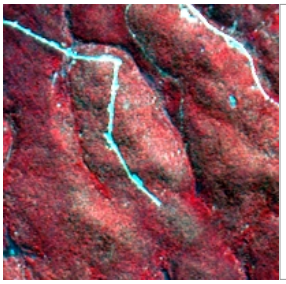
GROUND DATA PROCESSING & PRODUCTION OF THE LEVEL 1 HIGH RESOLUTION MAPS



VALERI 2004

Camerons site
(broadleaf forest)

Philippe Rossello, Frédéric Baret

June 2007

CONTENTS

1. Introduction **2**
2. Available data **2**
2.1. SPOT Image 2
2.2. Hemispherical images 3
2.3. Sampling strategy 6
2.3.1. Principles 6
2.3.2. Evaluation based on NDVI values 6
2.3.3. Evaluation based on classification 7
2.3.4. Using convex hulls 9
3. Determination of the transfer function for the 6 biophysical variables: LAIeff, LAItrue, LAI57eff, LAI57true, fCover, fAPAR **10**
3.1. The transfer functions considered 10
3.2. Results 11
3.2.1. Choice of the method 11
3.2.2. Choice of the band combination 12
3.3. Applying the transfer function to the Camerons SPOT image extraction 22
4. Conclusion **24**
5. Acknowledgements **24**

1. Introduction

This report describes the production of high resolution, level 1, biophysical variable maps for the Camerons site in 2004. 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 (hemispherical images). For each Elementary Sampling Unit (ESU), the hemispherical images were processed using the CAN-EYE software (Version 5) developed at INRA-CSE. The derived biophysical variable maps are:

• four Leaf Area Index (LAI) are considered: effective LAI (LAIeff) and true LAI (LAItrue) derived from the measurement of the gap fraction as a function of the view zenith angle; effective LAI57 (LAI57eff) and true LAI57 (LAI57true) derived from the gap fraction at 57.5°, which is independent on leaf inclination. Effective LAI and effective LAI57 do not take into account clumping effect. LAItrue and LAI57true are derived using the method proposed by Lang and Xiang[1] (1986);

• cover fraction (fCover): it is the percentage of soil covered by vegetation. To improve the spatial sampling, fCover was computed over 0 to 10° zenith angle;

• fAPAR: it is the fraction of Absorbed Photosynthetically Active Radiation (PAR = 400-700nm). fAPAR is defined either instantaneously (for a given solar position) or integrated all over the day. Following a study based on radiative transfer model simulations, it has been shown that the root mean square error between instantaneous fAPAR computed every 30 minutes and the daily fAPAR is the lowest for instantaneous fAPAR at 10h00 AM (solar time, RMSE = 0.021). Therefore, the derivation of fAPAR from CAN-EYE corresponds to the instantaneous black sky fAPAR at 10h00 AM.

The Camerons site corresponds to a natural autralian forest over gently undulated topography. Vegetation is made of several local species for which the distribution depends mainly on past fire events and topographic situations. Larger trees are 30 m maximum height, dominating younger individuals and shorter species. The understorey is relatively dry with some grasses and bushes sparsely distributed. The whole landscape appears mainly clumped at the shoot and tree levels. The site is approximately 4 x 4 km with coordinates described in Table 1:

	UTM 50, South WGS-84 (units = meters)		Geographic Lat/Lon WGS-84 (units = degrees)	
	Easting	Northing	Lat.	Lon.
Upper left corner	427998.1240	6395013.7590	-32.57999475	116.23285645
Lower right corner	432038.1240	6390973.7590	-32.61669188	116.27560436
Center	430018.1240	6392993.7590	-32.59834510	116.25422605

Table 1. Description of the site coordinates.

The ground measurements were carried out from 2nd March to 4th March 2004, while the high spatial resolution image (SPOT4, HRVIR1, resolution: 20 m) was acquired on 6th April 2004.

2. Available data

2.1. SPOT Image

The SPOT image was acquired on 6th April 2004 by HRVIR1 on SPOT4. It was geo-located by SPOT image (SPOTView Basic product). The projection is UTM 50, South, WGS-84. No atmospheric correction was applied to the image since no atmospheric data were available. However, as the SPOT 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 4 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) SPOT channels: the soil line is well marked and no saturated points are observed.

[1] Lang, A.R.G. and Xiang, Y., 1986. Estimation of leaf area index from transmission of direct sunlight in discontinuous canopies. Agric. For. Meteorol., 37: 229-243.

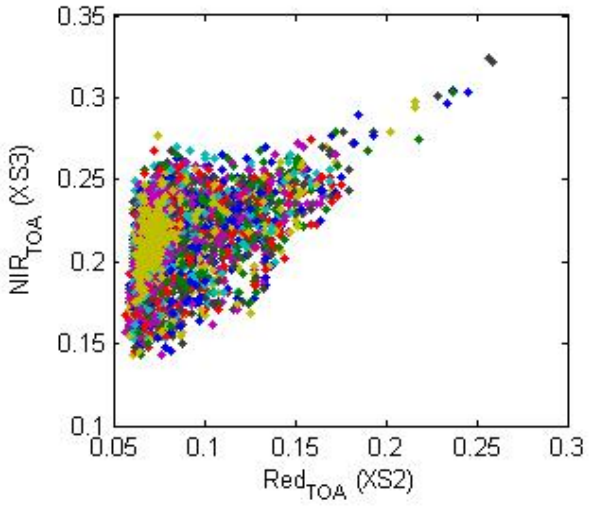


Figure 1. Red/NIR relationship on the SPOT image for Camerons, 2004.

2.2. Hemispherical images

The hemispherical images were processed using the CAN-EYE software (Version 5) to derive the biophysical variables. Figure 2 and Figure 3 show the distribution of the several variables over the 29 sampled ESUs. As Camerons site is mainly covered of forest (except X7), the hemispherical images were acquired from above the understorey and from below the canopy (trees). 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.

Note that LAI (effective and true) derived from directional gap fraction and LAI derived from gap fraction at 57.5° (effective and true) are consistent (Figure 2 and Figure 3). Effective LAI (LAIeff, LAI57eff) varies from 0.19 to 1.96, while true LAI (LAItrue, LAI57true) varies from 0.21 to 4.17. However, note that the site is relatively homogeneous in terms of LAI (Figure 2). LAIeff and LAI57eff are lower than LAItrue and LAI57true, due to the clumping observed for several ESUs. The relationship between fAPAR and LAI is in agreement with what is expected (Beer-Lambert law) while the fCover-LAI relationship is more noisy (Figure 3).

To build the relationships between biophysical variables and SPOT data, the reflectance of a given forest 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 20 m (maximum height: 30 m) and consequently the fish-eye observes an area of $\pi \times [20 \times \tan(60°)]^2 = 3770$ m², *i.e.* close to the area of 9 SPOT pixels (= 3600m²) when using a maximum view zenith angle of 60°.

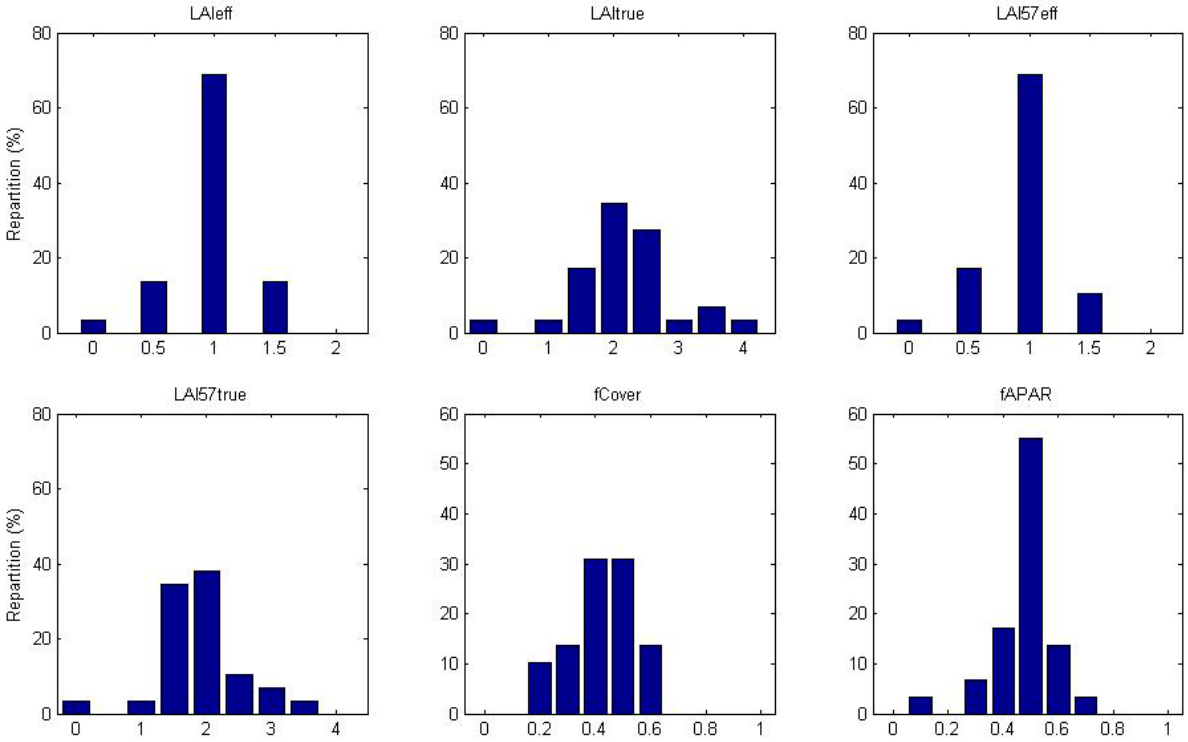


Figure 2. Distribution of the measured biophysical variables over the ESUs.

