the reflectance.

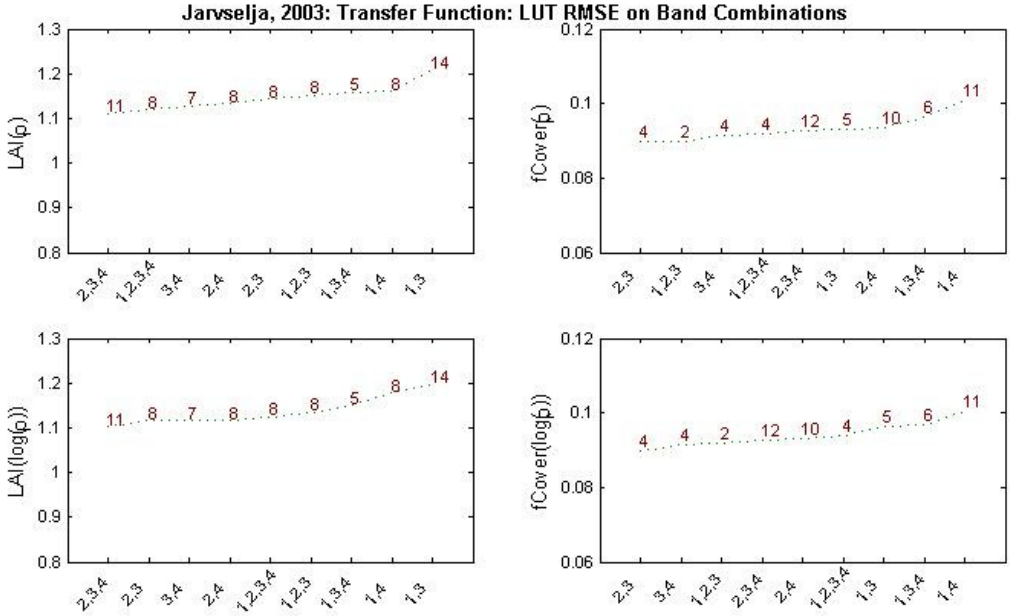

Figure 8. Transfer function: test of LUT applied on different band combinations. Band combinations are given in abscissa. The estimated biophysical variable is given in ordinate. Top graphs correspond to regression made on reflectance (ρ): the root mean square error is presented in green. The numbers indicate the number of elements selected in the LUT to compute the resulting biophysical variables. Bottom graphs correspond to LUT using the logarithm of the reflectance.

3.2.2 Choice of the band combination

For the effective LAI, the XS3, XS4 combination on log(ρ) was selected since it provides the best results in terms of weighted root mean square error and it provides only one weight lower than 0.7 and the lowest

cross-validation RMSE value. Moreover, Figure 9 shows that there is no bias (a null LAI is estimated quasi 0).The weights associated to each ESU are all higher to 0.7 (Figure 10).

