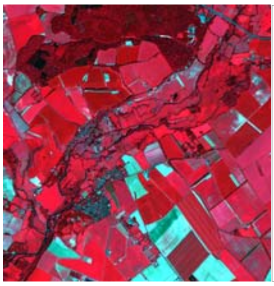
VALERI PROJECT

Website: http://www.avignon.inra.fr/valeri

CAMPAIGN 2006

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



VALERI 2006

Chilbolton site
(crops and forest)

Philippe Rossello

June 2009

CONTENTS

1. Introduction

This report describes the production of the high resolution, level 1, biophysical variable maps for the Chilbolton site in June 2006 (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 determination of a transfer function between reflectance values of the SPOT image acquired during or around the ground campaign and biophysical variable measurements (LAI2000 data and hemispherical images).

The derived biophysical variable maps are:

- Leaf Area Index (LAI): 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.

The area is mainly "agricultural fields which were growing barley, wheat, oats and oilseed rape or left fallow. The forest comprises conifer plantations and areas of broadleaf woodland". The site is flat (for more information, see campaign report: http://www.avignon.inra.fr/valeri).

The site coordinates are described in Table 1:

	United Kingdom, Ordnance Survey of Great Britain '36 (units = meters)		Geographic Lat/Lon WGS-84 (units = degrees)	
	Easting	Northing	Lat	Lon
Upper left corner	437298.9463	143403.6475	51.18859062	-1.46768170
Lower right corner	442518.9463	137983.6475	51.13949214	-1.39363124
Center	439908.9463	140693.6475	51.16404723	-1.43063682

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

2. Available data

2.1. SPOT image

The SPOT image was acquired the 10th June 2006 by HRG 1 on SPOT5. It was radiometrically and geometrically corrected by SPOT image (product: SPOTView Precision, level 3). The projection is United Kingdom, Ordnance Survey of Great Britain '36. The atmospheric correction[1] was performed by UMR 1114 INRA-UAPV EMMAH.

Figure 1 shows the relationship between Red and near infrared (NIR) SPOT channels: the soil line is marked and no saturated point is observed.

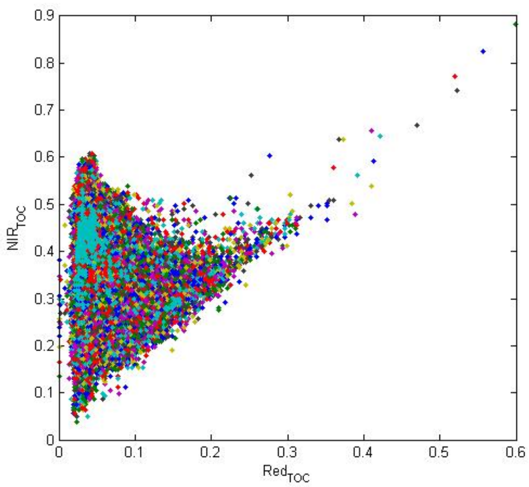


Figure 1. Red/NIR relationship on the SPOT image for Chilbolton, 2006

[1]Aerosol optical thickness: AOT_550 (nm) = 0.114 (Chilbolton AERONET site); water vapor content (gcm-2): 1.643 (Chilbolton AERONET site); ozone content: 0.323 atm.cm (TOMS observations); air Pressure: 1007 hPa.

2.2. Ground measurements

The ground measurements were performed from 14th to 17th June 2006. For each Elementary Sampling unit (ESU), LAI2000 data or hemispherical images were used to estimate biophysical variables. The type of measurement is specified in the GPS file (http://www.avignon.inra.fr/valeri).

2.2.1. LAI2000 data

The biophysical variables (LAI, fCover) were derived from LAI2000 instrument. According to the sampling protocol, 48 measurements were taken for each ESU. 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).

2.2.2. Hemispherical images

The hemispherical images were acquired from above the understorey and from below the canopy. The two sets of acquisition were processed separately to derive LAI (effective and true), LAI57 (effective and true), fCover, and fAPAR. The ESU biophysical variable was then computed as:

• LAIeff, LAI57eff, LAItrue, LAI57true: LAI(above) + LAI(below).

• fCover: 1-(1-fCover(above))*(1-fCover(below)). This assumes independency between the gaps inside the understorey and those inside the trees which is not true at all the scales but it is the only way to get the total fCover. However, for the local scales considered, this might be true as a first order approximation.