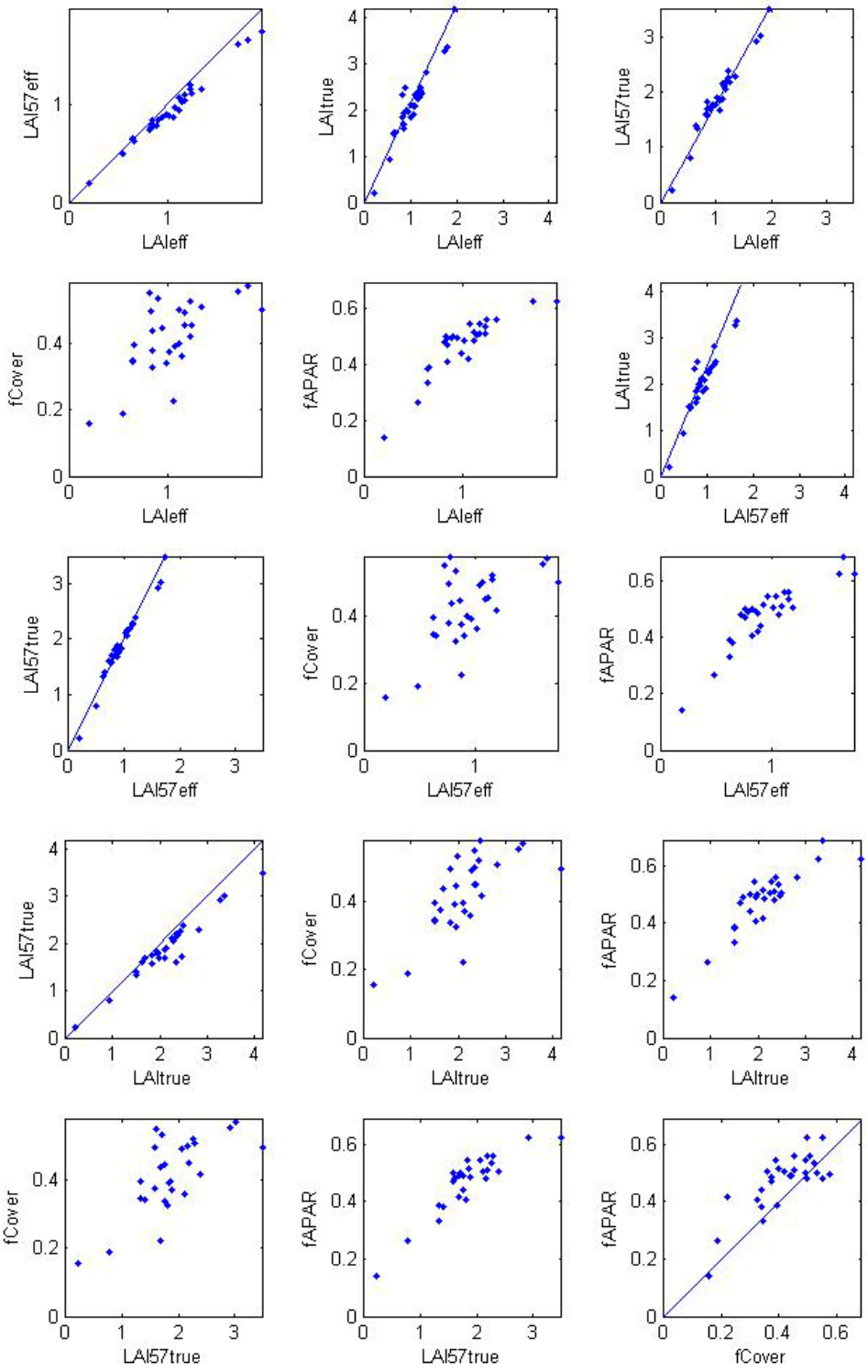


Figure 3. Relationships between the different biophysical variables.

2.3. Sampling strategy

2.3.1. Principles

The sampling of each ESU (Figure 4) is based at least on twelve elementary photographs from above the understorey and from below the canopy (except X7: only above the understorey).

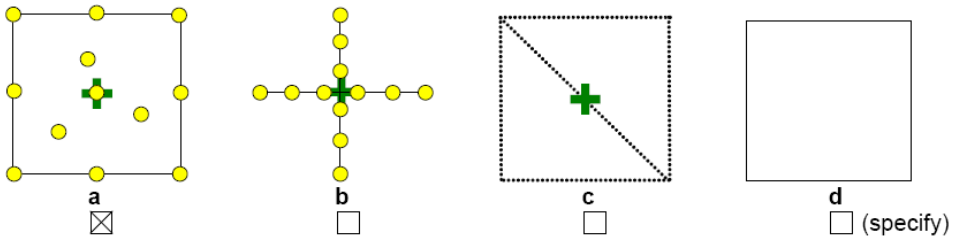


Figure 4. The sampling strategy for each ESU of the Camerons site

Figure 5 shows that the 29 ESUs are evenly distributed over the site (4 x 4 km). As the SPOT geo-location and the GPS measurements are not associated to errors, all the ESUs have been kept for the computation of the transfer function.

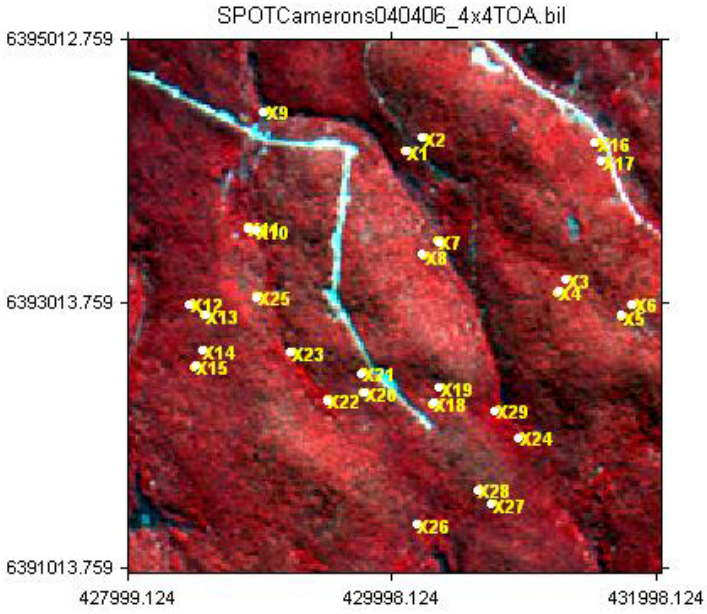


Figure 5. Distribution of the ESUs around the Camerons site.

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 6). As the number of pixels is drastically different for the ESUs and whole site (WS = 40000 in case of a 4 x 4 km SPOT image, resolution 20m), 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.

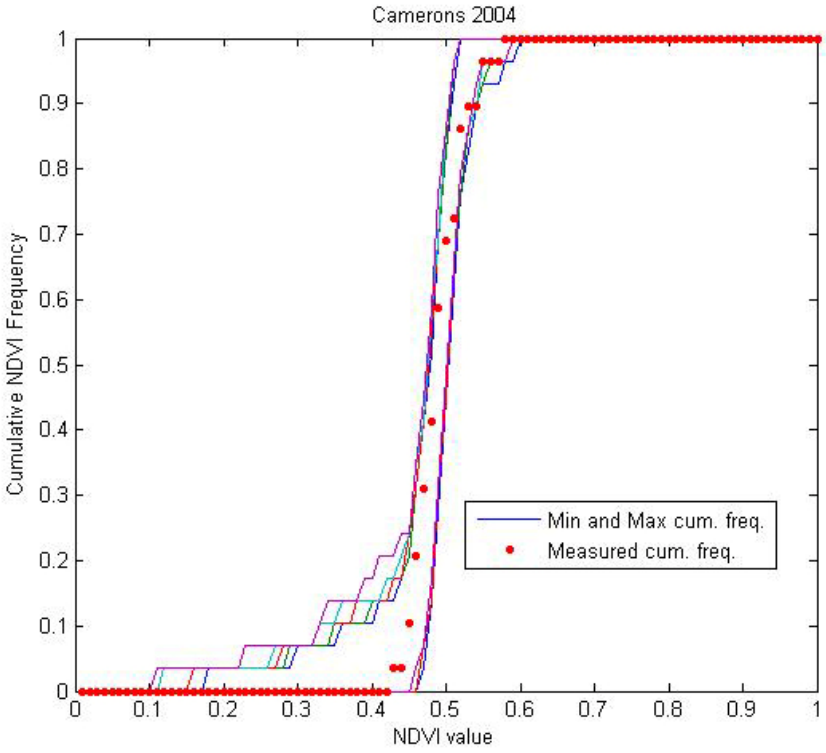


Figure 6. Comparison of the ESU NDVI distribution and the NDVI distribution over the whole image.

