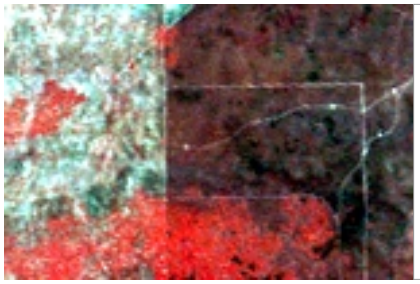
GROUND DATA PROCESSING & PRODUCTION OF THE LEVEL 1 HIGH RESOLUTION MAPS



VALERI 2004

Gnangara site
(evergreen broadleaf forest)

Philippe Rossello

March 2008

CONTENTS

1. Introduction

This report describes the production of high resolution, level 1, biophysical variable maps for the Gnangara site in March 2004. 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 (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 Gnangara site corresponds to "a natural australian forest and shrubland over gently undulated topography. It is characterized by large patches with characteristics depending mainly on past fire events. Due to recent fire events, a large fraction of patches was characterized by very low vegetation amounts with few individual large trees that survived across fires. The understorey is generally sparse and dry. The whole landscape appears mainly clumped at the shoot and tree and patch levels". The ground measurements were carried out from 27th February to 3rd March 2004. The site is approximately 3 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	391620.0000	6512220.0000	-31.51983970	115.85852266
Lower right corner	396180.0000	6509160.0000	-31.54786290	115.90622001
Center	393900.0000	6510690.0000	-31.53385346	115.88236774

Table 1. Description of the site coordinates.

2. Available data

2.1. LANDSAT Image

The LANDSAT 5 TM image (resolution: 30 m) was acquired on 3rd March 2004. It was radiometrically and geometrically corrected by USGS (level 1G systematic). The projection is UTM 50, South, WGS-84. In order to reduce the residual error in the systematic 1G product, a rectification was performed from an aerial image taken in 2003 (CSIRO Land and Water): +135 meters in Easting and +185 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 x 4 km site. Therefore, it will be taken into account everywhere in the same way.

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

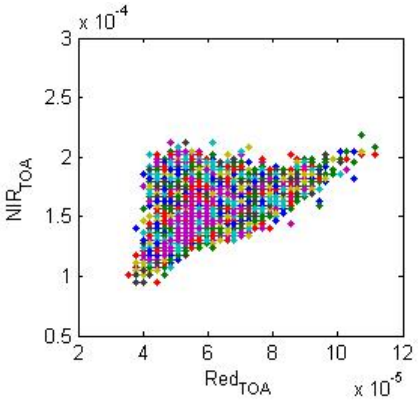


Figure 1. Red/NIR relationship on the LANDSAT image for Gnangara, 2004.

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 31 sampled ESUs. As Gnangara site is mainly covered of broadleaf forest, 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 0.21 to 0.86, while true LAI (LAItrue, LAI57true) varies from 0.45 to 2.3. The site is homogeneous (Figure 2) in terms of LAI, even if the range of true LAI values is relatively large. 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).

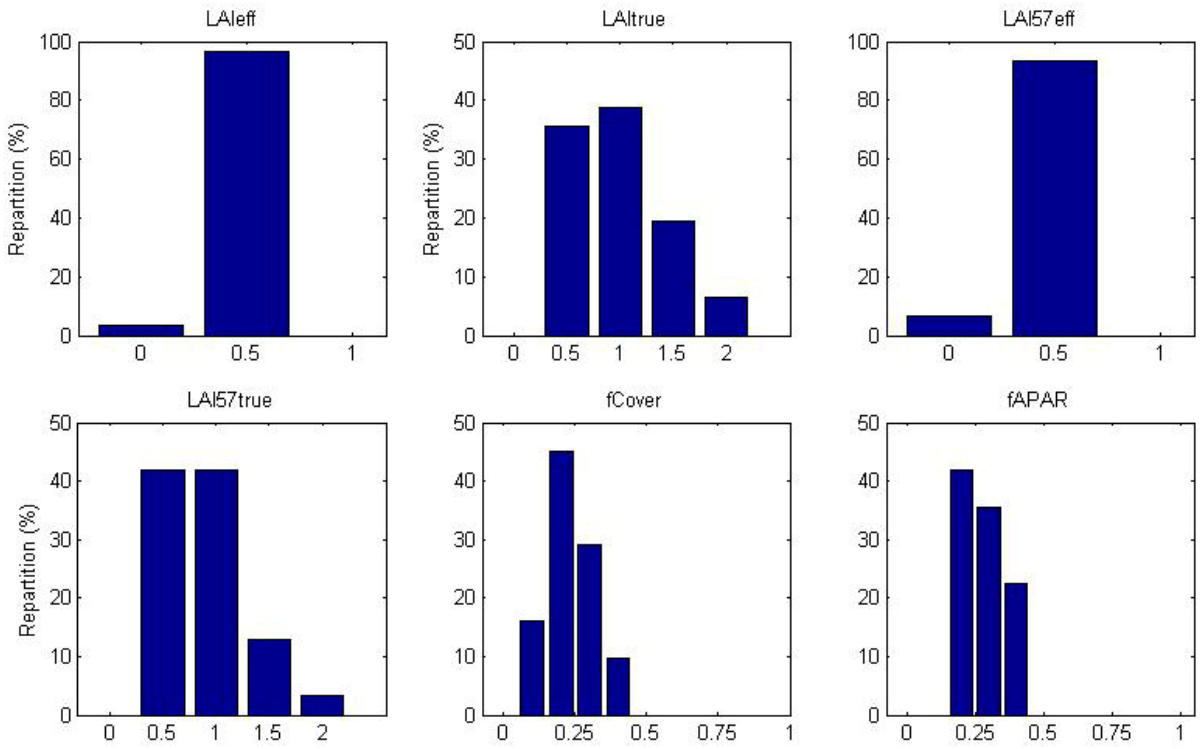


Figure 2. Distribution of the measured biophysical variables over the ESUs.