• fAPAR: 1-(1-fAPAR(below))*(1-fAPAR(above)), since 1-fAPAR can be considered equivalent to a gap fraction. Here again, the same independency between the two layers has to be assumed.

As LAI2000 instrument only estimates LAIeff and fCover, it is not possible to produce true LAI and fAPAR maps.

2.2.3. Distribution of the measured biophysical variables

Figure 2 shows the distribution of the several variables over the 28 sampled ESUs. Effective LAI varies from 1.93 to 5.64 and fCover from 0.35 to 0.99. Three ESUs located on bare soil (LAI = 0) were added to order to improve the representativeness of ESUs (§2.3.1). This range shows a heterogeneous site in terms of LAI. To build the relationships between biophysical variables and SPOT data, the reflectance of a given forest ESU (hemispherical images) 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 20 m and consequently the fish-eye observes an area of $\pi\ x\ [20\ x\ \tan(60°)]^2 = 3770\ m^2$, *i.e.* close to the area of 9 SPOT pixels (= 3600m²) when using a maximum view zenith angle of 60°.

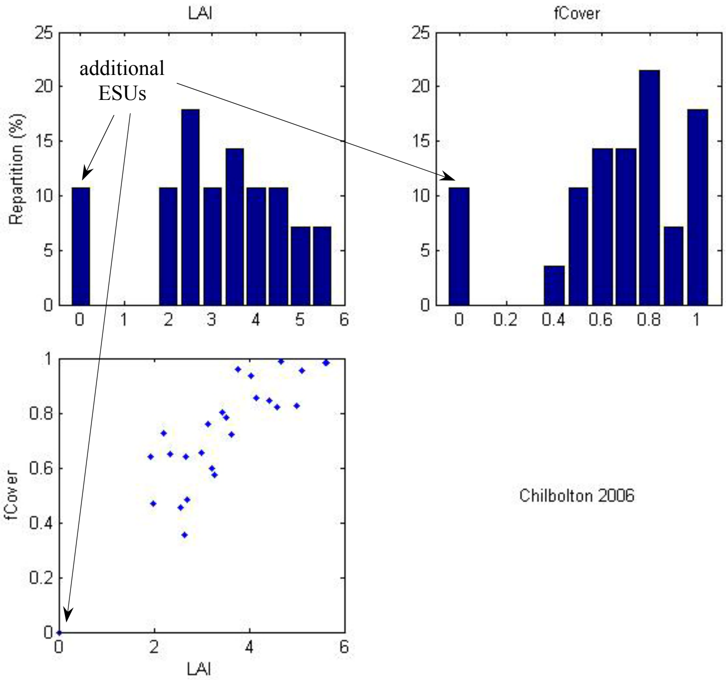


Figure 2. Distribution of the measured biophysical variables over the ESUs.

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.

Figure 3 shows that the 25 ESUs corresponding to ground measurements are evenly distributed over the site (5 x 5 km). Three ESUs (E18, E19 and E20) were added to improve the representativeness of the land cover. They are located on bare soil. For these ESUs, LAI value is equal to 0.

All ESUs have been kept for the computation of the transfer function:

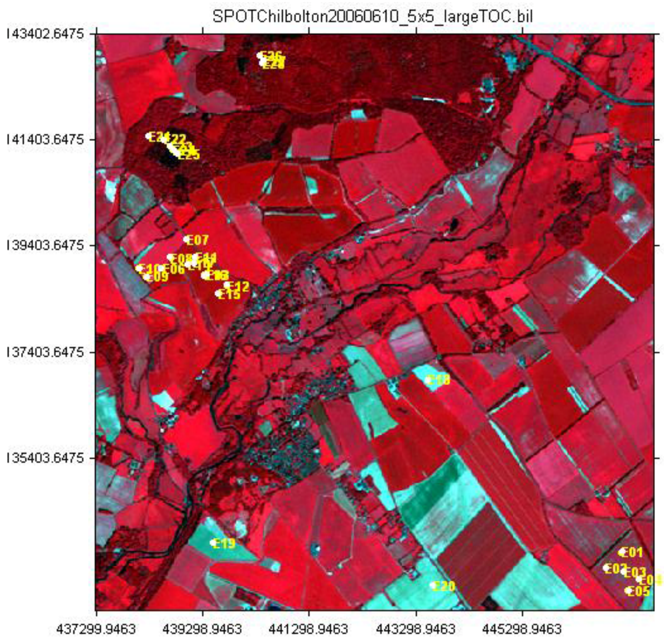


Figure 3. Distribution of the ESUs around the Chilbolton site.