Figure 6 shows that the NDVI distribution of the 29 ESUs is very good over the whole site (comprised between the 5 highest and lowest cumulative frequencies), even if the cumulative frequency curve is sometimes very close to the boundaries (mainly for high NDVI values comprised between 0.53 and 0.57). Note that NDVIs lower than 0.43 have not been sampled although they are present in the image. The site is homogeneous in terms of NDVI since the highest and lowest distributions are close.

2.3.3. Evaluation based on classification

A non supervised classification based on the k_means method (Matlab statistics toolbox) was applied to the 4 reflectances of the SPOT image to distinguish if different behaviours on the image for the biophysical variable-reflectance relationship exist.

A number of 4 classes was chosen (Figure 7). Except class 3, the distribution of the classes on the image and on the ESUs is comparable. Class 3 corresponds to wide forest roads. As no ESU belongs to this class, a value equal to zero (LAI = 0) is attributed.

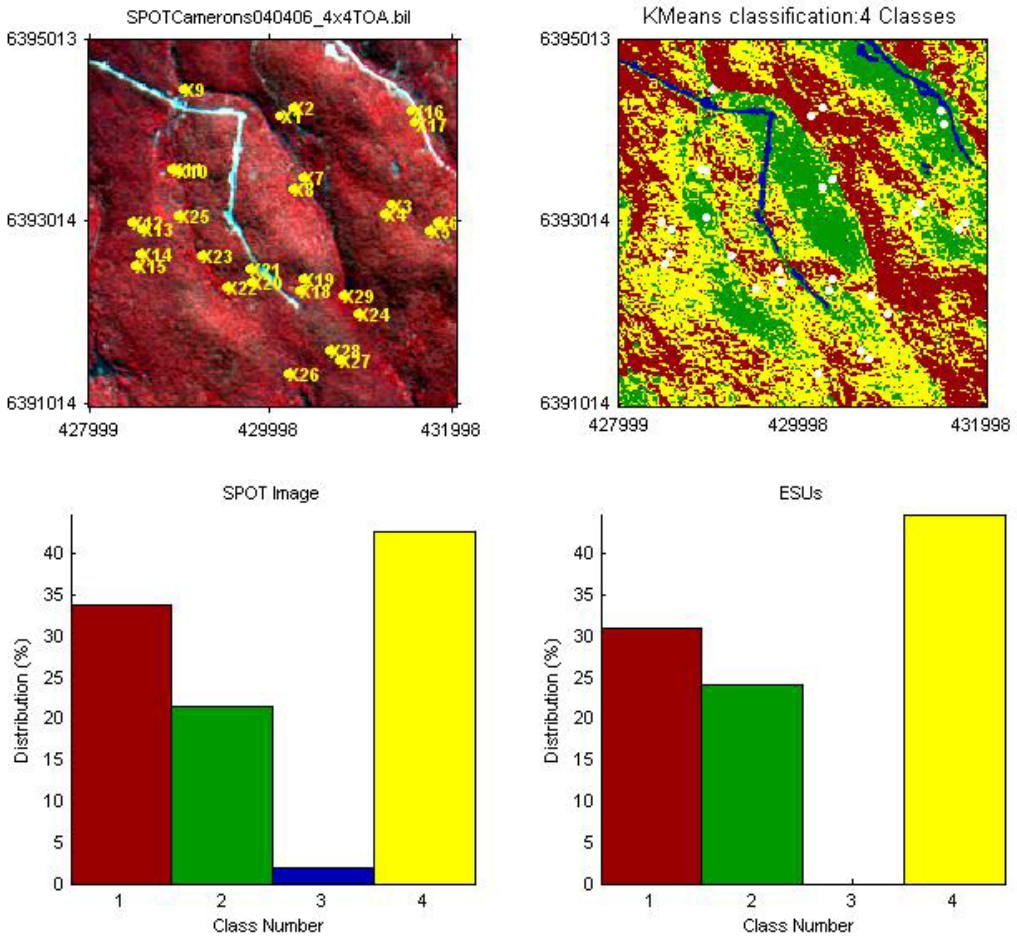


Figure 7. Classification of the SPOT image. Comparison of the class distribution between the SPOT image and sampled ESUs.

Figure 8 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.

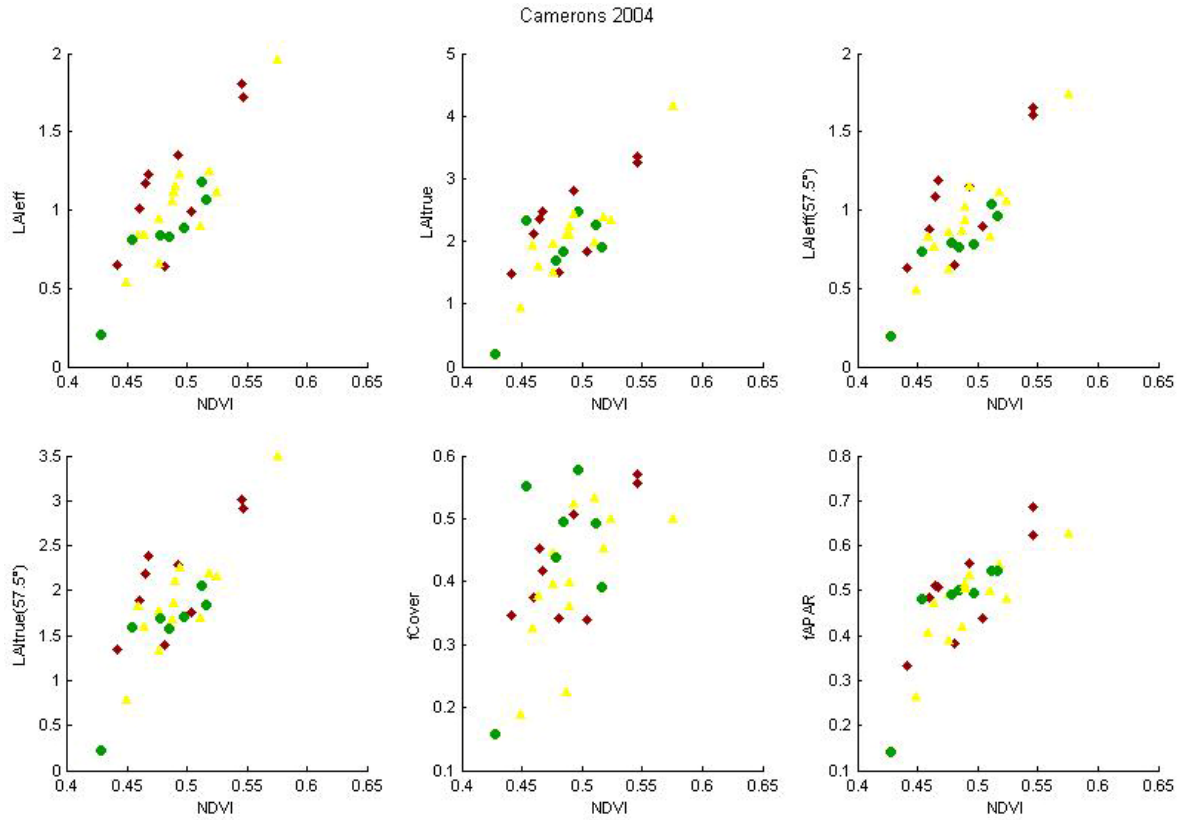


Figure 8. NDVI-Biophysical Variable relationships as a function of SPOT classes

The relation between NDVI and biophysical variables is consistent. No different behaviour between the classes can be observed.

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 of the SPOT image. A flag image, is computing over the reflectances (Figure 9). 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.

