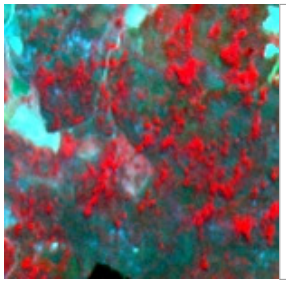
GROUND DATA PROCESSING & PRODUCTION OF THE LEVEL 1 HIGH RESOLUTION MAPS



VALERI 2002

LARZAC site
(grassland)

Philippe Rossello, Marie Weiss

December 2005

CONTENTS

1. Introduction

This report describes the production of the high resolution, level 1, biophysical variable maps for the Larzac site in 2002 (see campaign report for more details about the site and the ground measurement campaign: annex or http://www.avignon.inra.fr/valeri). Level 1 map corresponds to the map derived from the determination of a transfer function between reflectance values of the 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 4.1) 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 description 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 the 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 Yueqin[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). The 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 land cover is mainly composed of grassland. The site is generally flat even if the altitude fluctuates between 760 and 860 m (for more information, see annex or campaign report: http://www.avignon.inra.fr/valeri).

The site coordinates are described in Table 1:

	UTM 31, North, WGS84 (units = meters)		Geographic Lat/Lon WGS84 (units = degrees)	
	Northing	Easting	Lat.	Lon.
Upper left corner	4866450	508350	43.951138	3.104055
Lower right corner	4863450	511350	43.924083	3.141388
Center	4864950	509850	43.937611	3.122722

Table 1. Description of the site coordinates.

2. Available data

2.1. SPOT Image

The SPOT image was acquired the 12th July 2002 by HRV1 on SPOT2 while the ground measurements were carried out from 01/07/2002 to 03/07/2002. The geometrical correction was carried out by the CEFE (http://www.cefe.cnrs.fr). Note that the radiometrical quality is good. The projection is UTM 31 North, WGS-84 (please, refer to the campaign report for more details: annex or http://www.avignon.inra.fr/valeri). 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 3 x 3 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 Yueqin, X., 1986. Estimation of leaf area index from transmission of direct sunlight in discontinuous canopies. Agric. For. Meteorol., 37: 229-243.

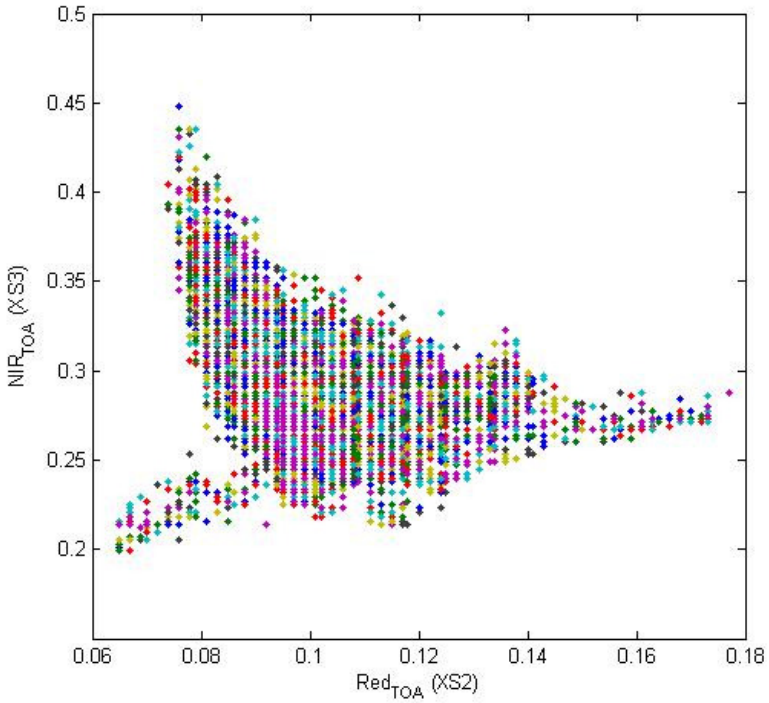


Figure 1. Red/NIR relationship on the SPOT image for Larzac, 2002.

2.2. Hemispherical images

The hemispherical images were processed using the CAN-EYE software (Version 4.1) to derive the biophysical variables. Figure 2 and Figure 3 show the distribution of the several variables over the 27 sampled ESUs. As Larzac is a grassland site, all the hemispherical images were acquired from above the canopy. 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 0.2 to 1.4, while true LAI (LAItrue, LAI57true) varies from 0.24 to 1.78. This range shows a quite homogeneous site in terms of LAI. For values, LAIeff and LAI57eff are lower than LAItrue and LAI57true. This is 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.