2.3.2. Evaluation based on NDVI values

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 4). As the number of pixels is drastically different for the ESU and whole site (WS = 250000 in case of a 5 x 5 km image at 10 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.

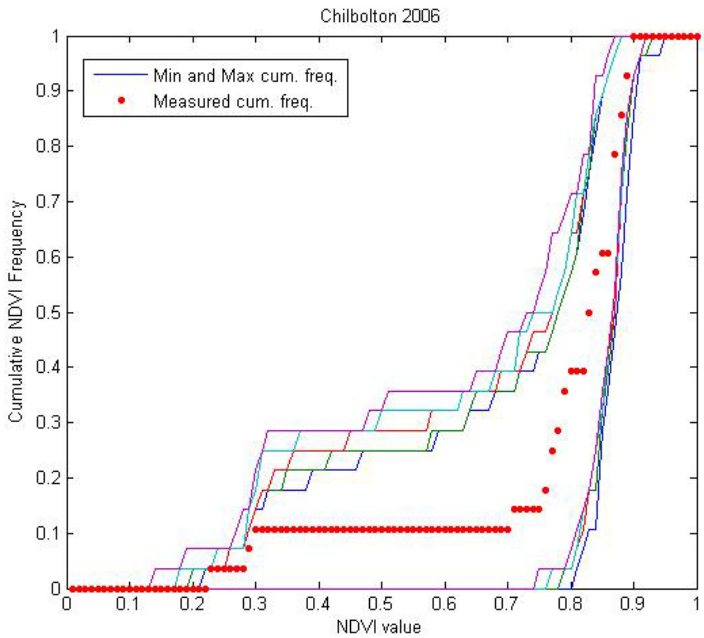


Figure 4. Comparison of the ESU NDVI distribution and the NDVI distribution over the whole image.

Figure 4 shows that NDVI distribution of the 28 ESUs is good over the whole site even if NDVIs between 0.31 and 0.70 have not been sampled although they are present in the image. They correspond to bare soil, fallows, grassland, crops… The sampling of different types of crops or different developmental stages of vegetation is probably in question. Note that Chilbolton site is very heterogeneous in terms of NDVI since the highest and lowest distributions are not close (except for NDVI values > 0.8).

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 SPOT image to distinguish if different behaviours on the image for the biophysical variable-reflectance relationship exist.

A number of 5 classes was chosen (Figure 5). The distribution of classes 1, 2 and 3 on the image and on the ESUs is similar. Class 3 (bare soil) corresponds to three additional ESUs (§2.3.1). The distribution of classes 4 and 5 is different. Class 4 (ESUs = winter wheat) is over-sampled and class 5 (ESUs = oilseed rape) is under-represented. However, the evaluation based on classification is satisfactory.