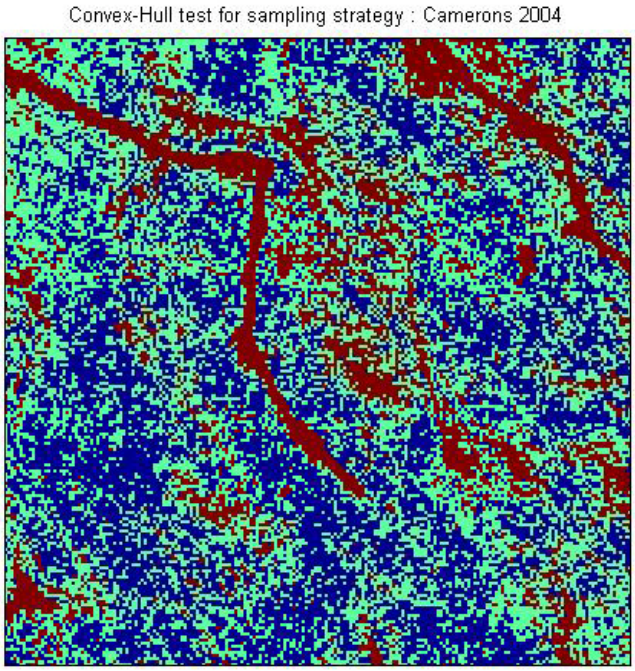


Figure 9. Evaluation of the sampling based on the convex hulls. The map is shown: blue and light blue correspond to the pixels belonging to the 'strict' and 'large' convex hulls and red to the pixels for which the transfer function is extrapolating.

This map shows that the representativeness of the ESUs is good, even if pixels are outside the two convex-hulls. They correspond to forest roads, bare soil, NDVI pixels lower than 0.43 which have not been sampled (§2.3.2).

3. Determination of the transfer function for the 6 biophysical variables: LAIeff, LAItrue, LAI57eff, LAI57true, fCover, fAPAR

3.1. The transfer functions 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).

The relationship between NDVI and LAI (§2.3.3) being consistent, the 'REG' method is applied to classes 1, 2 and 4. As class 3 corresponds to wide forest roads and no ESU belongs to this class, the value 0 is attributed to it.

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

A single transfer function was thus computed for classes 1, 2 and 4. Figure 10 shows the results obtained for all the possible band combinations using either the reflectance (ρ) or the logarithm of the reflectance (log(ρ)). Even if the regression made on the log(ρ) sometimes provides slightly better results, the results using the reflectance (ρ) were selected for all the variables. The transfer function using the log(ρ) indeed creates coplanar points which do not allow the determination of the 'strict' and 'large' convex hulls.

Note that 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).

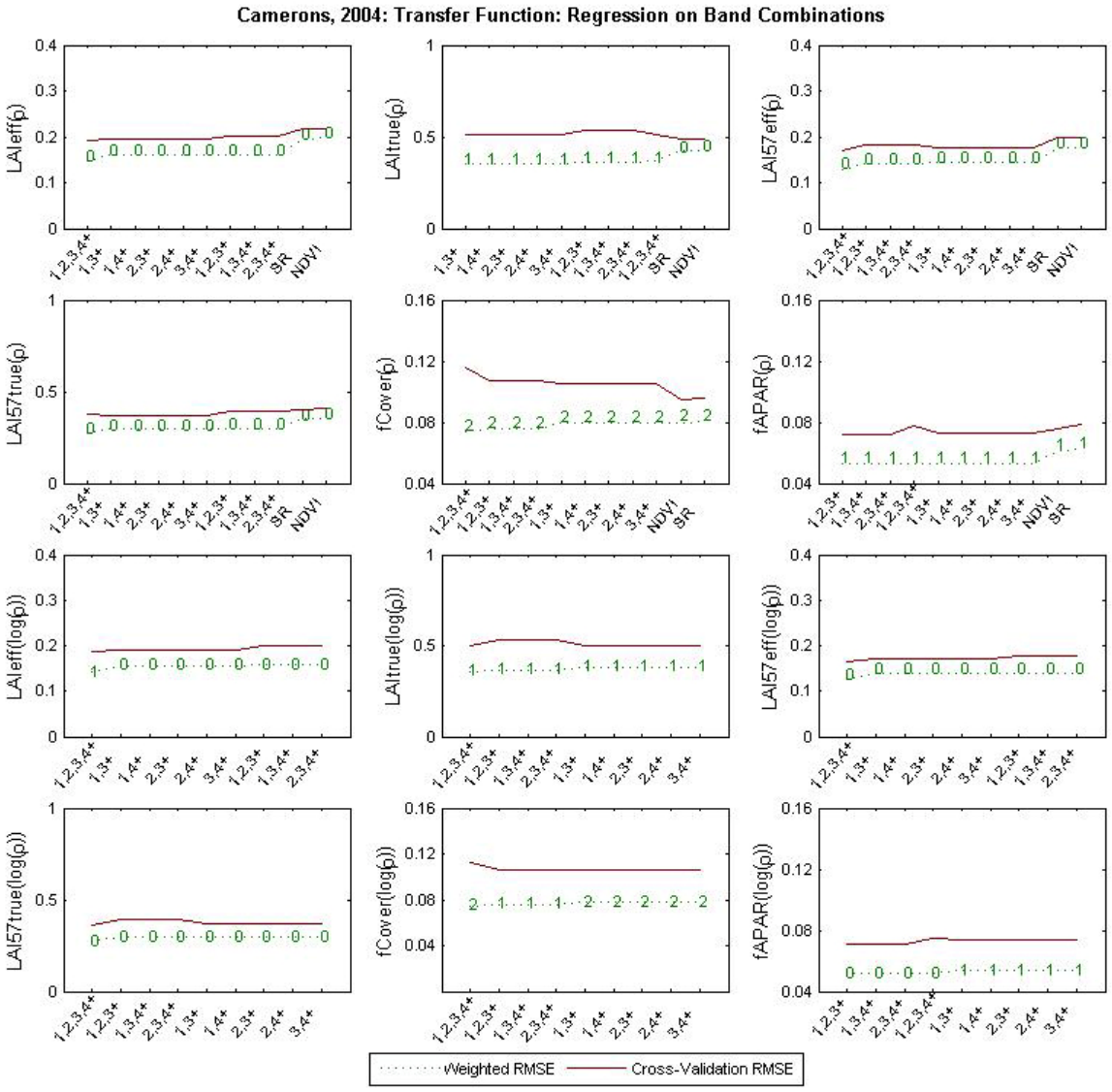


Figure 10. Transfer function: test of multiple regression applied on different band combinations. Band combinations are given in abscissa. The estimated biophysical variable is given in ordinate. Top graphs correspond to regression made on reflectance (ρ): the weighted root mean square error (RMSE) is 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 LAIeff, the XS1, XS2, XS3, XS4, RN combination on reflectance (Figure 11 and Figure 12) was selected since it provides the best results: cross-validation RMSE (lowest value), weighted RMSE (lowest value) and RMSE (lowest value). Note that zero weight is lower than 0.7.