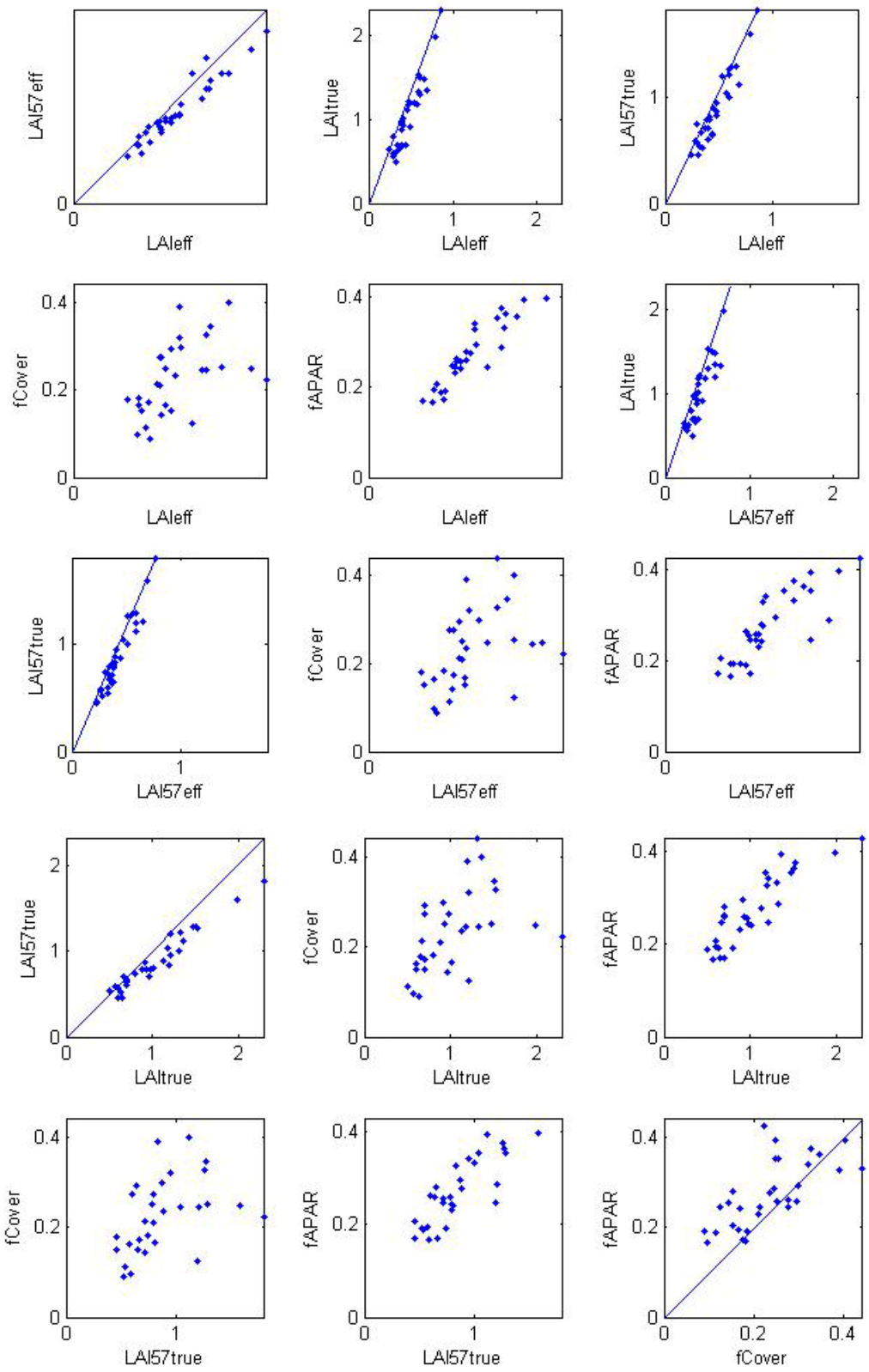


Figure 3. Relationships between the different biophysical variables.

2.3. Sampling strategy

2.3.1. Principles

The sampling of each ESU is based at least on twelve elementary images from above the understorey and from below the canopy. Figure 4 shows that the 31 ESUs are evenly distributed over the site (3 x 4 km). Note that the recently burnt area is mainly sampled on the borders. As the LANDSAT geo-location and the GPS measurements are not associated to errors, all the ESUs have been kept for the computation of the transfer function.

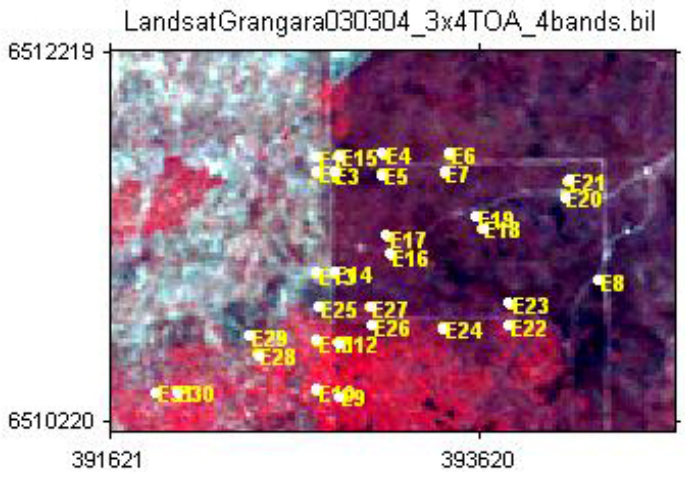


Figure 4. Distribution of the ESUs around the Gnangara site.

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 5). As the number of pixels is drastically different for the ESUs and whole site (WS = 15504 in case of this 3 x 4 km LANDSAT image, resolution 30m), 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.

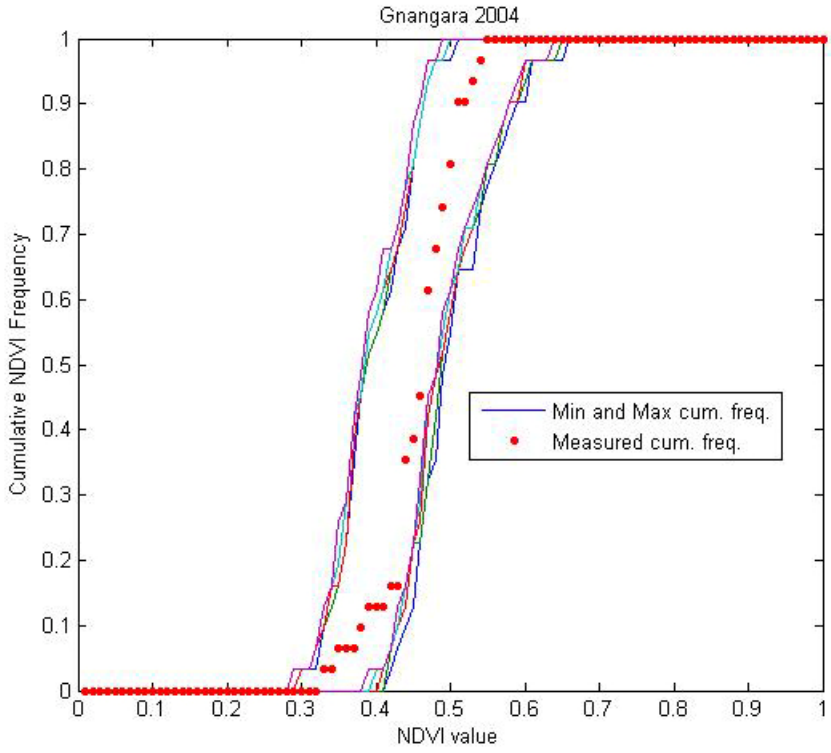


Figure 5. Comparison of the ESU NDVI distribution and the NDVI distribution over the whole image.

Figure 5 shows that the NDVI distribution of the 31 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 close to the boundaries (mainly for NDVI values comprised between 0.43 and 0.46). Note that NDVIs lower than 0.33 have not been sampled although they are present in the image. The site is relatively 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 4 reflectances (band 2 to band 5: http://www.eurimage.com/products/docs/landsat.pdf) 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 6). The distribution of the classes on the image and on the ESUs is rather similar. The classes 2 and 3 are comparable. The class 4 is under-represented, while the class 1 appears to be over-sampled.