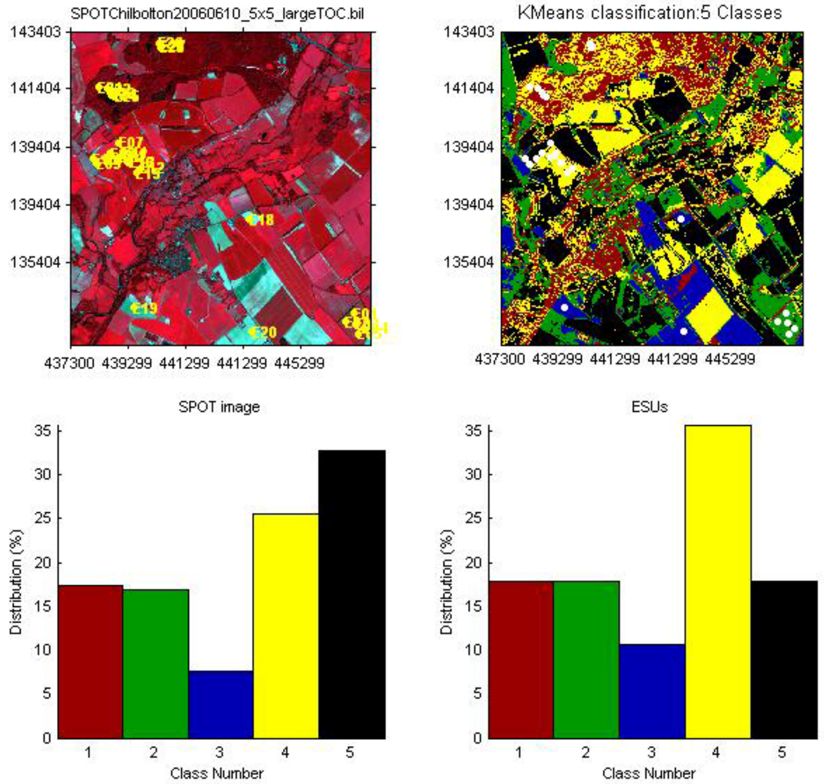


Figure 5. Classification of the SPOT image and comparison of the class distribution between the satellite image and sampled ESUs.

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

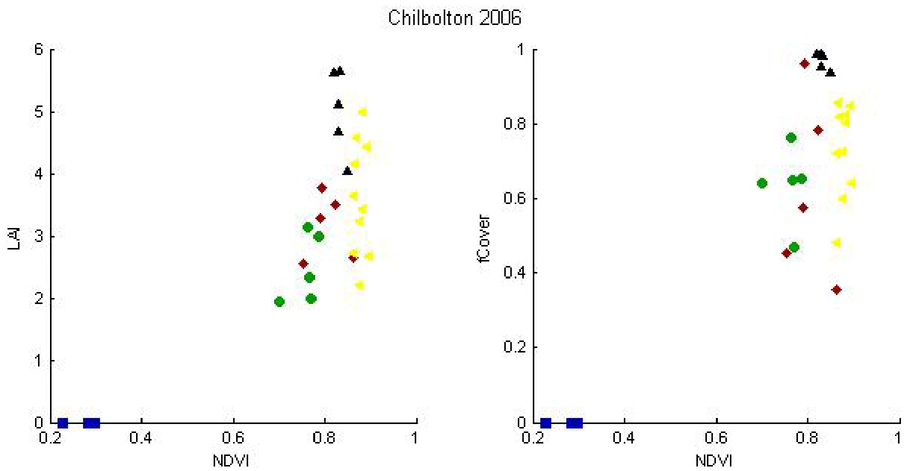


Figure 6. NDVI-biophysical variable relationships as a function of SPOT classes

The relationship between NDVI and biophysical variables on classes 4 and 5 (yellow = winter wheat; black = oilseed rape) are not very consistent. The additional ESUs (E18, E19 and E20) improve the relationships between the biophysical variables and corresponding NDVI, even if the results are not optimal. Note that NDVI values between 0.31 and 0.70 is not sampled (§2.3.2). However, a single transfer function will be generated (§3.2).

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 4 bands (green, red and

NIR, SWIR in this case) of the SPOT 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 SPOT 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.

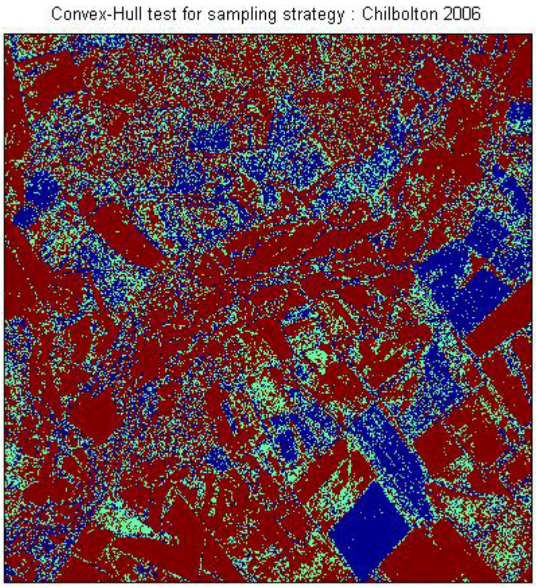


Figure 7. Evaluation of the sampling based on the convex hulls. The map is shown at the bottom: blue and light blue correspond to the pixels belonging to the 'strict' and 'large' convex hulls and red to the pixels for which the transfer function is extrapolating.