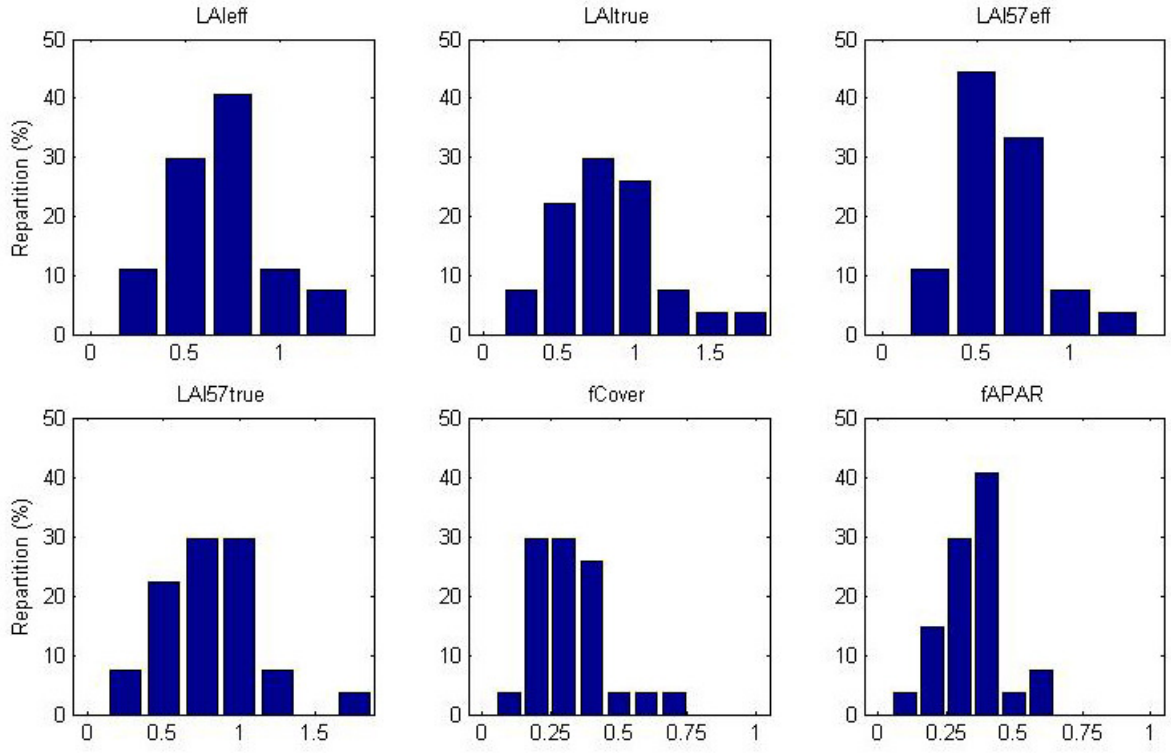


Figure 2. Distribution of the measured biophysical variables over the ESUs.