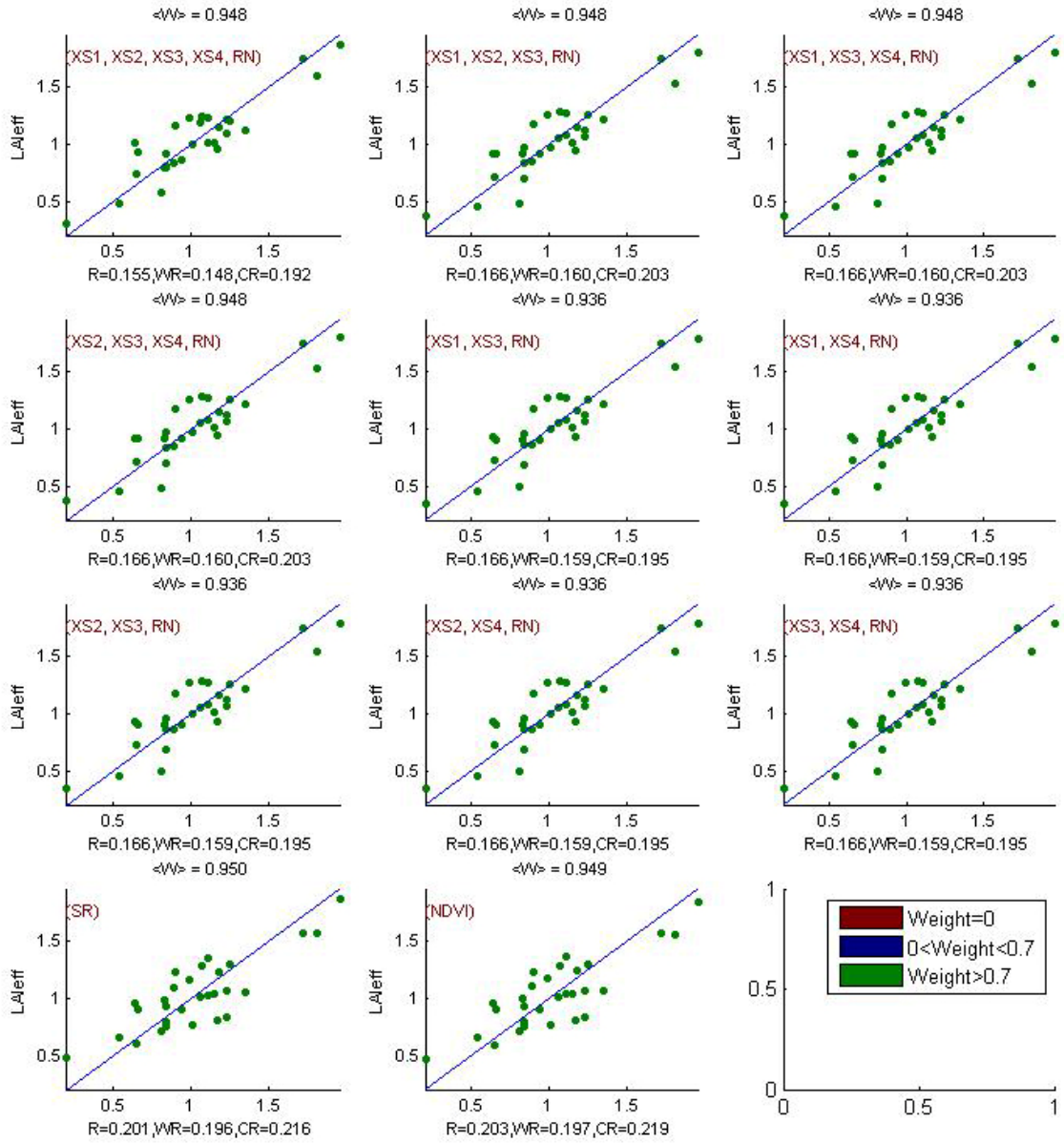


Figure 11. Effective Leaf Area Index: results for regression on reflectance using different band combinations. R is the root mean square error computed between LAIeff and estimated LAIeff. WR is the weighted root mean square error and CR is the cross validation root mean square error.

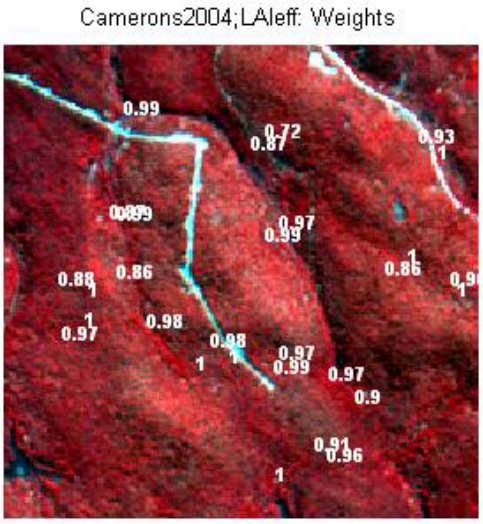


Figure 12. Weights associated to each ESU for the determination of LAIeff transfer function.

For the LAItrue, the XS1, XS2, XS3, XS4, RN combination on reflectance (Figure 13 and Figure 14) was selected since it provides a good compromise between the cross-validation RMSE, the weighted RMSE and the RMSE (lowest value). Note that one weight is lower than 0.7.

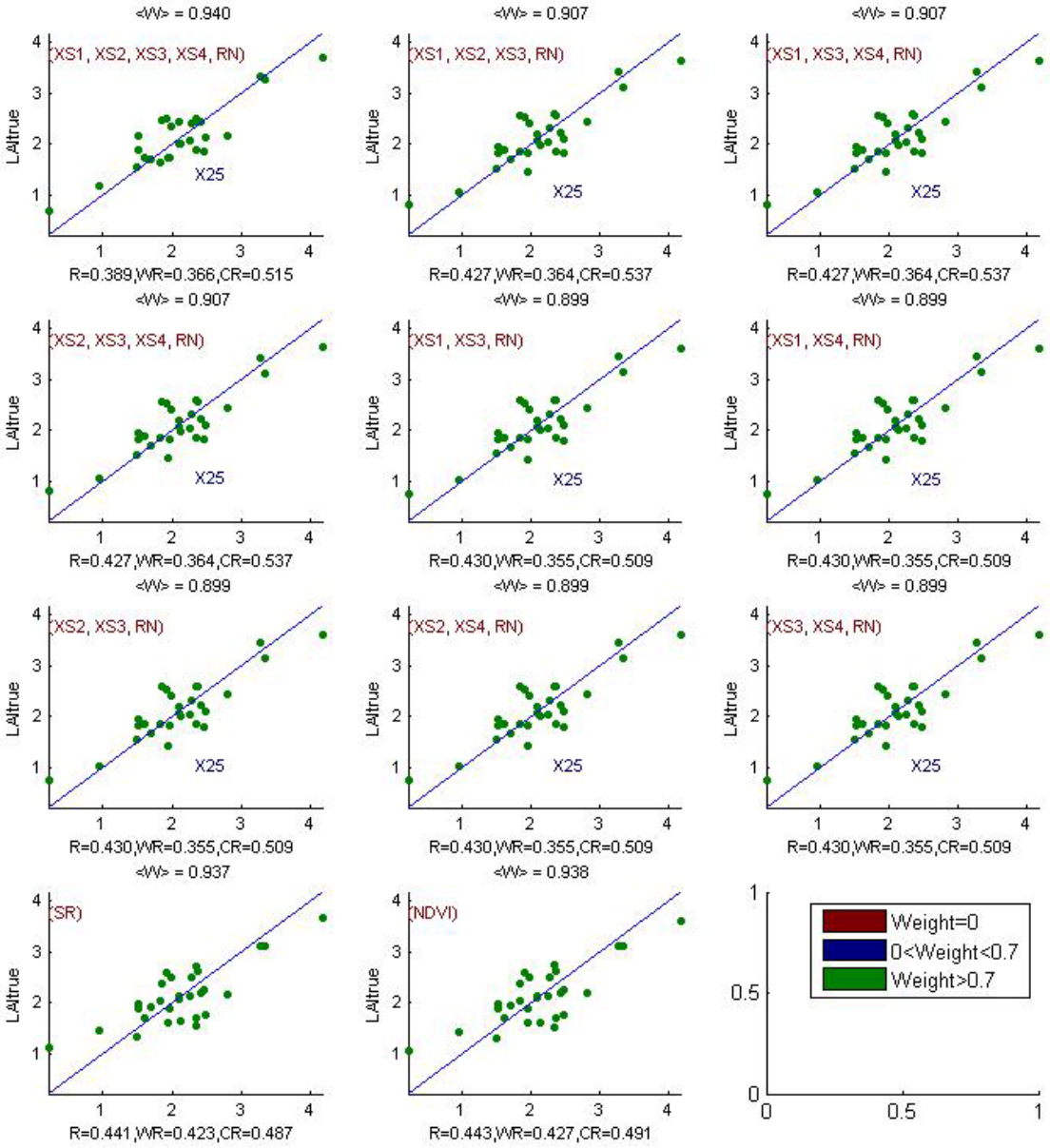


Figure 13. True Leaf Area Index: results for regression on reflectance using different band combinations. R is the root mean square error computed between LAItrue and estimated LAItrue. WR is the weighted root mean square error and CR is the cross validation root mean square error.

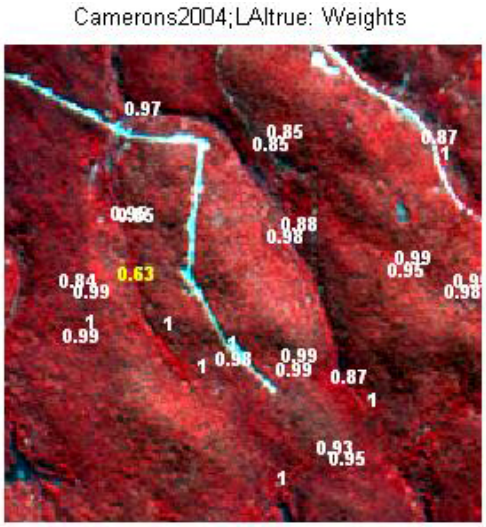


Figure 14. Weights associated to each ESU for the determination of LAItrue transfer function.

For the LAI57eff, the XS1, XS2, XS3, XS4, RN combination on reflectance (Figure 15 and Figure 16) was selected since it provides the best results: cross-validation RMSE (lowest value), weighted RMSE (lowest value) and RMSE (lowest value). Note that zero weight is lower than 0.7.