The flag map shows that the representativeness of the ESUs is insufficient. Pixels outside the two convex-hulls are numerous. They correspond to bare soil, fallows, crops, forest… The strict convex-hull mainly corresponds to winter wheat and oilseed rape.

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

3.1. The transfer function considered

Two types of transfer functions are usually tested in the frame of the VALERI project:

- 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 SPOT 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).

3.2. Results

3.2.1. Choice of the method

For all classes, '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.

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)): the regression made on the reflectance provides better results. The results using the reflectance are thus selected for LAI and fCover.

The Red*NIR ('+' or RN) combination is added to all the band combinations (except NDVI and SR). Please read the document (http://www.avignon.inra.fr/valeri/table_methods/new_linear.pdf): "A method to improve the relation between the biophysical variables".

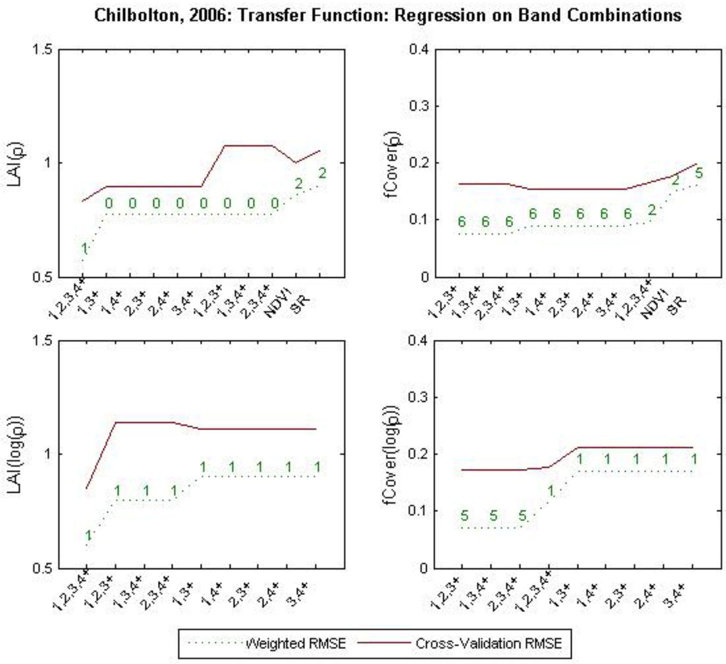


Figure 8. Transfer function: test of multiple regression applied on different band combinations. Band combinations are given in abscissa. The estimated biophysical variable is given in ordinate. Top graphs correspond to regression made on reflectance (ρ): the weighted root mean square error (RMSE) is 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 XS1, XS2, XS3, XS4, RN combination on reflectance (Figure 9 and Figure 10) was selected since it provides the best results. One weight is lower than 0.7.