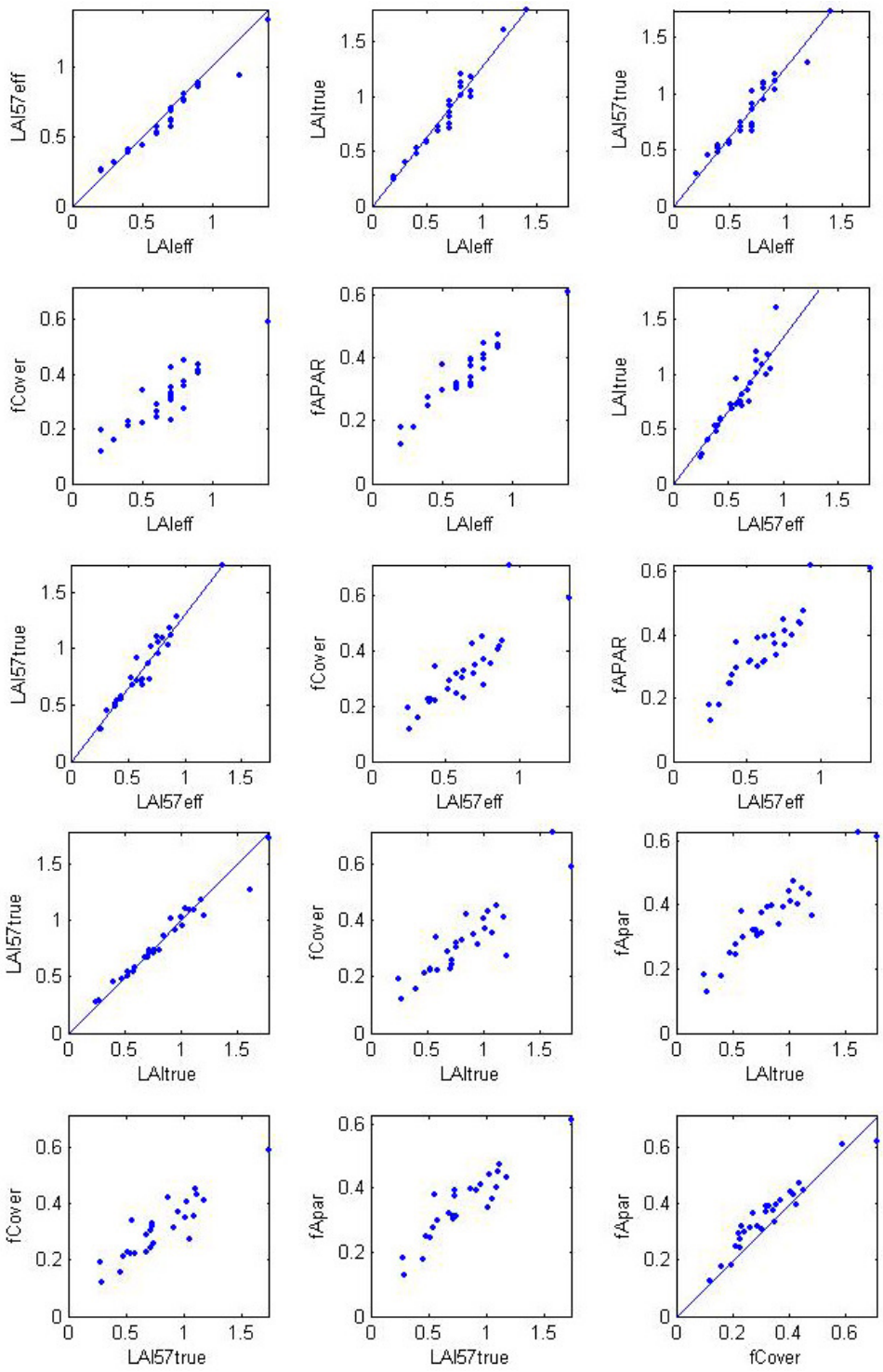


Figure 3. Relationships between the different biophysical variables

2.3. Sampling strategy

2.3.1. Principles

The sampling strategy is defined in the campaign report: http://www.avignon.inra.fr/valeri. The sampling of each ESU is based on twelve elementary photographs.

Figure 4 shows that the 27 ESUs are evenly distributed over the site (3 x 3 km). The processing of the ground data has shown that:

• considering that SPOT geo-location and GPS measurements are associated to errors, we found that processed LAI for ESUs 1C, 2C and 3B did not correspond to the SPOT pixel in terms of reflectance as compared to the knowledge of the land use: they have been shifted by 1 or 2 pixels.

Finally the 27 ESUs have been kept for the computation of the transfer function.

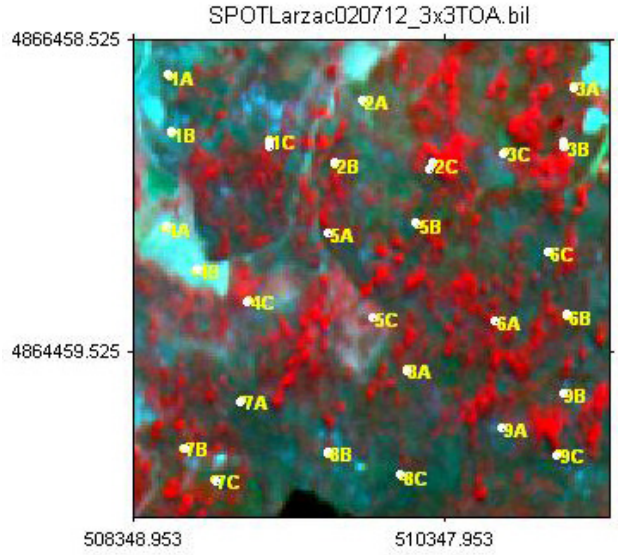


Figure 4. Distribution of the ESUs around the Larzac site.

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 5). As the number of pixels is drastically different for the ESU and whole site (WS=22500 in case of a 3 x 3 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.