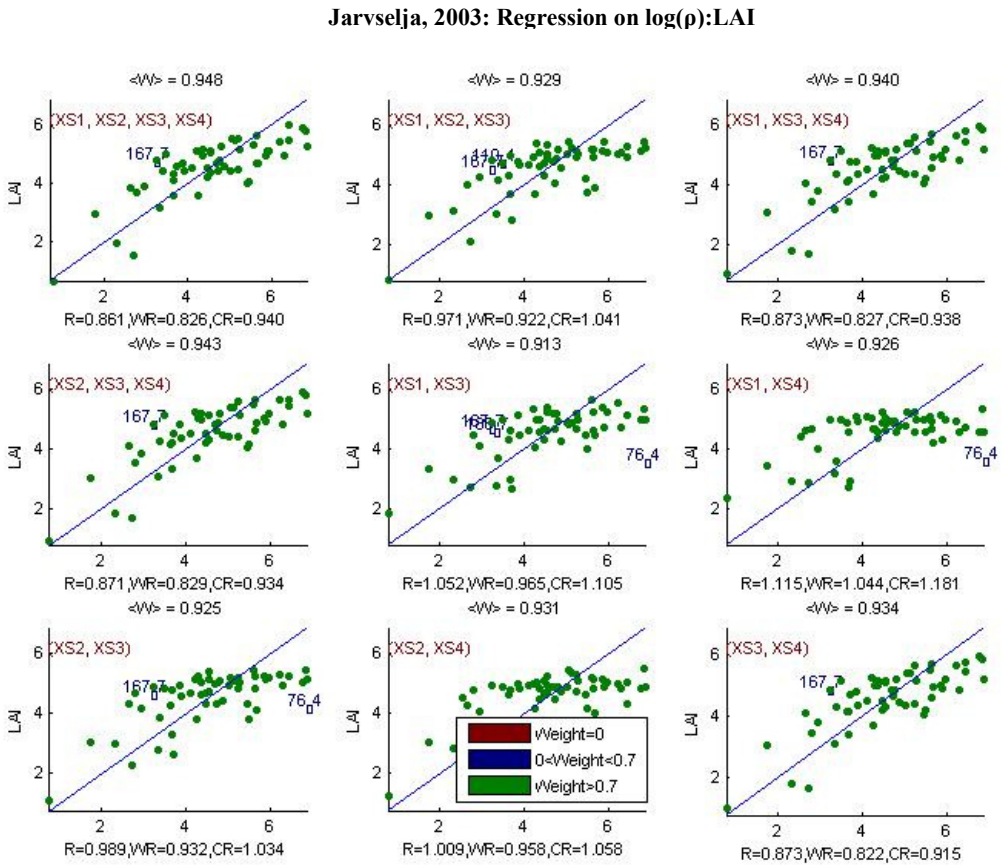

Figure 9. Effective Leaf Area Index: results for regression using different band combinations. R is the root mean square error computed between LAI and estimated LAI. WR is the weighted root mean square error and CR is the cross validation root mean square error.

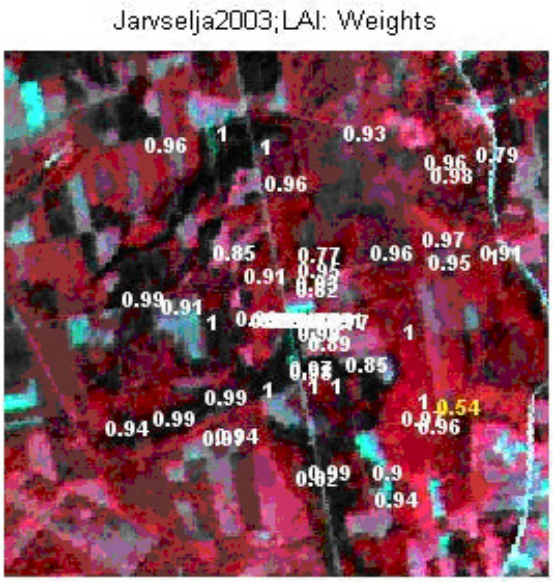

Figure 10. Weights associated to each ESU for the determination of LAI transfer function.

For the fCover, the robust regression algorithm was not able to converge since there was no outlier in the data. Therefore a simple multiple regression was applied. The XS1, XS2, XS3 combination on reflectance was selected since it provides the lowest values in terms of RMSE and cross-validation RMSE (Figure 11).