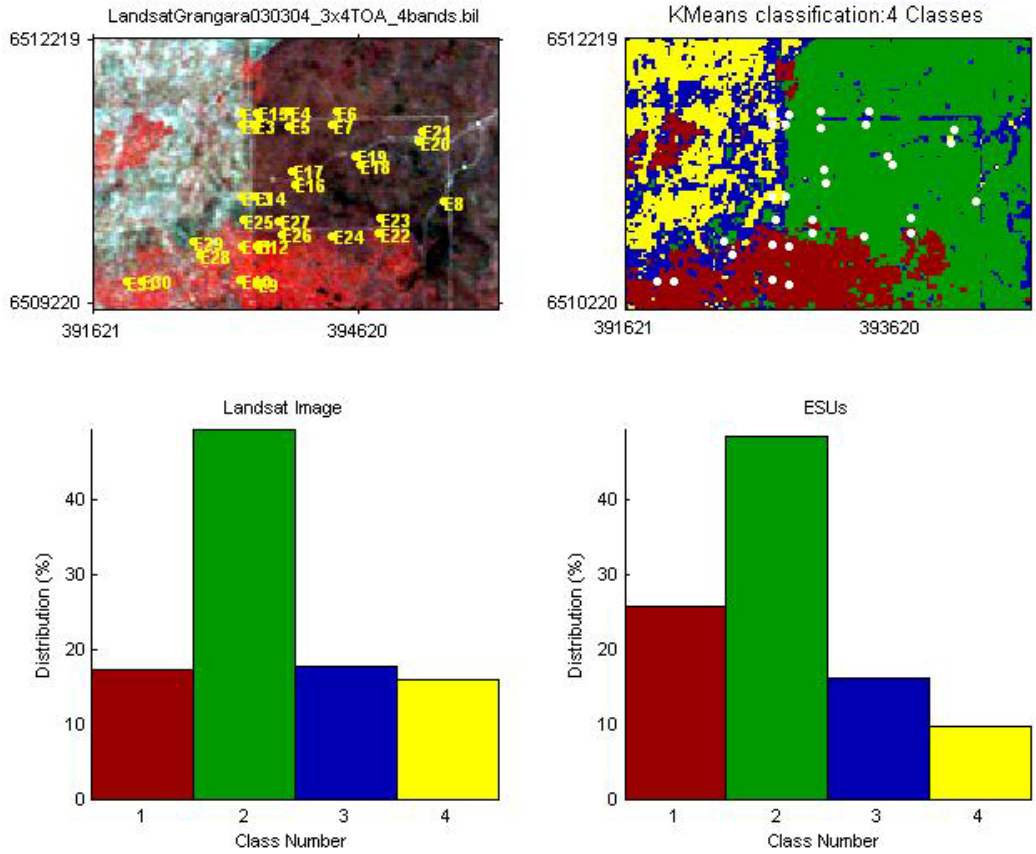


Figure 6. Classification of the LANDSAT image. Comparison of the class distribution between the LANDSAT image and sampled ESUs.

Figure 7 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.

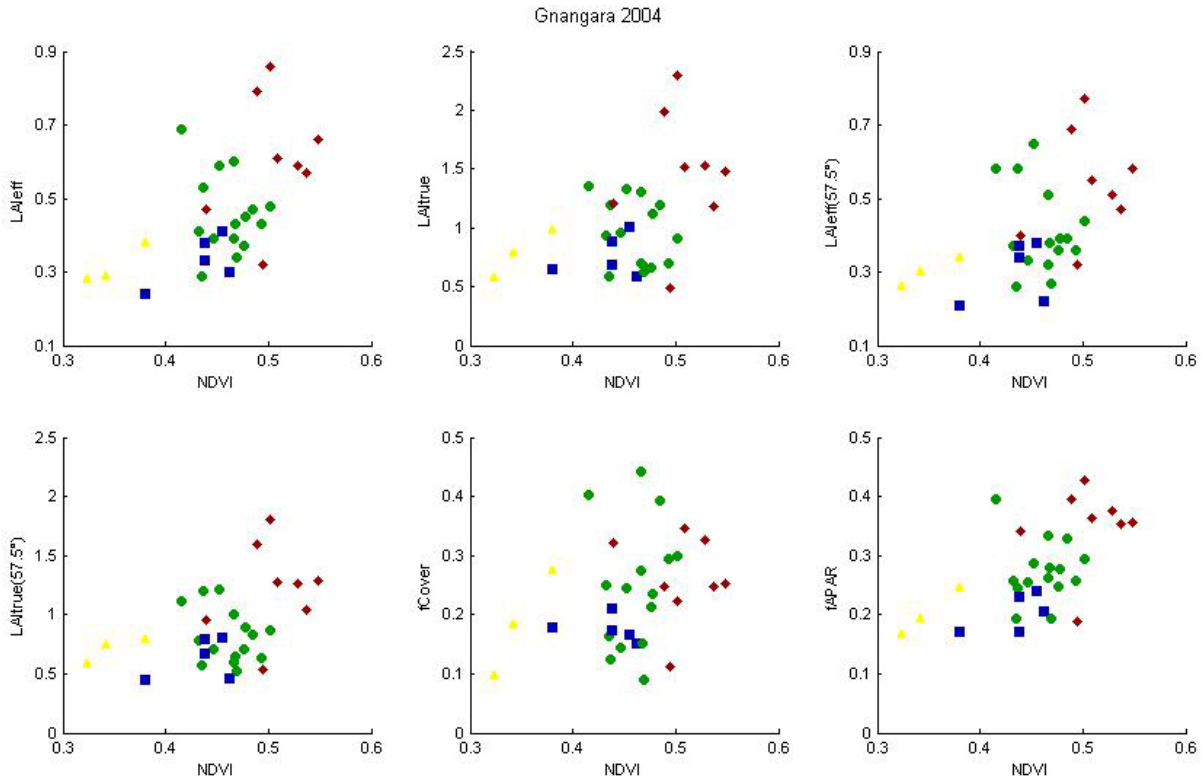


Figure 7. NDVI-Biophysical Variable relationships as a function of LANDSAT classes

The relationship between NDVI and biophysical variables is not very consistent. Even if the NDVI-biophysical variable relationships are not optimal, a single transfer functions will be generated (§3.2.1).

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 four bands of the LANDSAT image (green, red, near infrared, mid infrared). A flag image, is computing over the reflectances (Figure 8). 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.

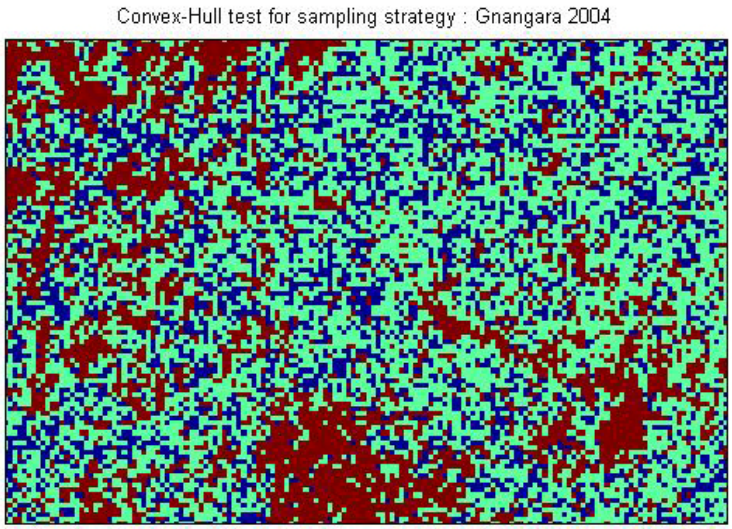


Figure 8. Evaluation of the sampling based on the convex hulls. The map is shown: blue and light blue correspond to the pixels belonging to the 'strict' and 'large' convex hulls and red to the pixels for which the transfer function is extrapolating.

This map shows that the representativeness of the ESUs is not optimal since pixels outside the strict convex-hulls are numerous. However, the extrapolation is not very large. Pixels outside the two convex-hulls correspond to bare soil, lower and higher NDVI pixels, recently burnt areas…

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 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).

Even if the relationship between NDVI and LAI (§2.3.3) is not very consistent, the 'REG' method is applied to all the classes since the results are satisfactory.