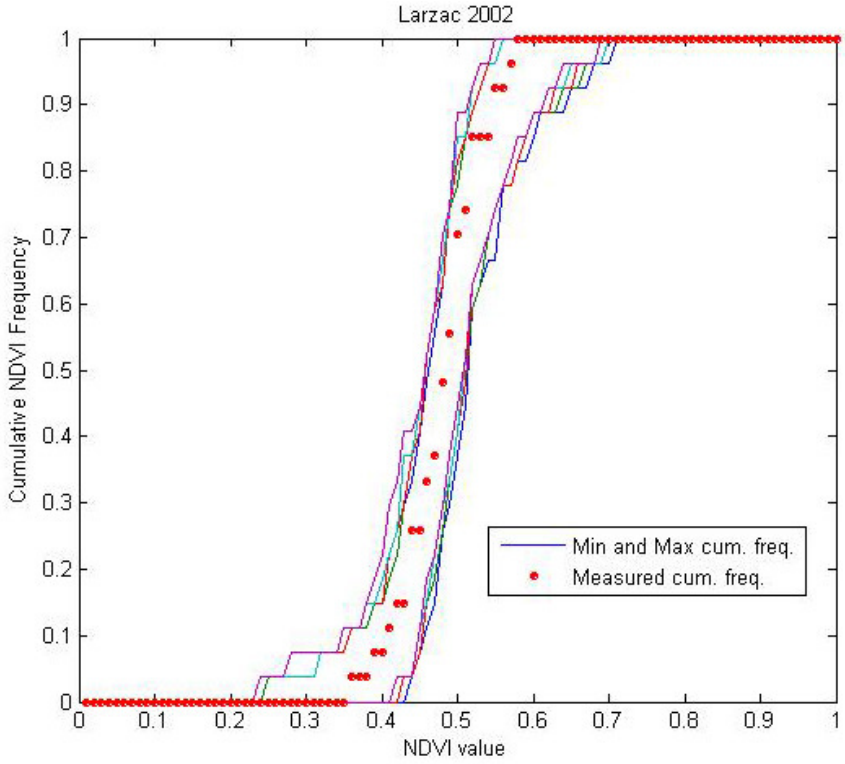


Figure 5. Comparison of the ESU NDVI distribution and the NDVI distribution over the whole image.

Figure 5 shows that the NDVI distribution of the 27 ESUs is very good over the whole site (comprised between the 5 highest and lowest cumulative frequencies). Note that NDVIs lower than 0.35 have not been sampled either although they are present in the image. They may correspond to bare soil, roads or paths. The site is quite 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 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 6). The distribution of the classes on the image and on the ESUs is rather similar. Classes 2 and 5 are under-represented while classes 1, 3 and 4 appear to be over-sampled.

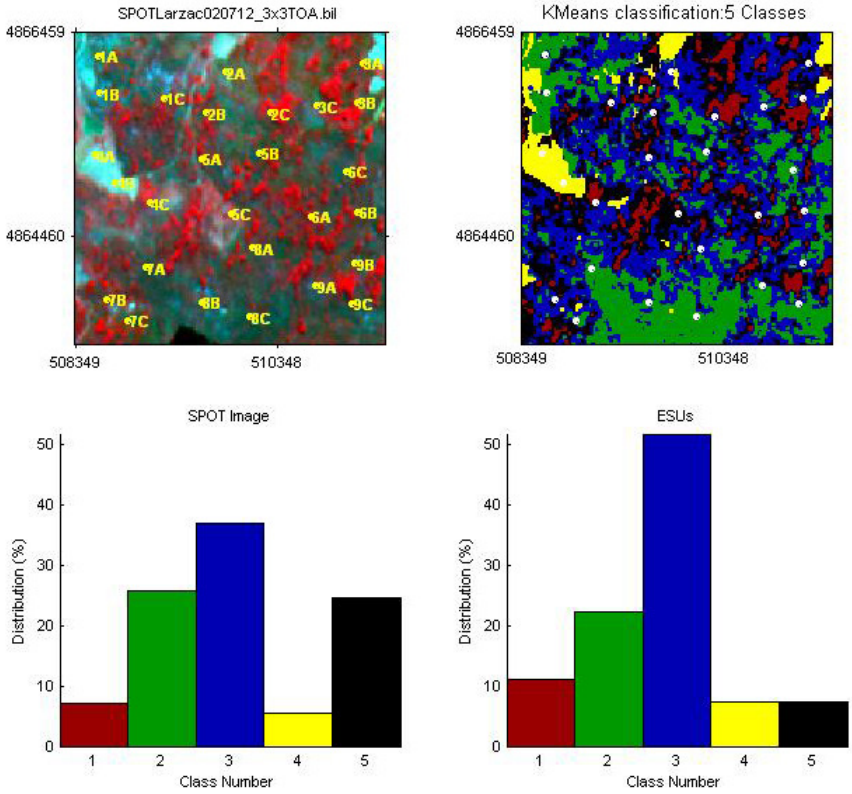


Figure 6. Classification of the SPOT image. Comparison of the class distribution between the SPOT image and sampled ESUs.

Figure 7 shows the different relationships observed between the biophysical variables and the corresponding NDVI on the ESUs, as a function of the SPOT classes determined from non supervised classification.