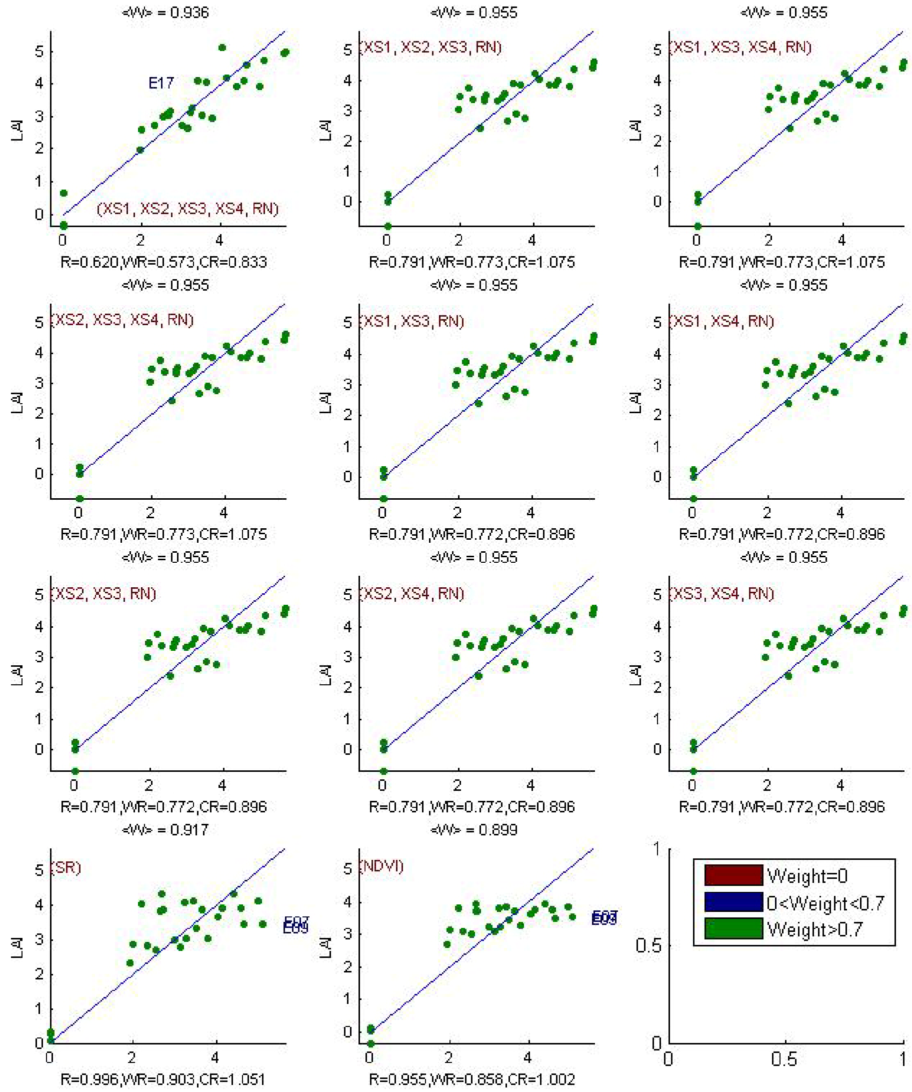


Figure 9. Leaf Area Index: results for regression on reflectance using different band combinations. R is the root mean square error computed between LAI and estimated LAI. WR is the weighted root mean square error and CR is the cross validation root mean square error.

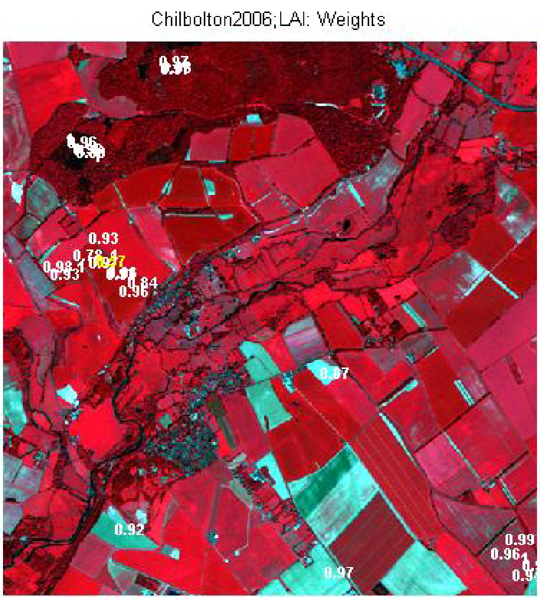


Figure 10. Weights associated to each ESU for the determination of LAI transfer function.

For the fCover, the XS1, XS2, XS3, XS4, RN combination on reflectance (Figure 11 and Figure 12) was selected since it provides a good compromise between cross-validation RMSE, weighted RMSE and RMSE. Two weights are lower than 0.7.