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

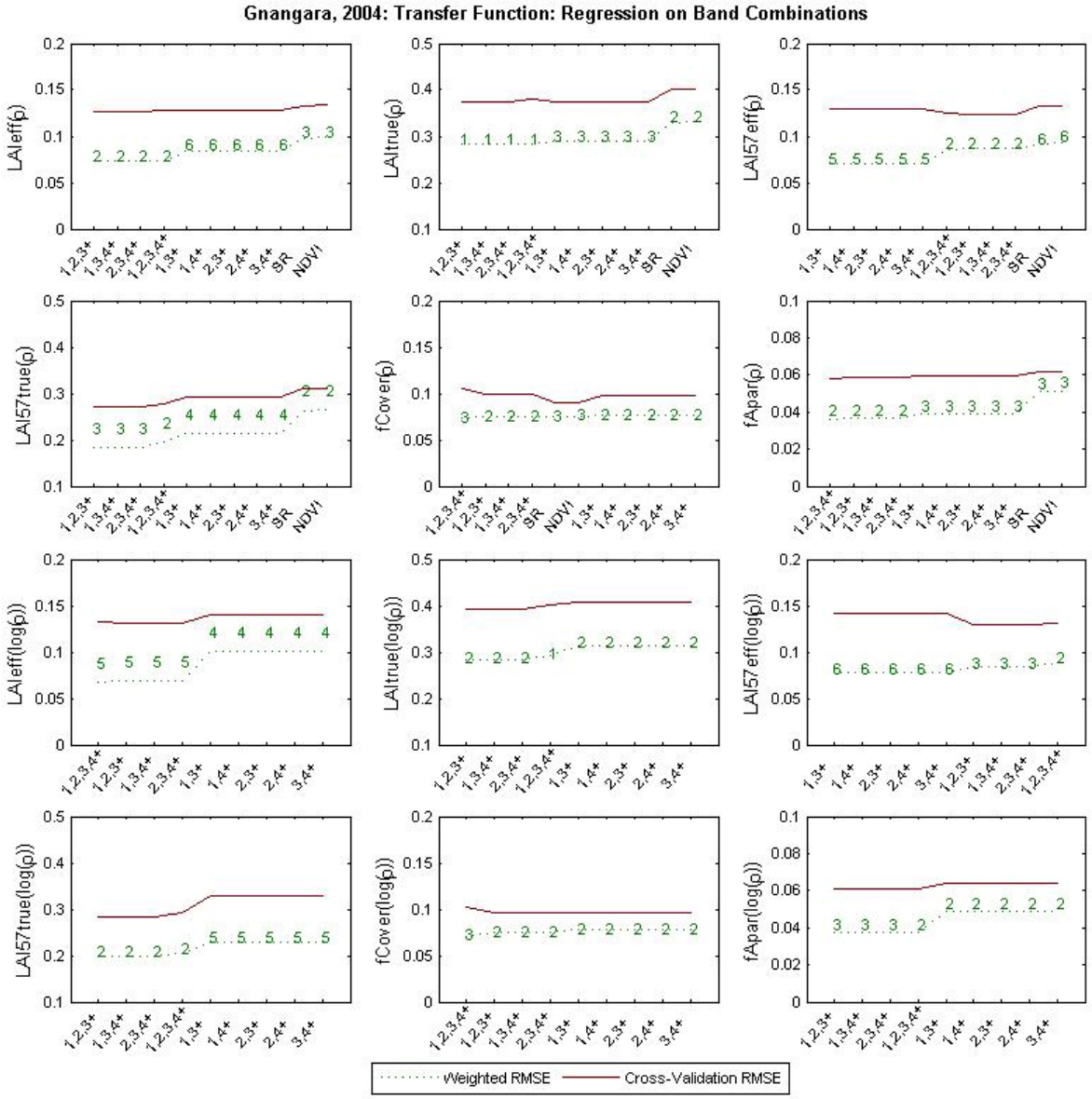
3.2.1. Choice of the method

A single transfer function was thus computed for all the classes.

Figure 9 shows the results obtained for all the possible band combinations using either the reflectance (ρ) or the logarithm of the reflectance (log(ρ)). The results using the reflectance (ρ) were selected for all the variables since they provide the best results (except fCover, but the transfer function using the log(ρ) creates coplanar points which do not allow the determination of the 'strict' and 'large' convex hulls. However, the results are very close).

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).

The LANDSAT bands correspond to: 1 = green (band 2) ; 2 = red (band 3) ; 3 = near infrared (band 4) ; 4 = mid infrared (band 5) ; RN = red* near infrared.



1: green (band 2) ; 2: red (band 3) ; 3: near infrared (band 4) ; 4: mid infrared (band 5) ; RN: red*NIR

Figure 9. 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 B3, B4, B5, RN combination on reflectance (Figure 10 and Figure 11) was selected since it provides the best results. Note that one weight is lower than 0.7. The following band combinations provide the same results: [B2, B3, B4, RN]; [B2, B4, B5, RN].

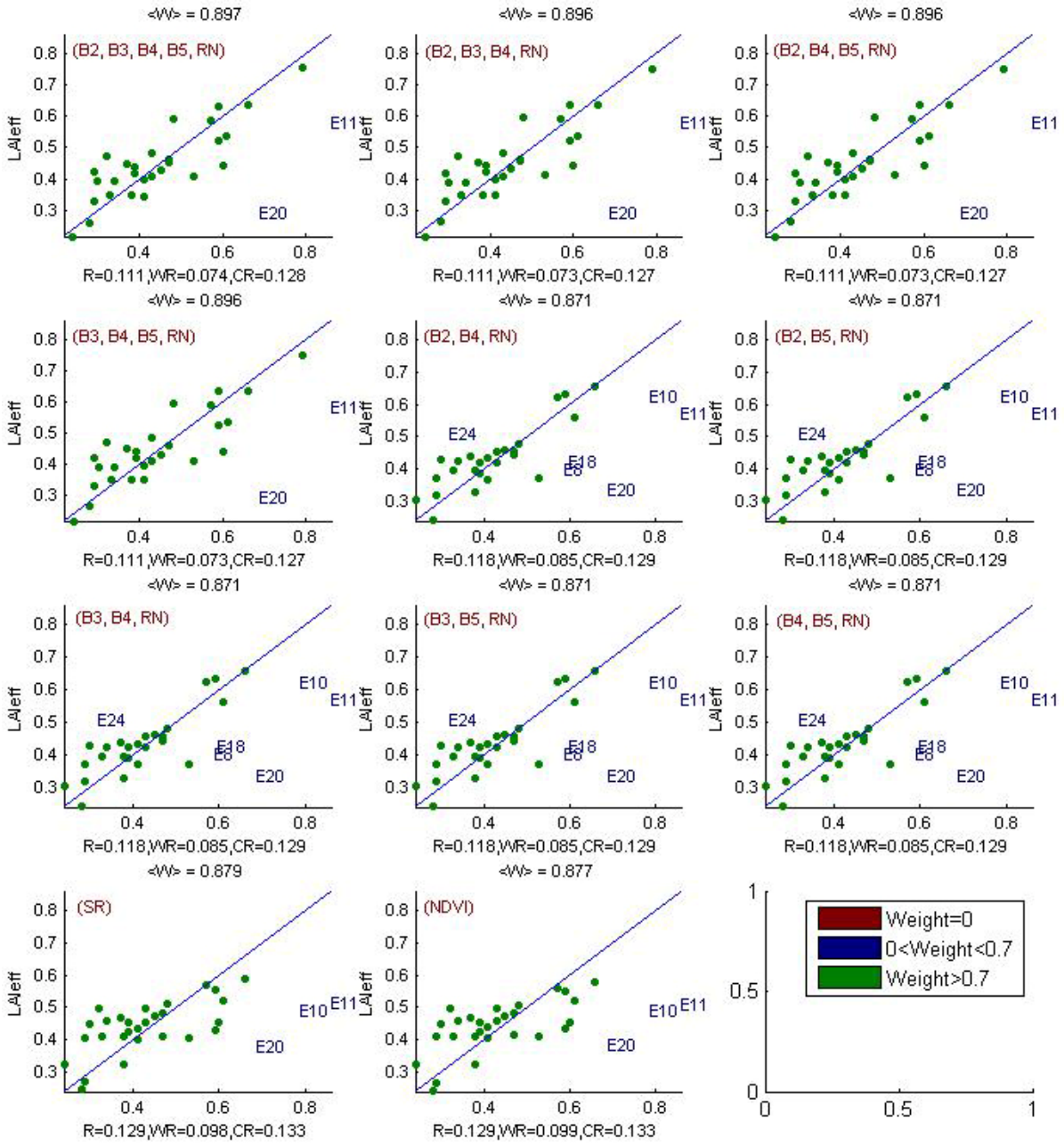


Figure 10. Effective Leaf Area Index: results for regression on reflectance using different band combinations. R is the root mean square error computed between LAIeff and estimated LAIeff. WR is the weighted root mean square error and CR is the cross validation root mean square error.

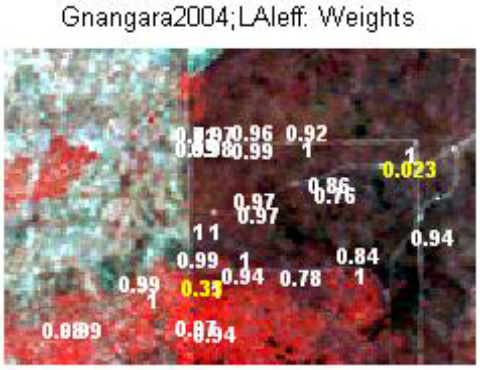


Figure 11. Weights associated to each ESU for the determination of LAIeff transfer function.