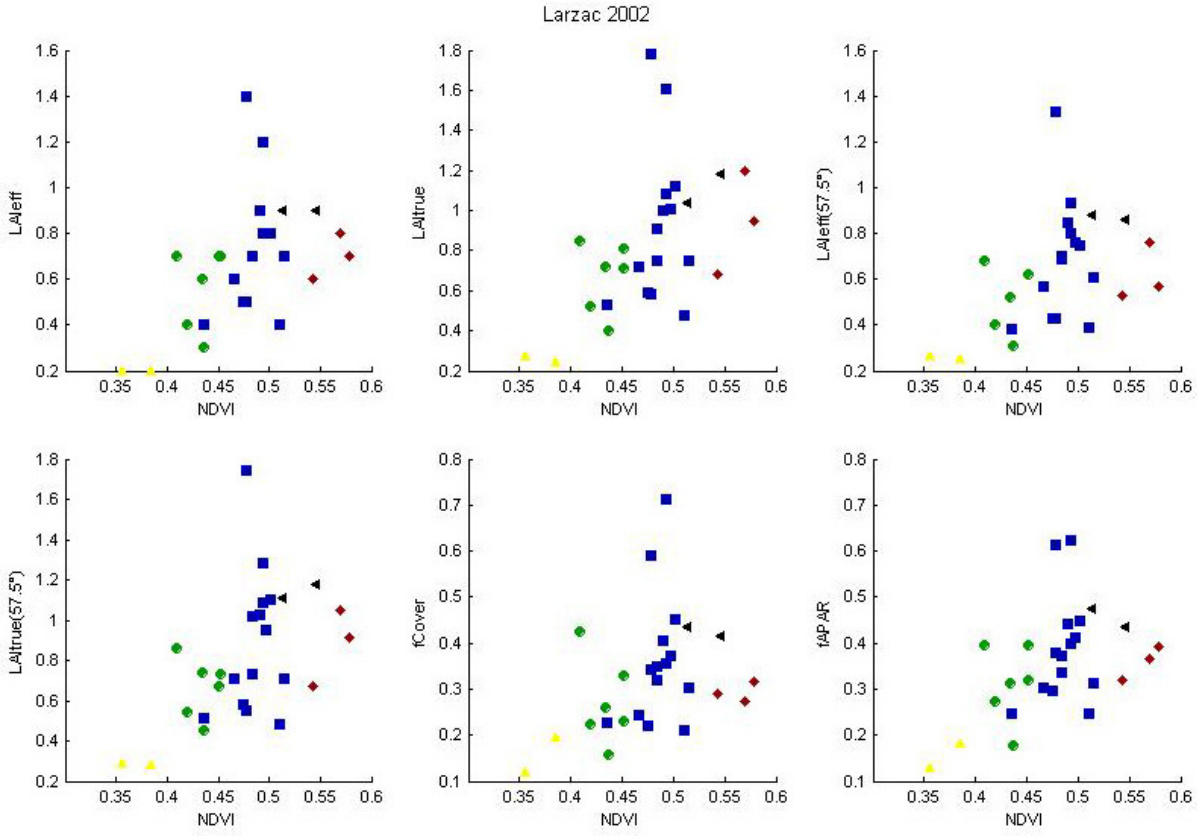


Figure 7. NDVI-Biophysical Variable relationships as a function of SPOT classes

Even if no different behaviour between the classes can be observed, two ESUs (1A and 5B in blue class) differ from the others: the biophysical variable values are relatively high while NDVIs are rather low. Note that the biophysical variable values are generally low. The Larzac site in 2002 is mainly characterized by scattered grassland and shrubs. However, a single transfer function per variable will be generated.

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 8). 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.

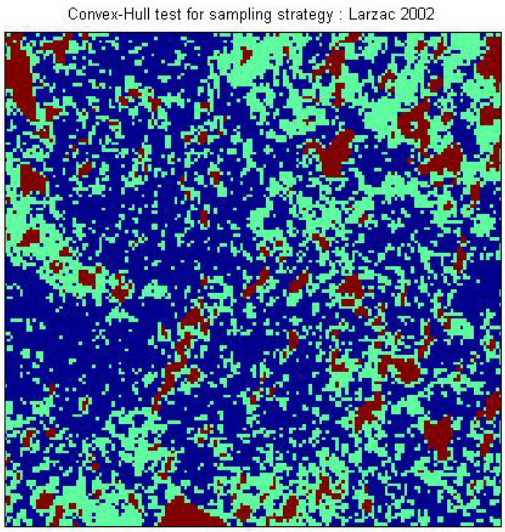


Figure 8. Evaluation of the sampling based on the convex hulls. The map is shown at the bottom: blue and light blue correspond to the pixels belonging to the 'strict' and 'large' convex hulls and red to the pixels for which the transfer function is extrapolating.

This map shows that the representativeness of the ESUs is very good, even if pixels are outside the two convex-hulls. They mainly correspond to wooded surfaces or bare soil.