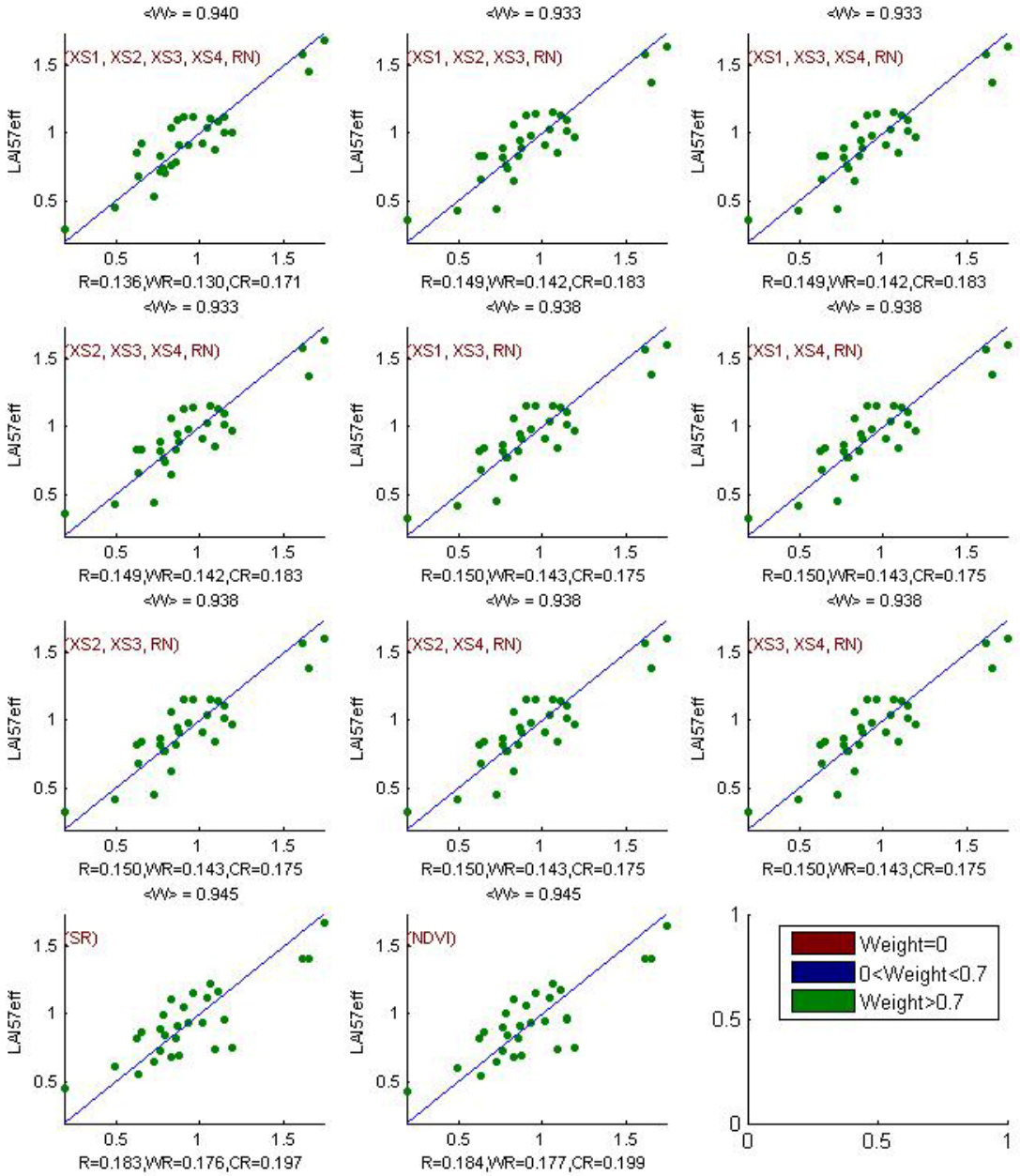


Figure 15. Effective LAI at 57.5°: results for regression on reflectance using different band combinations. R is the root mean square error computed between LAI57eff and estimated LAI57eff. WR is the weighted root mean square error and CR is the cross validation root mean square error.

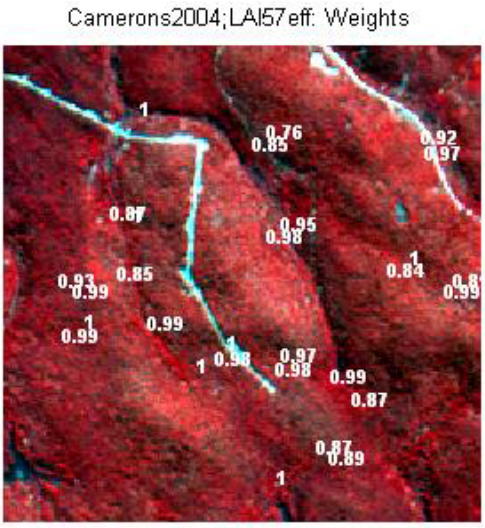


Figure 16. Weights associated to each ESU for the determination of LAI57eff transfer function.

For the LAI57true, the XS1, XS2, XS3, XS4, RN combination on reflectance (Figure 17 and Figure 18) was selected since it provides good results: cross-validation RMSE (among the lowest values), weighted RMSE (lowest value) and RMSE (lowest value). Note that zero weight is lower than 0.7.

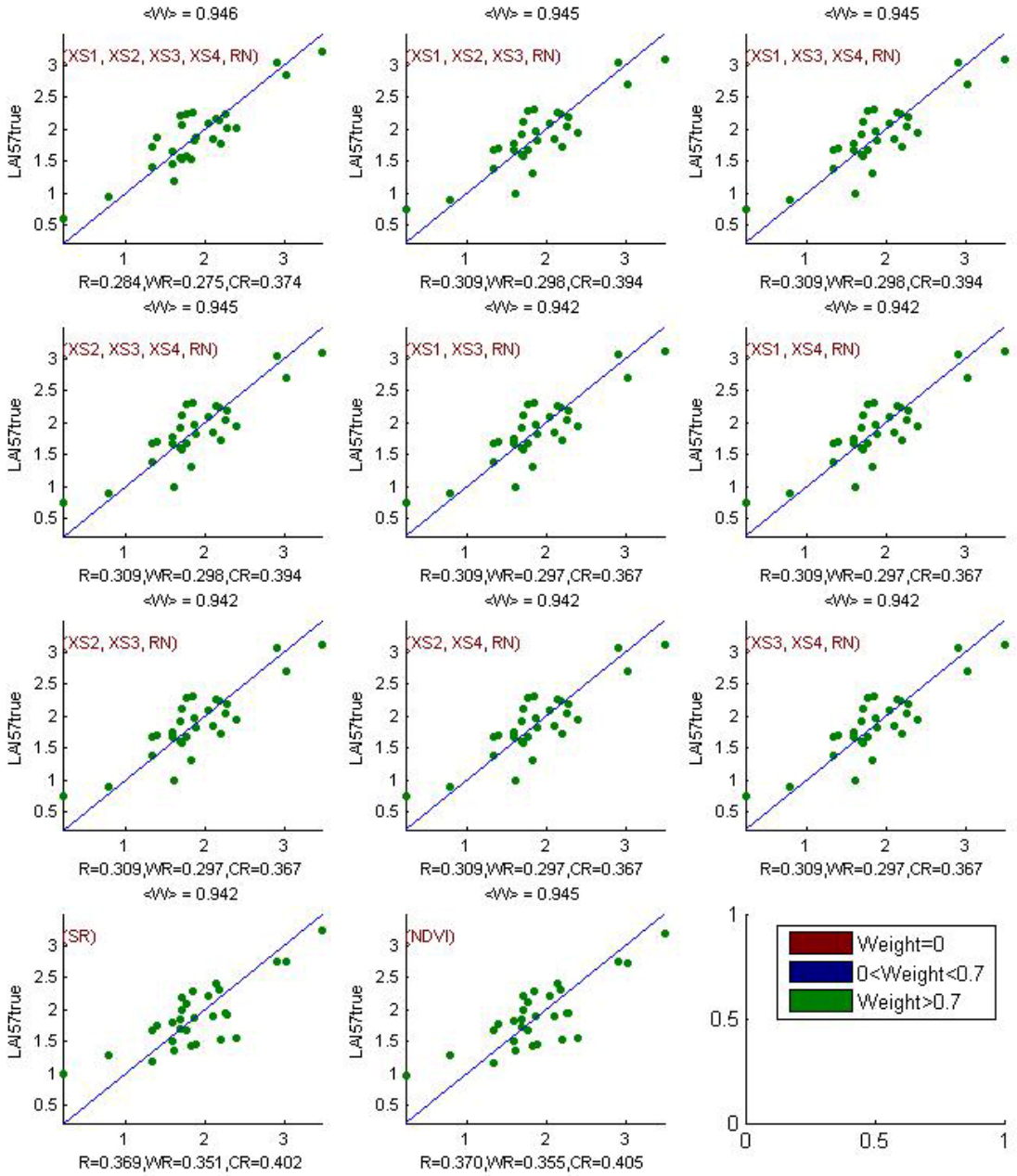


Figure 17. True Leaf Area Index at 57.5°: results for regression on reflectance using different band combinations. R is the root mean square error computed between LAI57true and estimated LAI57true. WR is the weighted root mean square error and CR is the cross validation root mean square error.

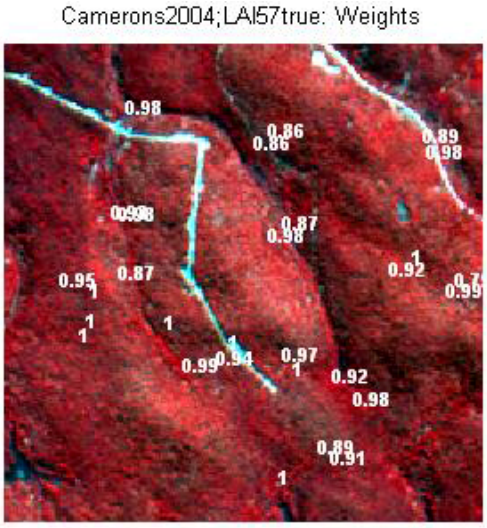


Figure 18. Weights associated to each ESU for the determination of LAI57true transfer function.