For the LAItrue, the B3, B4, B5, RN combination on reflectance (Figure 12 and Figure 13) was selected since it provides the best results. Note that one weight is lower than 0.7. The following band combinations provide the same results: [B2, B3, B4, RN]; [B2, B4, B5, RN].

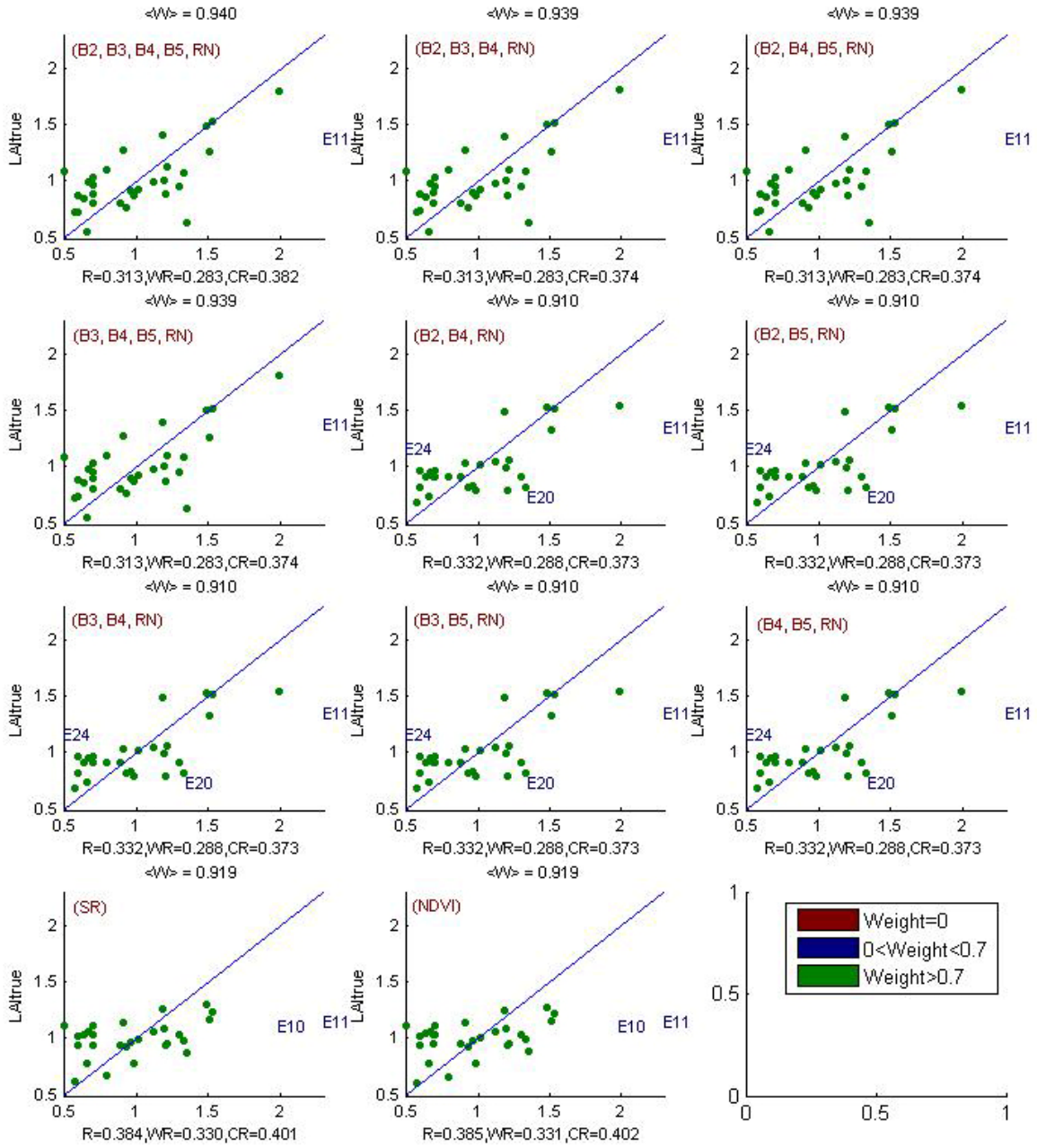


Figure 12. True Leaf Area Index: results for regression on reflectance using different band combinations. R is the root mean square error computed between LAItrue and estimated LAItrue. WR is the weighted root mean square error and CR is the cross validation root mean square error.

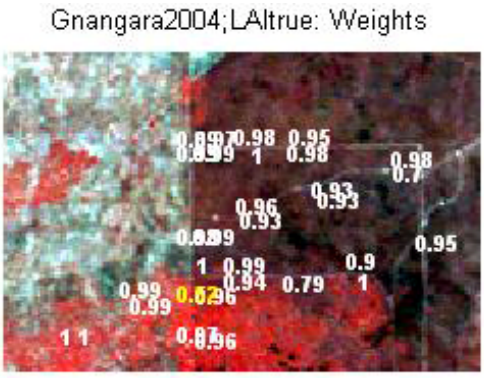


Figure 13. Weights associated to each ESU for the determination of LAItrue transfer function.

For the LAI57eff, the B3, B4, B5, RN combination on reflectance (Figure 14 and Figure 15) was selected since it provides the best results. Note that two weights are lower than 0.7. The following band combinations provide the same results: [B2, B3, B4, RN]; [B2, B4, B5, RN].