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

3.1. The transfer function considered

For each class determined in §2.3, the following transfer function was 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 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 regression is tested using either the reflectance or the logarithm of the reflectance for any band combination as well as the simple ratio or NDVI. As the method has poor extrapolation capacities, a flag image, based on the convex hulls is computing over reflectances.

3.2. Results

3.2.1. Choice of the method

For the 5 classes, a unique transfer function is computed. 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)): for LAIeff, LAItrue, LAI57eff, LAI57true, fCover and fAPAR, the results using the reflectance are selected. The regression made on the logarithm of the reflectance sometimes provides better results but it overestimates the biophysical variable values. Therefore the results using the reflectance and the logarithm of the reflectance are not consistent. Depending on the biophysical variable, the choice of the method proves to be difficult because the results are 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).

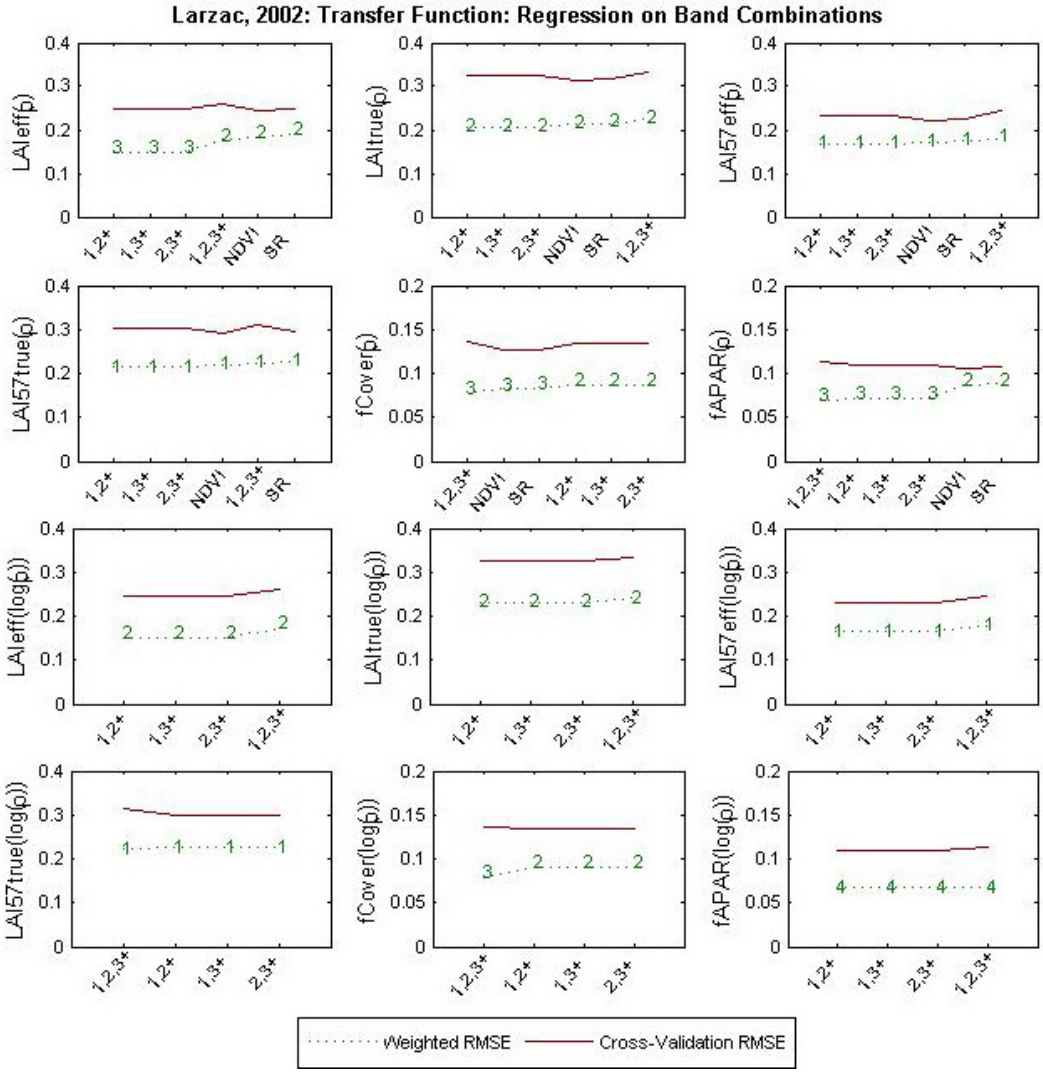


Figure 9. Transfer function: test of multiple regression applied on different band combinations. Band combinations are given in abscissa. The estimated biophysical variable is given in ordinate. Top graphs correspond to regression made on reflectance (ρ): the weighted root mean square error (RMSE) is 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 XS2, XS3, RN (Figure 10 and Figure 11) combination on reflectance was selected since it provides a good compromise between the cross-validation RMSE, the number of weights lower than 0.7 (three) and the weighted root mean square error (the lowest value). The following band combinations provide the same results: [XS1,XS2,RN]; [XS1,XS3,RN].