For the fCover, the NDVI combination on reflectance (Figure 19 and Figure 20) was selected since it provides a good compromise between the cross-validation RMSE (lowest value), the weighted RMSE and the RMSE. Note that two weights are lower than 0.7.

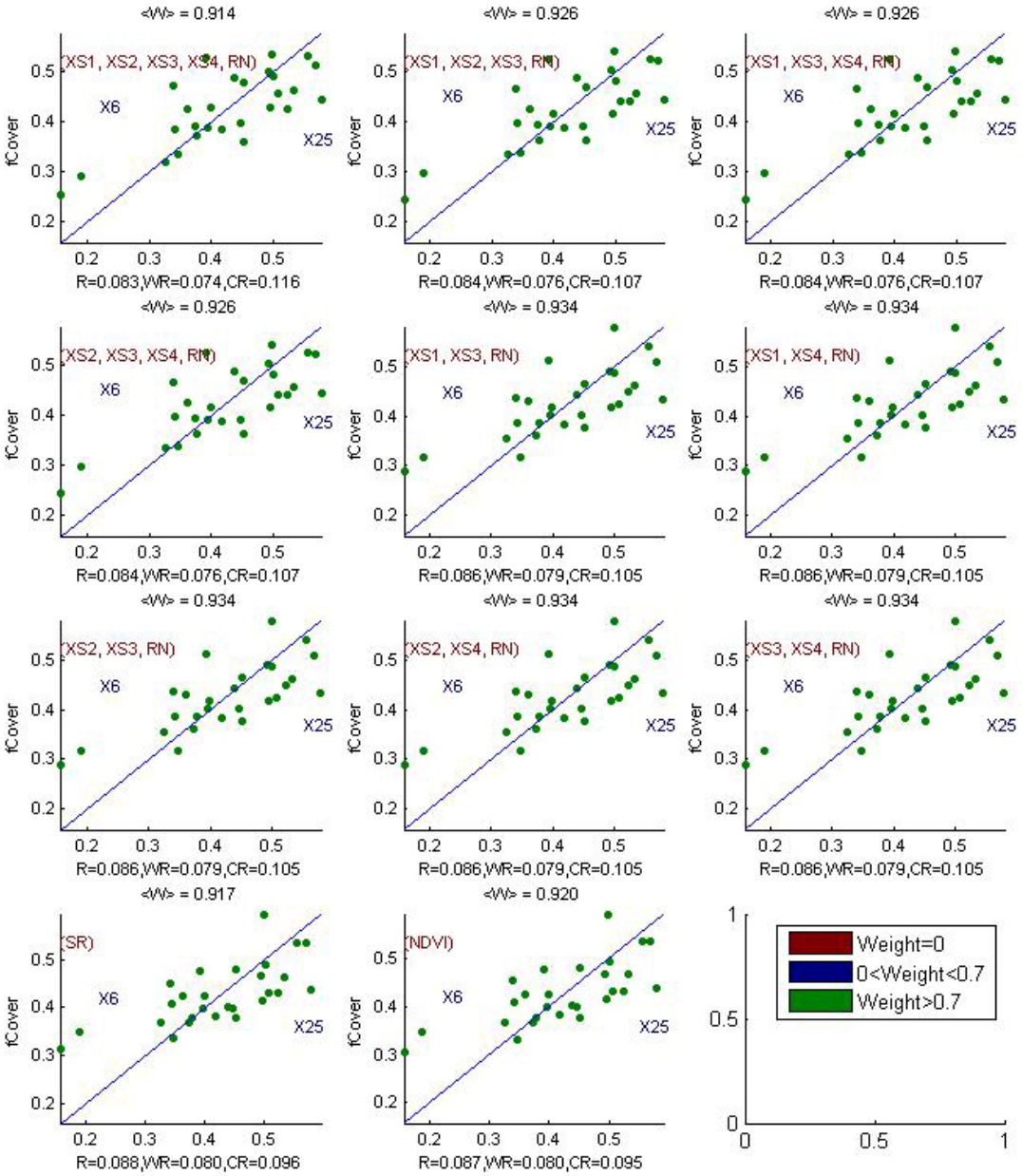


Figure 19. fCover: results for regression on reflectance using different band combinations. R is the root mean square error computed between fCover and estimated fCover. WR is the weighted root mean square error and CR is the cross validation root mean square error.

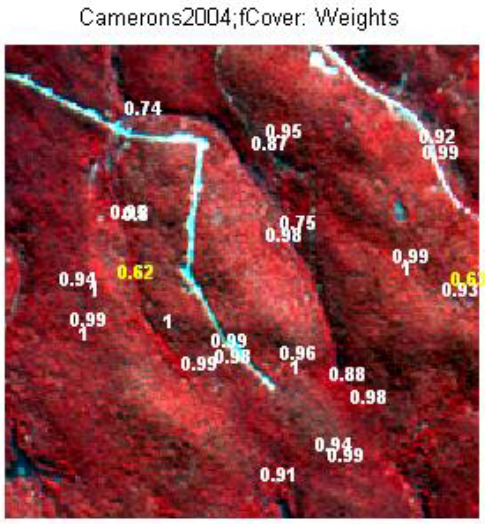


Figure 20. Weights associated to each ESU for the determination of fCover transfer function.

For the fAPAR, the XS2, XS3, XS4, RN combination on reflectance (Figure 21 and Figure 22) was selected since it provides the best results: cross-validation RMSE (lowest value), weighted RMSE (lowest value) and RMSE (lowest value). Note that one weight is lower than 0.7.

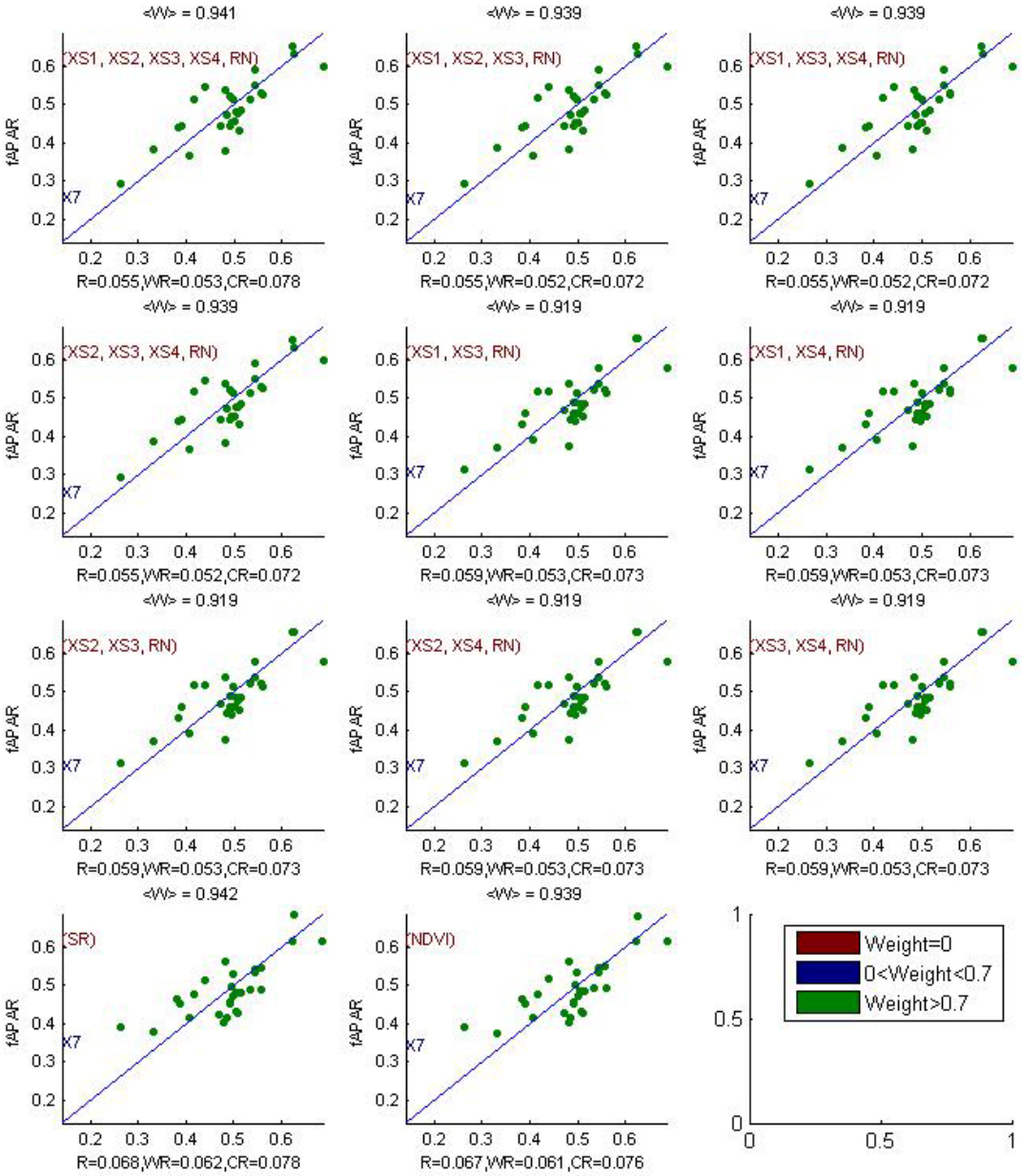


Figure 21. fAPAR: results for regression on reflectance using different band combinations. R is the root mean square error computed between fAPAR and estimated fAPAR. WR is the weighted root mean square error and CR is the cross validation root mean square error.