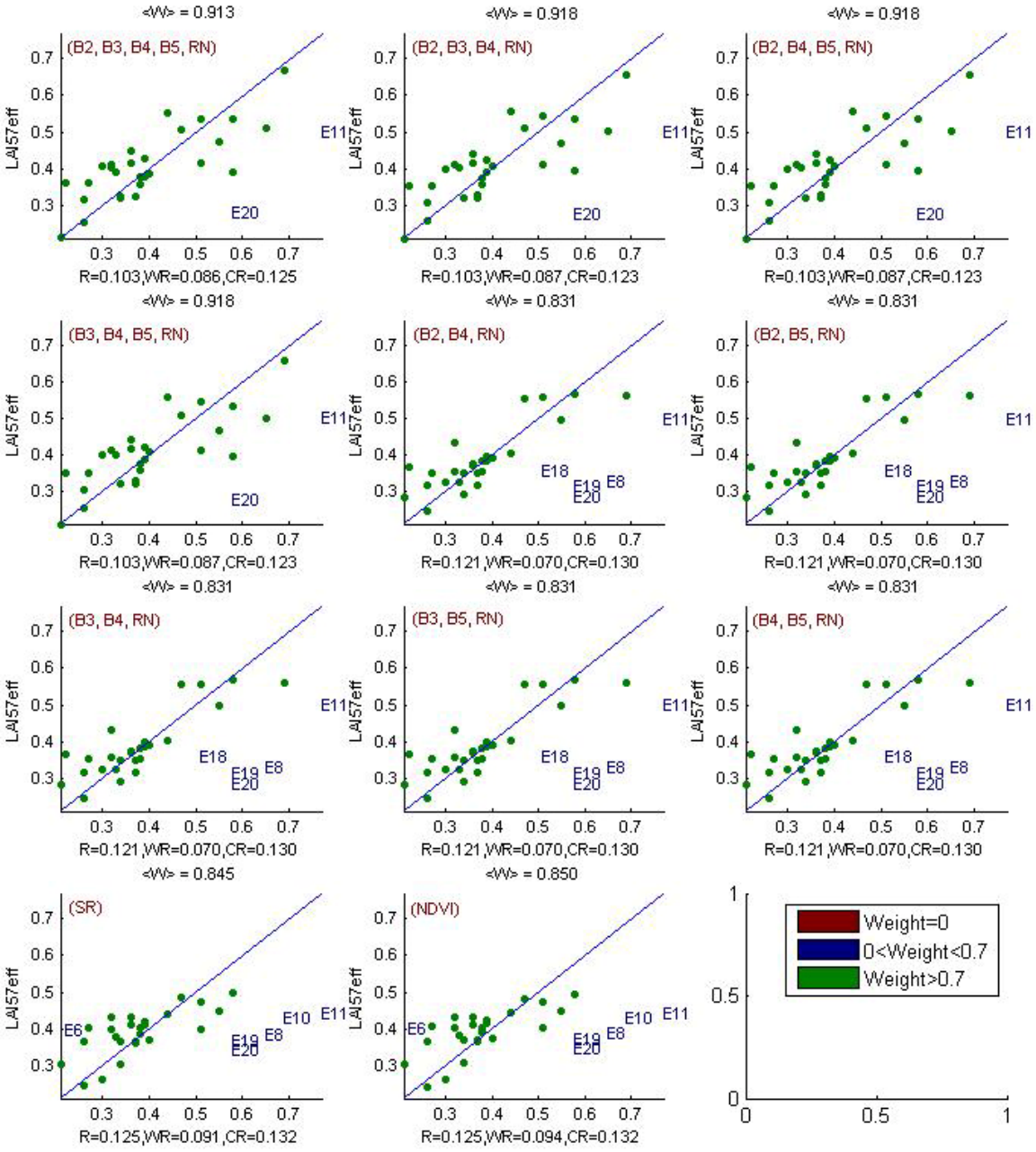


Figure 14. Effective LAI at 57.5°: results for regression on reflectance using different band combinations. R is the root mean square error computed between LAI57eff and estimated LAI57eff. WR is the weighted root mean square error and CR is the cross validation root mean square error.

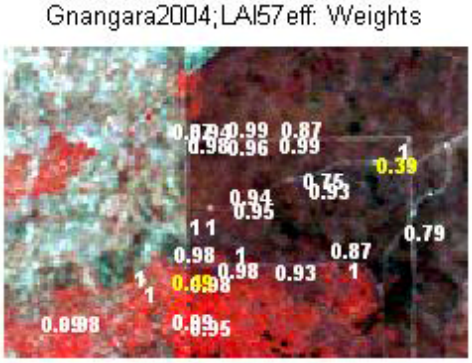


Figure 15. Weights associated to each ESU for the determination of LAI57eff transfer function.

For the LAI57true, the B3, B4, B5, RN combination on reflectance (Figure 16 and Figure 17) was selected since it provides the best results. Note that three weights are lower than 0.7. The following band combinations provide the same results: [B2, B3, B4, RN]; [B2, B4, B5, RN].

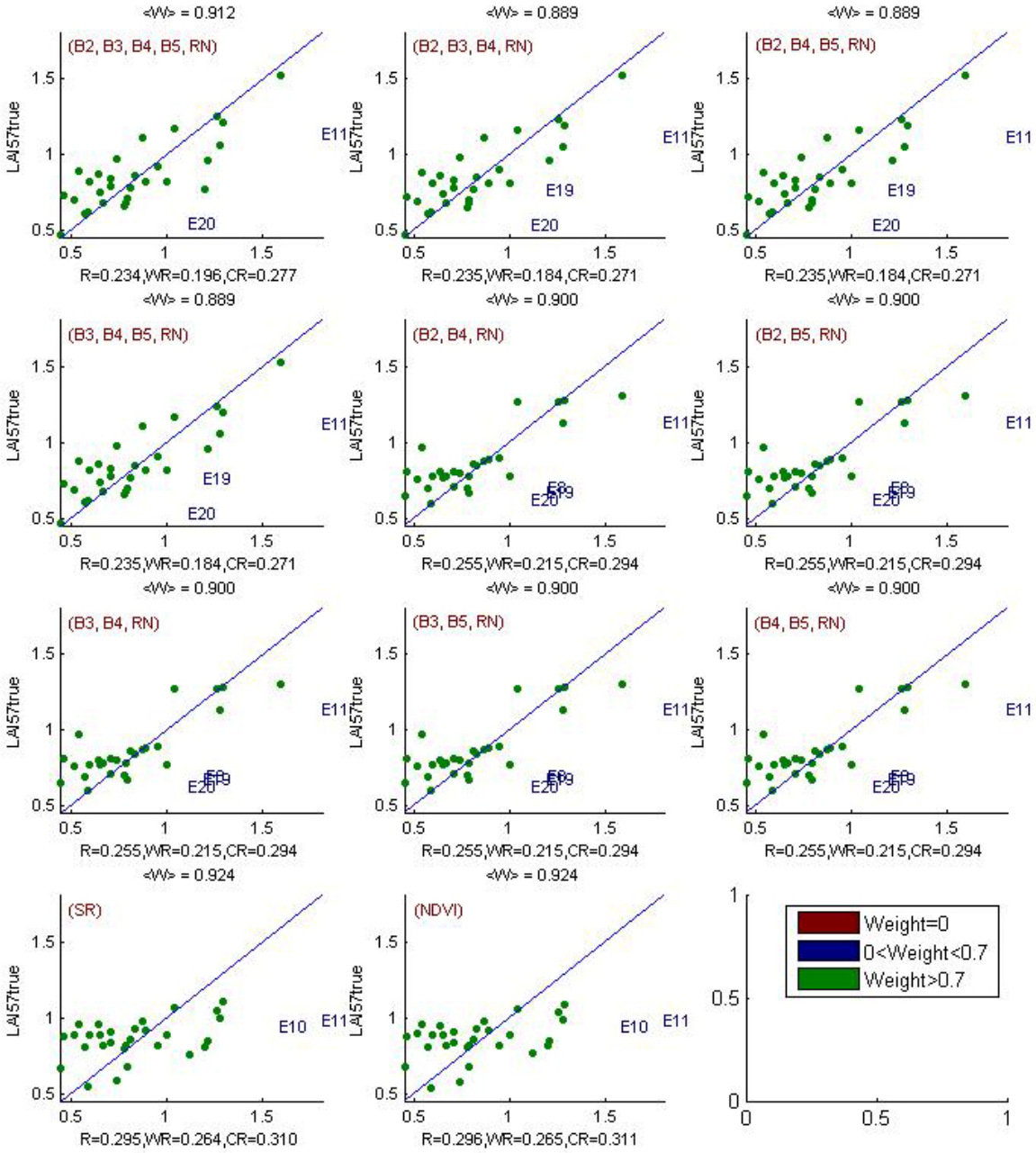


Figure 16. True Leaf Area Index at 57.5°: results for regression on reflectance using different band combinations. R is the root mean square error computed between LAI57true and estimated LAI57true. WR is the weighted root mean square error and CR is the cross validation root mean square error.

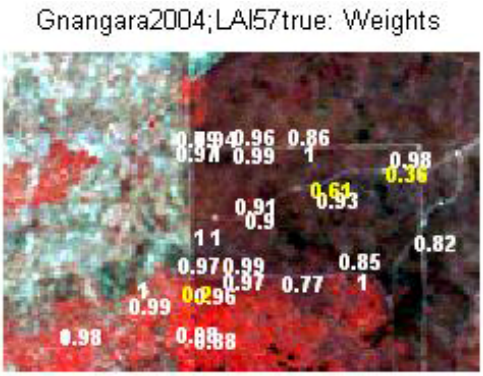


Figure 17. Weights associated to each ESU for the determination of LAI57true transfer function.

For the fCover, the B3, B4, B5, RN combination on reflectance (Figure 18 and Figure 19) was selected since it provides a good compromise between the cross-validation RMSE, the weighted RMSE and the RMSE (lowest value). Note that two weights are lower than 0.7. The following band combinations provide the same results: [B2, B3, B4, RN]; [B2, B4, B5, RN].