Jarvselja, 2003: Regression on reflectance: fCover

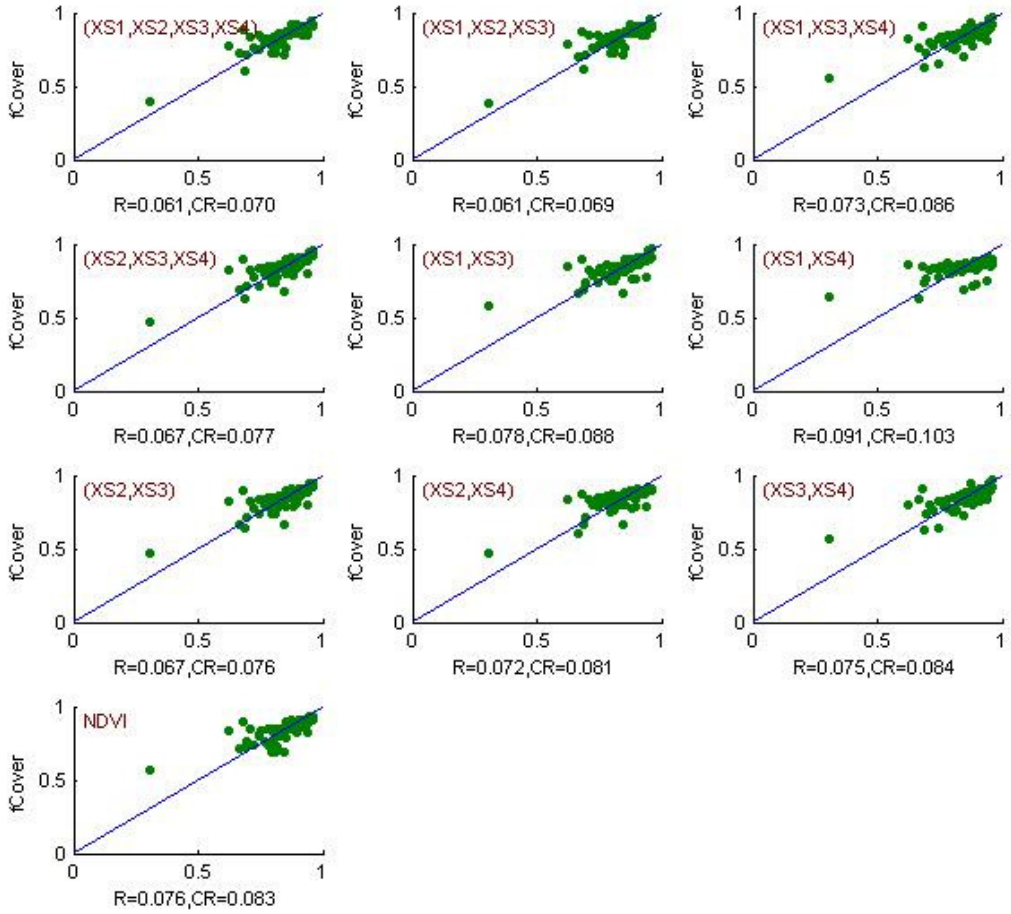

Figure 11. fCover: results for regression using different band combinations

Variable	Band Combination	RMSE	Weighted RMSE	Cross-valid RMSE
Effective LAI	-2.4142 + 6.0736 log(XS3) - 7.5157 log(XS4)	0.873	0.822	0.915
fCover	0.9044 + 16.3533 XS1 -26.8851 XS2 + 0.3247 XS3	0.061	-	0.069

Table 2. Transfer function applied to the whole site for the different biophysical variables, and corresponding errors

3.3 Applying the transfer function to the Järvselva SPOT image extraction

Figure 12 presents the biophysical variable maps obtained with the transfer function described in Table 2. The maps obtained for the two variables are consistent, showing similar patterns: low LAI values where low fCover are observed and conversely.