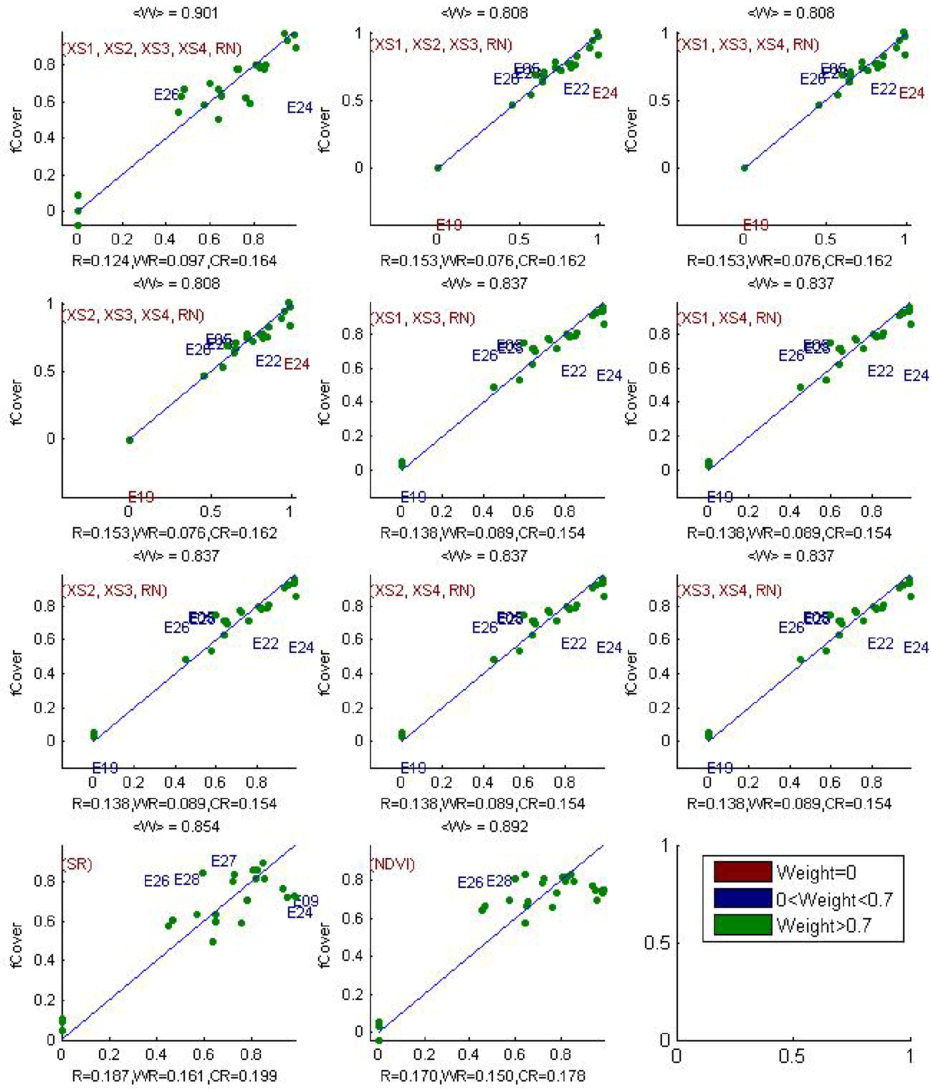


Figure 11. fCover: results for regression on reflectance using different band combinations. R is the root mean square error computed between fCover and estimated fCover. WR is the weighted root mean square error and CR is the cross validation root mean square error.

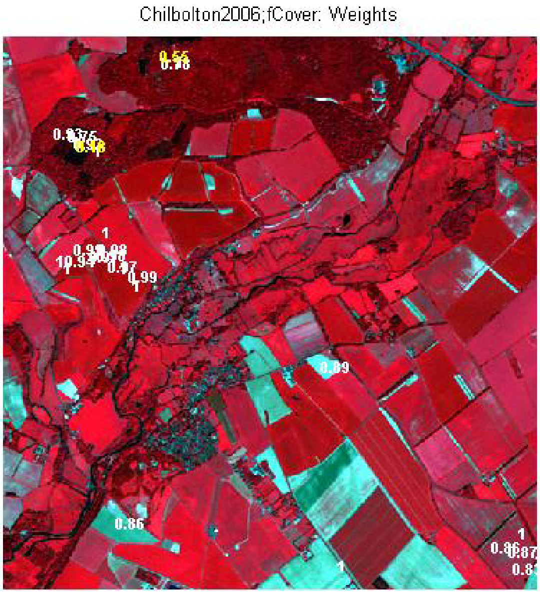


Figure 12. Weights associated to each ESU for the determination of fCover transfer function.