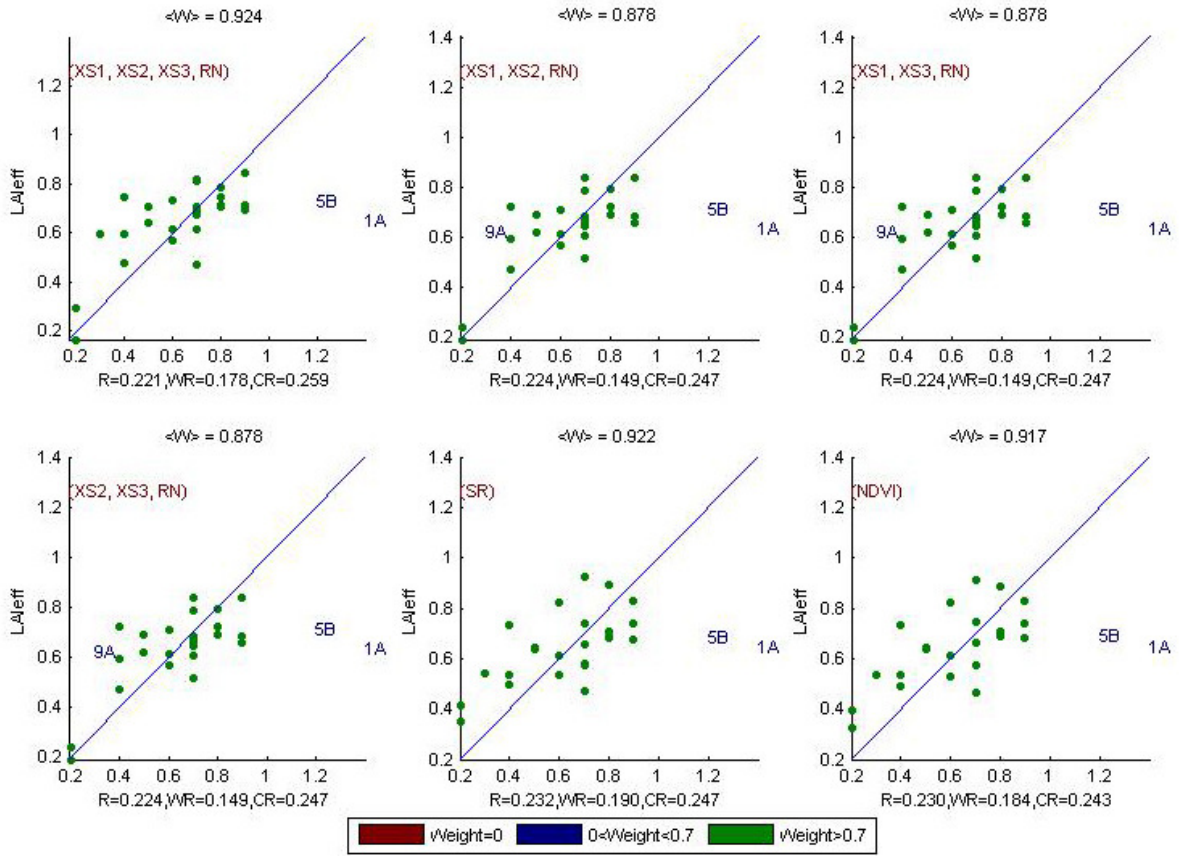


Figure 10. Effective Leaf Area Index: results for regression on reflectance using different band combinations. R is the root mean square error computed between LAIeff and estimated LAIeff. WR is the weighted root mean square error and CR is the cross validation root mean square error.

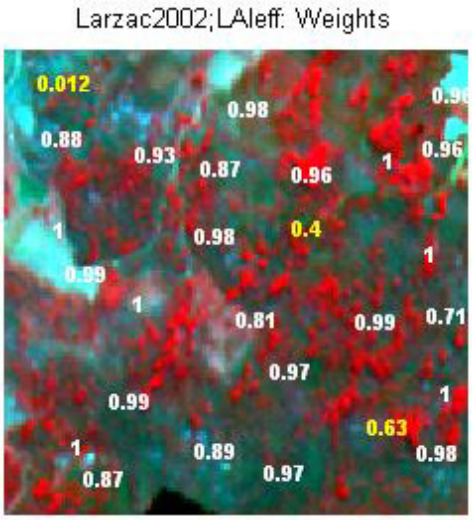


Figure 11. Weights associated to each ESU for the determination of LAIeff transfer function.