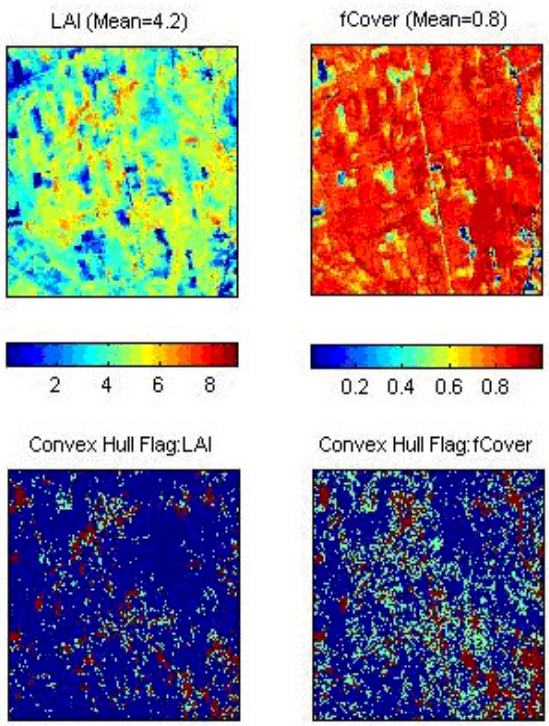

Figure 12. High resolution biophysical variable maps applied on the Järvselva site (top). Associated Flags are shown at the bottom: blue and light blue corresponds to the pixels belonging to the 'strict' and 'large' convex hulls and red to the pixels for which the transfer function is extrapolating.

The flag maps are different between LAI and fCover. This is explained by the fact that the summits of the convex-hulls are defined using two (LAI) or three (fCover) bands. Moreover, there is little extrapolation of the transfer function all over the site (especially for LAI).

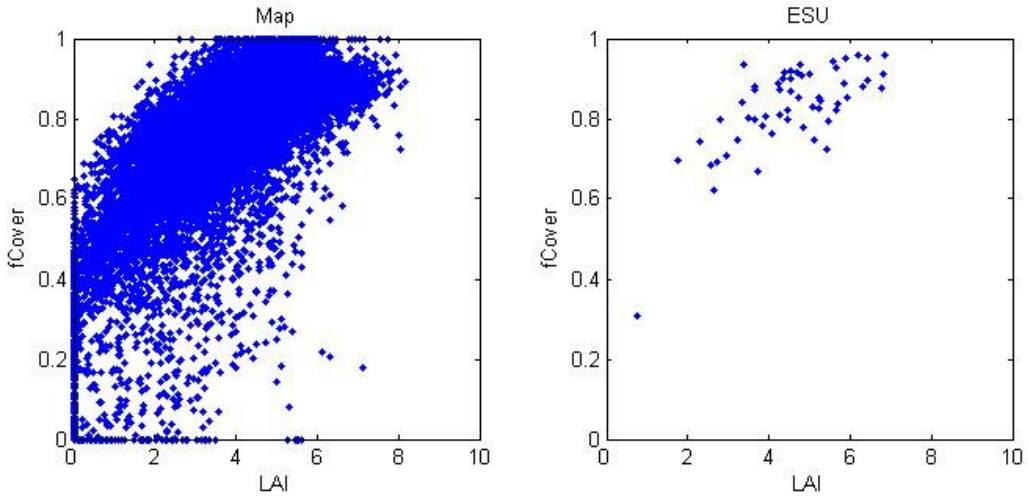

Figure 13. Relationships between LAI and fCover from SPOT image and measurements over the ESUs

Figure 13 shows the relationships observed between LAI and fCover estimated with the transfer function over the whole SPOT image (left) and measurements over the ESUs (right). Note that the scatter plots are consistent.

4 Conclusion

The transfer functions are finally obtained by using 56 ESUs. The relationships observed between LAI and fCover estimated with the transfer function over the whole SPOT image and measurements over the ESUs is similar. For the LAI variable, the regression coefficients are computed by relating the variable itself to logarithm of the reflectance while the fCover variable is related to reflectance. Only the XS3 band is common to both regressions. The maps obtained for the biophysical variables are consistent and the flag associated to each map show that the transfer function is used as an extrapolator in little areas.

The biophysical variable maps are available in Lambert-Est-92 projection coordinates (Datum: WGS-84) at 20m resolution.

5 Acknowledgements

We want to thank :

- people of the Tartu Observatory who participated to the campaign: **Mait Lang**, **Tõnu Lükk**, **Hannes Böttcher** and **Tiit Nilson**.
- **Marie Weiss** for her advices and her precious help for the processing of these data.