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

Variable	Band Combination	RMSE	Weighted RMSE	Cross-valid RMSE
LAI	5.3361 + 8.4408(XS1) - 47.4188(XS2) + 0.3204(XS3) - 21.4674 (XS4) + 163.4993(RN)	0.620	0.573	0.833
fCover	0.8322 + 1.5471(XS1) - 11.9504(XS2) + 0.2256(XS3) - 2.1851(XS4) + 32.242(RN)	0.124	0.097	0.164

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 Chilbolton SPOT image extraction

Figure 13 presents the biophysical variable maps obtained with the transfer function described in Table 2 for all classes. The maps obtained for the two variables are consistent, showing similar patterns: low LAI values where low fCover are observed and conversely…

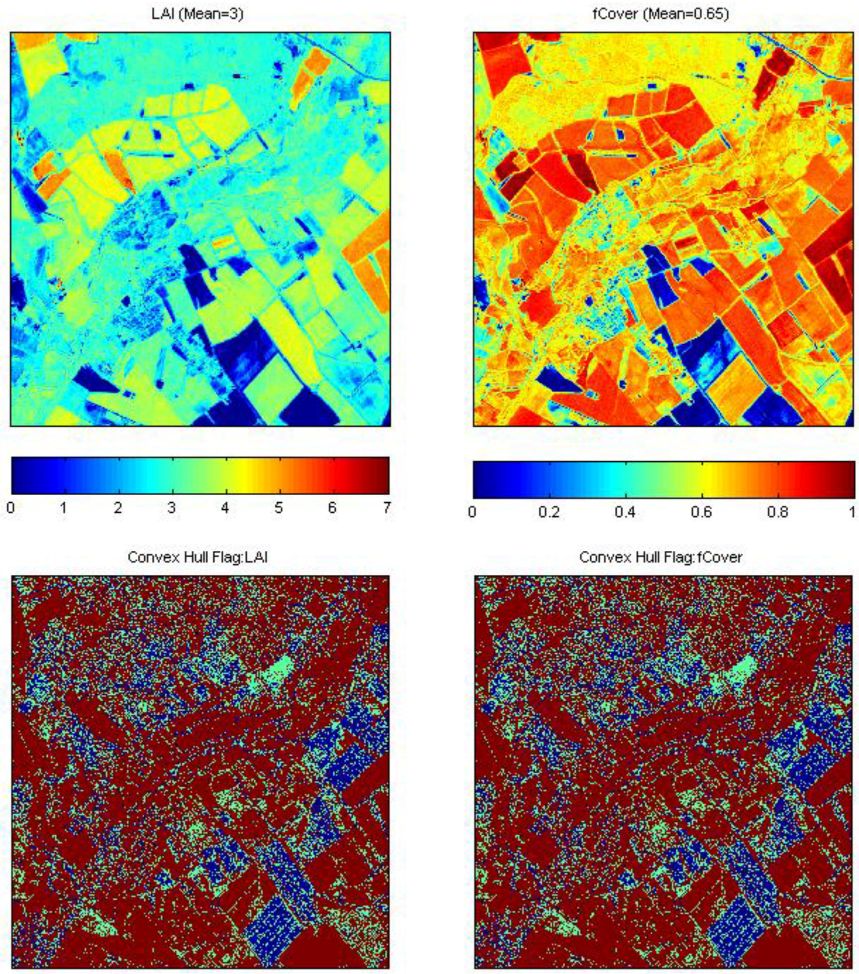


Figure 13. High resolution biophysical variable maps applied on the Chilbolton site (top). Associated Flags are shown at the bottom: blue and light blue correspond to the pixels belonging to the 'strict' and 'large' convex hulls and red to the pixels for which the transfer function is extrapolating.

The flag maps are comparable between the two biophysical variables. Note that the extrapolation is large. The pixels outside the two convex-hulls mainly correspond to bare soil, fallows, grassland, crops (§2.3.4)… In theory, the more the number of bands increases, the larger the extrapolation is.

4. Conclusion

The Chilbolton site is heterogeneous in terms of LAI and NDVI. The representativeness of the different ESUs is not optimal (§2.3.2). 'REG' method (§3.1) is applied to all classes. The results of the robust regression are good. The maps obtained for the biophysical variables are consistent. The flag associated to each map shows that the large extrapolation of the transfer function is mainly bounded to bare soil, fallows, crops… For LAI and fCover, the regression coefficients are computed by relating the variable itself to reflectance.

The biophysical variable maps are available in United Kingdom, Ordnance Survey of Great Britain '36 projection coordinates at 10m resolution.

5. Acknowledgements

We want to thank **E.J. Milton and participants of the experiment** (University of Southampton) for the organisation to the campaign.