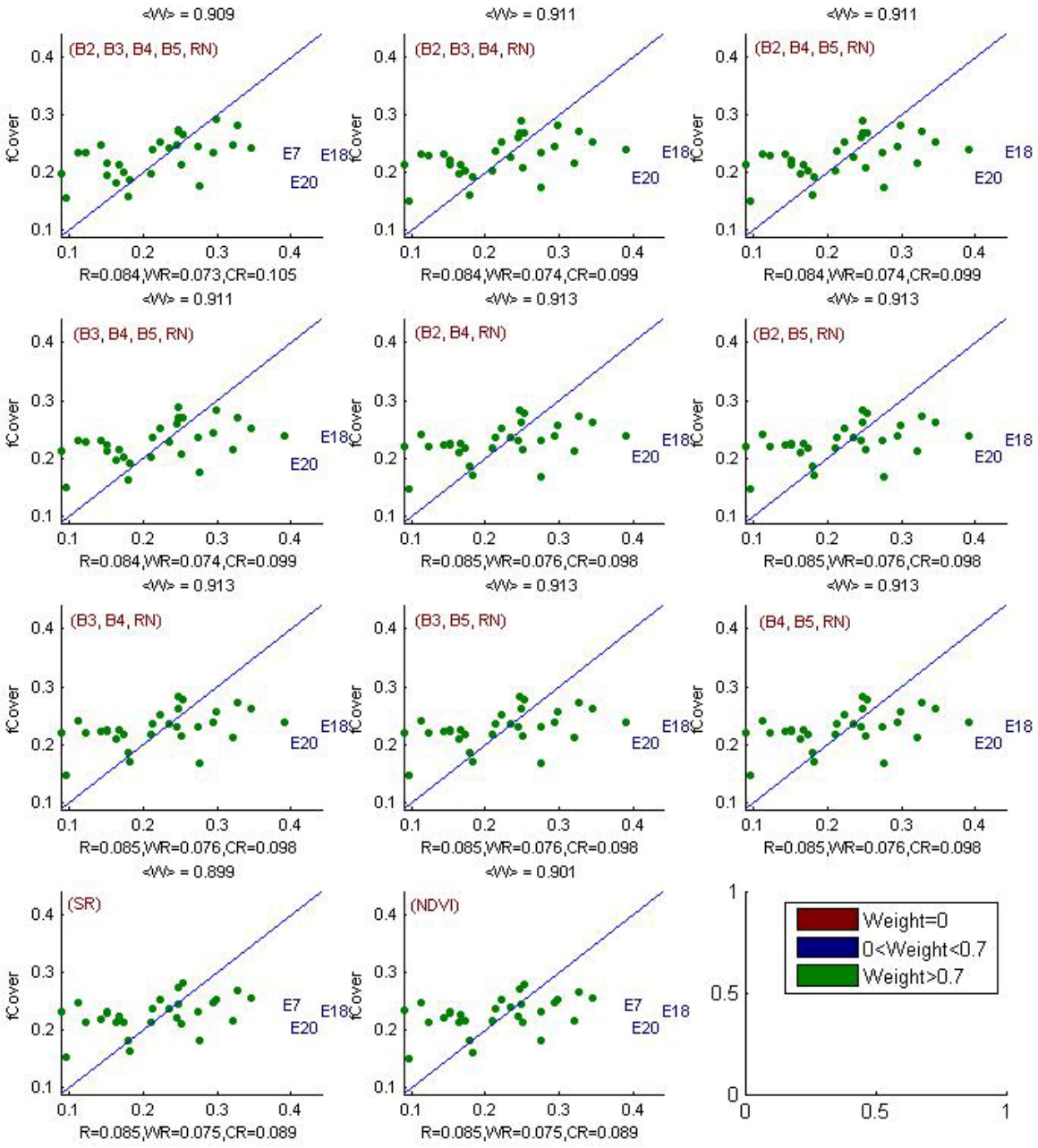


Figure 18. fCover: results for regression on reflectance using different band combinations. R is the root mean square error computed between fCover and estimated fCover. WR is the weighted root mean square error and CR is the cross validation root mean square error.

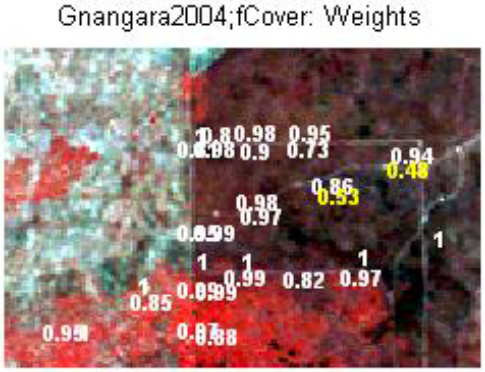


Figure 19. Weights associated to each ESU for the determination of fCover transfer function.

For the fAPAR, the B2, B3, B4, B5, RN combination on reflectance (Figure 20 and Figure 21) was selected since it provides the best results. Note that two weights are lower than 0.7.

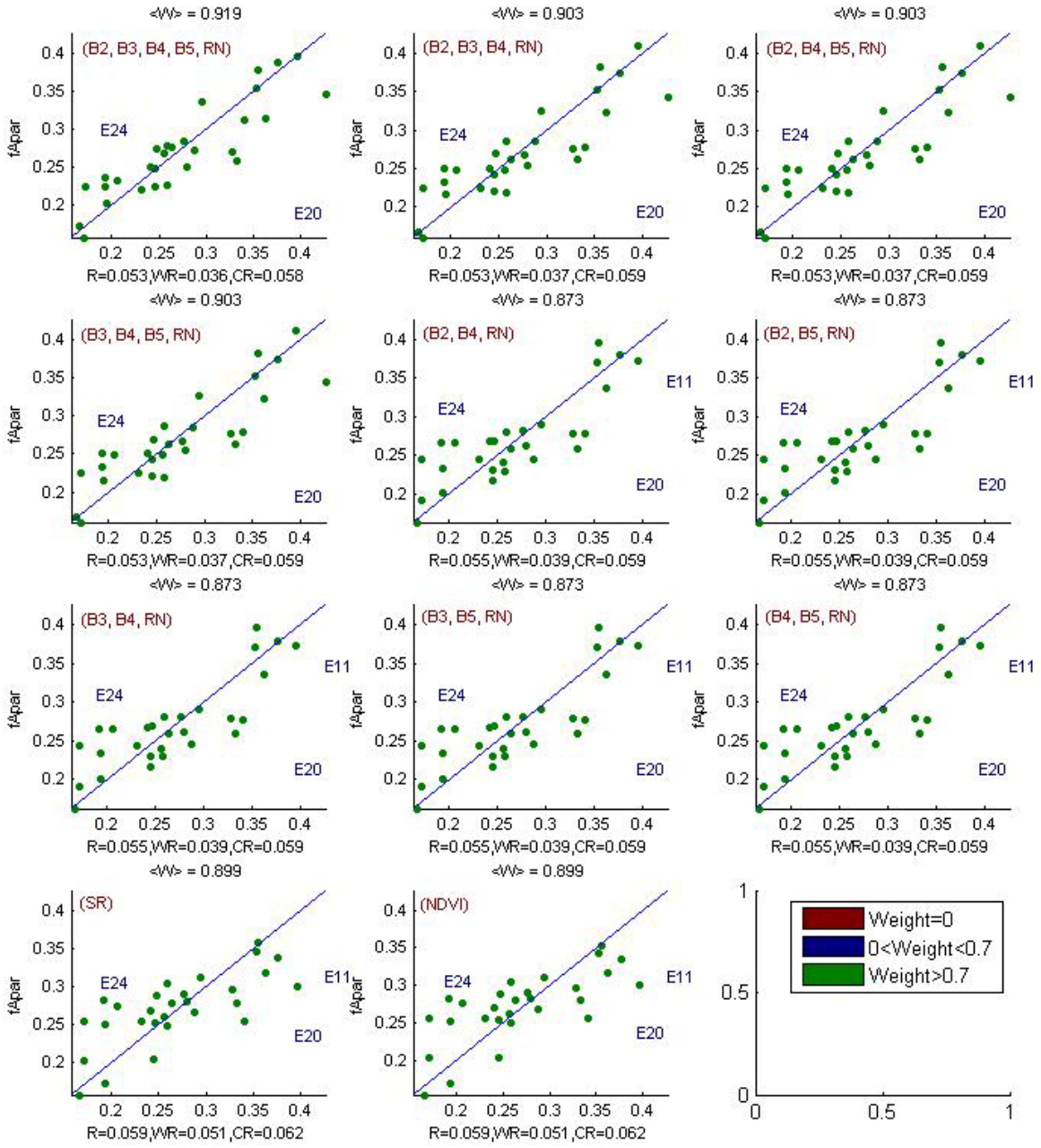


Figure 20. fAPAR: results for regression on reflectance using different band combinations. R is the root mean square error computed between fAPAR and estimated fAPAR. WR is the weighted root mean square error and CR is the cross validation root mean square error.