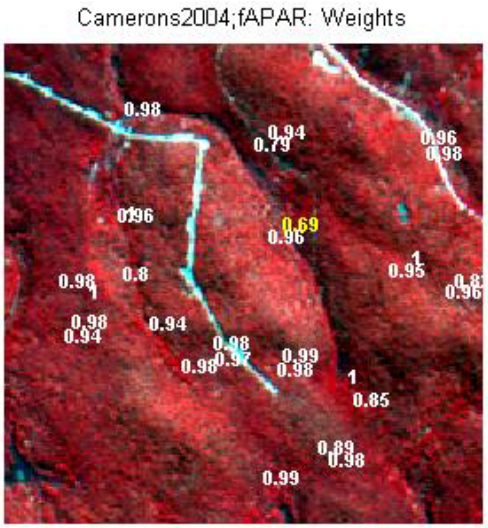


Figure 22. Weights associated to each ESU for the determination of fAPAR transfer function.

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

Variable	Band Combination	RMSE	Weighted RMSE	CR RMSE
LAIeff	7.1765 - 49.3027(XS1) - 49.9125(XS2) + 0.5391(XS3) - 9.3063(XS4) + 189.0003 (RN)	0.155	0.148	0.192
LAItrue	15.8733 - 26.2814(XS1) - 193.6275(XS2) - 27.6143(XS3) - 22.4951(XS4) + 759.7083(RN)	0.389	0.366	0.515
LAI57eff	5.6881 - 43.0014(XS1) - 31.5503(XS2) + 3.3233(XS3) - 9.6208(XS4) + 123.2072(RN)	0.136	0.130	0.171
LAI57true	17.3994 - 77.1595(XS1) - 154.1063(XS2) - 26.5898(XS3) - 19.1863(XS4) + 692.65(RN)	0.284	0.275	0.374
fCover	-0.5204 + 1.9314(NDVI)	0.087	0.080	0.095
fAPAR	6.7525 - 22.1321(XS2) - 76.1718(XS3) - 19.1707(XS4) + 336.8755(RN)	0.055	0.052	0.072

RN = Red*NIR

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 Camerons SPOT image extraction

Figure 23 presents the biophysical variable maps obtained with the transfer function described in Table 2 for the classes 1, 2 and 4 (class 3: LAI = 0; please read §3.1). The maps obtained for the six variables are consistent, showing similar patterns: low LAIeff values where low fCover or fAPAR are observed and conversely… The difference between effective LAI and true LAI is significant (see the average values in

Figure 23). This was expected when looking the LAIeff/LAItrue relationship, showing that for high LAI the difference between the two can be significant.

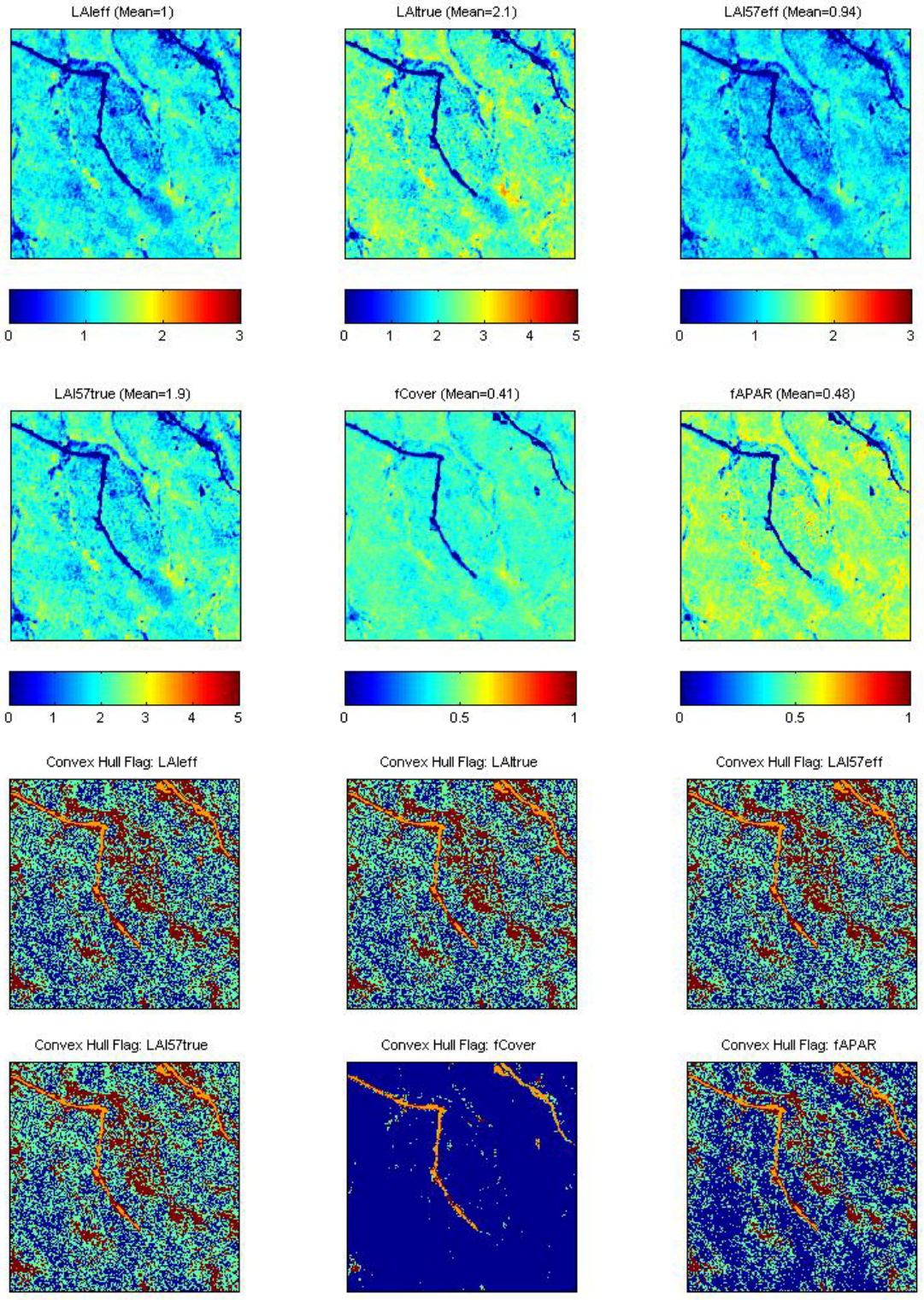


Figure 23. High resolution biophysical variable maps applied on the Camerons site (top). Associated flags are shown: blue and light blue correspond to the pixels belonging to the 'strict' and 'large' convex hulls, red to the pixels for which the transfer function is extrapolating and orange to the pixels for which the value 0 is attributed.

The flag maps are comparable between LAIeff, LAItrue, LAI57eff and LAI57true (the number and the bands used for the regression are the same). For fCover, the pixels outside the 'strict convex-hull' are very few. In theory, the more the number of bands increases, the larger the extrapolation is. The extrapolation mainly corresponds to forest roads, not sampled areas… The orange pixels correspond to forest roads (§2.3.3).

4. Conclusion

The 'REG' method is applied to the classes 1, 2 and 4 by using 29 ESUs, whereas the value 0 is attributed to class 3 which corresponds to wide forest roads (bare soil). The relationship between NDVI and LAI variables is consistent and the representativeness of the land cover of the different ESUs is very good. The results of the robust regression are also good and the maps obtained for the biophysical variables are consistent. The flag associated to each map show that the little extrapolation of the transfer function is mainly bounded to forest roads, areas which have not been sampled (§2.3.2)… For all the variables, the regression coefficients are computed by relating the variable itself to reflectance.

The biophysical variable maps are available in UTM, 50 South, projection coordinates (Datum: WGS-84) at 20 m resolution.

5. Acknowledgements

We thank people who participated to the field experiment: **Michael Martin**, **Chengchao**, **Michael Canci** (Water Corporation), **Ian Colquhoun**, **Naomi Kerp** (Alcoa), **Richard Silberstein**, **Geoff Hodgson**, **Simon Higginson**, **Kieran Coupe** (CSIRO Land and Water), **Jeremy Wallace**, **Graeme Behn** (CSIRO Mathematical and Information sciences).