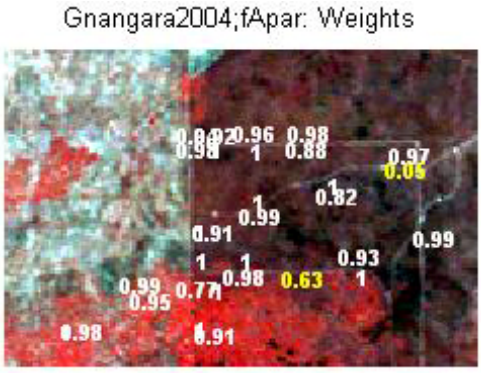


Figure 21. Weights associated to each ESU for the determination of fAPAR transfer function.

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

Variable	Band Combination	RMSE	Weighted RMSE	CR RMSE
LAIeff	4.6171 - 8.9358(B3) - 83.3589(B4) - 21.4752(B5) + 502.5433(RN)	0.111	0.073	0.127
LAItrue	10.0147 - 28.4446(B3) - 175.3893(B4) - 45.7155(B5) + 1112.5276(RN)	0.313	0.283	0.374
LAI57eff	4.493 - 9.5144(B3) - 76.9143(B4) - 21.5994(B5) + 476.6656(RN)	0.103	0.087	0.123
LAI57true	9.4802 - 34.9262(B3) - 151.7618(B4) - 43.965(B5) + 1017.5565(RN)	0.235	0.184	0.271
fCover	1.3646 - 3.9966(B3) - 20.268(B4) - 5.3122(B5) + 122.8487(RN)	0.084	0.074	0.099
fAPAR	1.6534 - 1.5862(B2) - 27.0411(B3) - 6.3416(B4) - 1.6075(B5) + 164.6202(RN)	0.053	0.036	0.058

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 Gnangara LANDSAT image extraction

Figure 22 presents the biophysical variable maps obtained with the transfer function described in Table 2 for all the classes. 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 22). This was expected when looking the LAIeff/LAItrue relationship, showing that for high LAI the difference between the two can be significant.

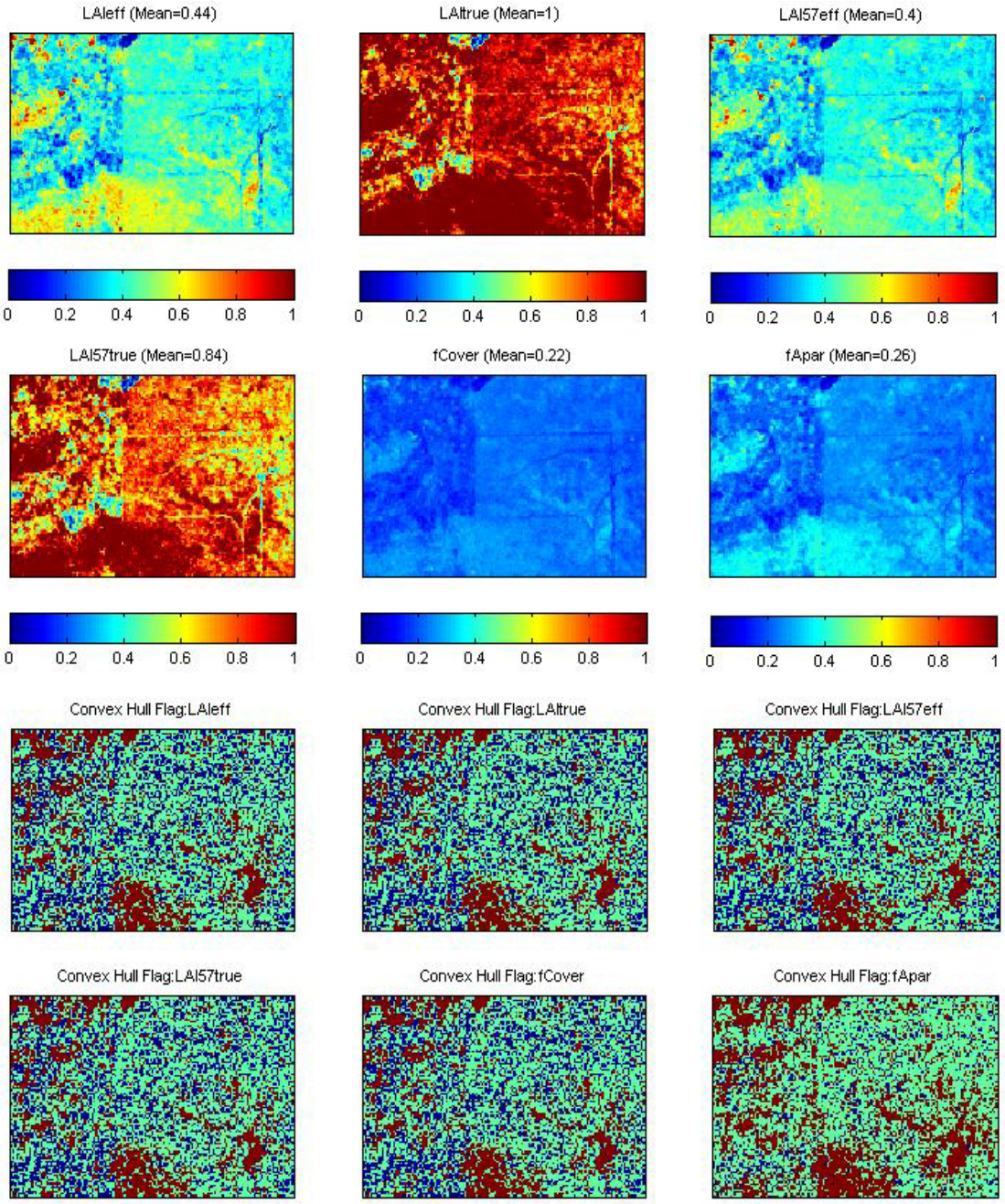


Figure 22. High resolution biophysical variable maps applied on the Gnangara site (top). Associated flags are shown: blue and light blue correspond to the pixels belonging to the 'strict' and 'large' convex hulls, red to the pixels for which the transfer function is extrapolating and orange to the pixels for which the value 0 is attributed.

The flag maps are comparable between LAIeff, LAItrue, LAI57eff, LAI57true and fCover (the number and the bands used for the regression are the same). In theory, the more the number of bands increases, the larger the extrapolation is. The extrapolation mainly corresponds to bare soil, lower and higher NDVI pixels, recently burnt areas… For fAPAR, the pixels inside the 'strict convex-hull' are very few.

4. Conclusion

The 'REG' method is applied to all the classes by using 31 ESUs. The relationship between NDVI and LAI variables is not very consistent but the representativeness of the land cover of the different ESUs is good. The results of the robust regression are satisfactory 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, lower and higher NDVI pixels (§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: **Ken McIntosh, Lidia Boniecka, Ian Colquhoun, Naomi Kerp** (Department of Environment, ALCOA), **Richard Silberstein**, **Geoff Hodgson**, **Simon Higginson**, **Kieran Coupe** and **Jeremy Wallace** (Floreat Laboratory, CSIRO), **Michael Martin**, **Chengchao**, **Michael Canci** (Water Corporation) and **Frédéric Baret** (INRA, Avignon).