For the LAItrue, the XS2, XS3, RN (Figure 12 and Figure 13) combination on reflectance was selected since it provides a good compromise between the cross-validation RMSE, the number of weights lower than 0.7 (two) and the weighted root mean square error (the lowest value). The following band combinations provide the same results: [XS1,XS2,RN]; [XS1,XS3,RN].

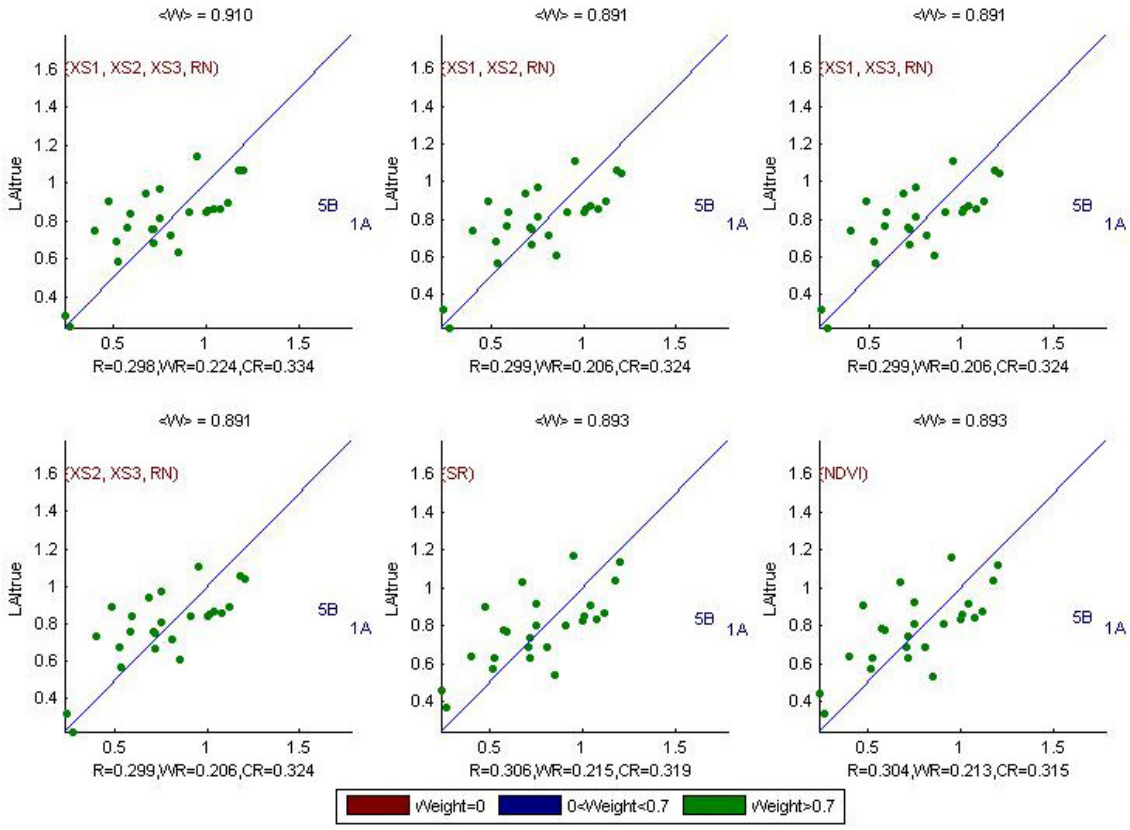


Figure 12. True Leaf Area Index: results for regression on reflectance using different band combinations. R is the root mean square error computed between LAItrue and estimated LAItrue. WR is the weighted root mean square error and CR is the cross validation root mean square error.

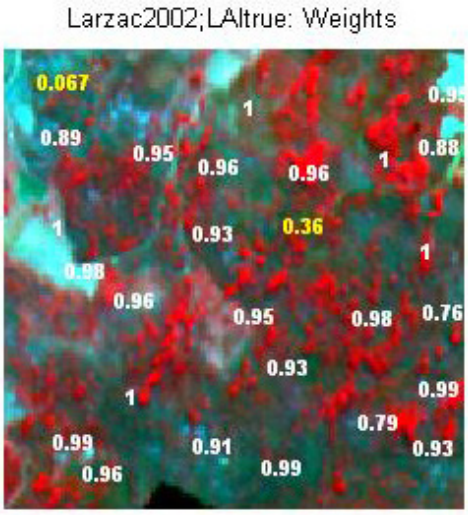


Figure 13. Weights associated to each ESU for the determination of LAItrue transfer function.

For the LAI57eff, the XS2, XS3, RN (Figure 14 and Figure 15) combination on reflectance was selected since it provides a good compromise between the cross-validation RMSE, the number of weights lower than 0.7 (one) and the weighted root mean square error (the lowest value). The following band combinations provide the same results: [XS1,XS2,RN]; [XS1,XS3,RN].

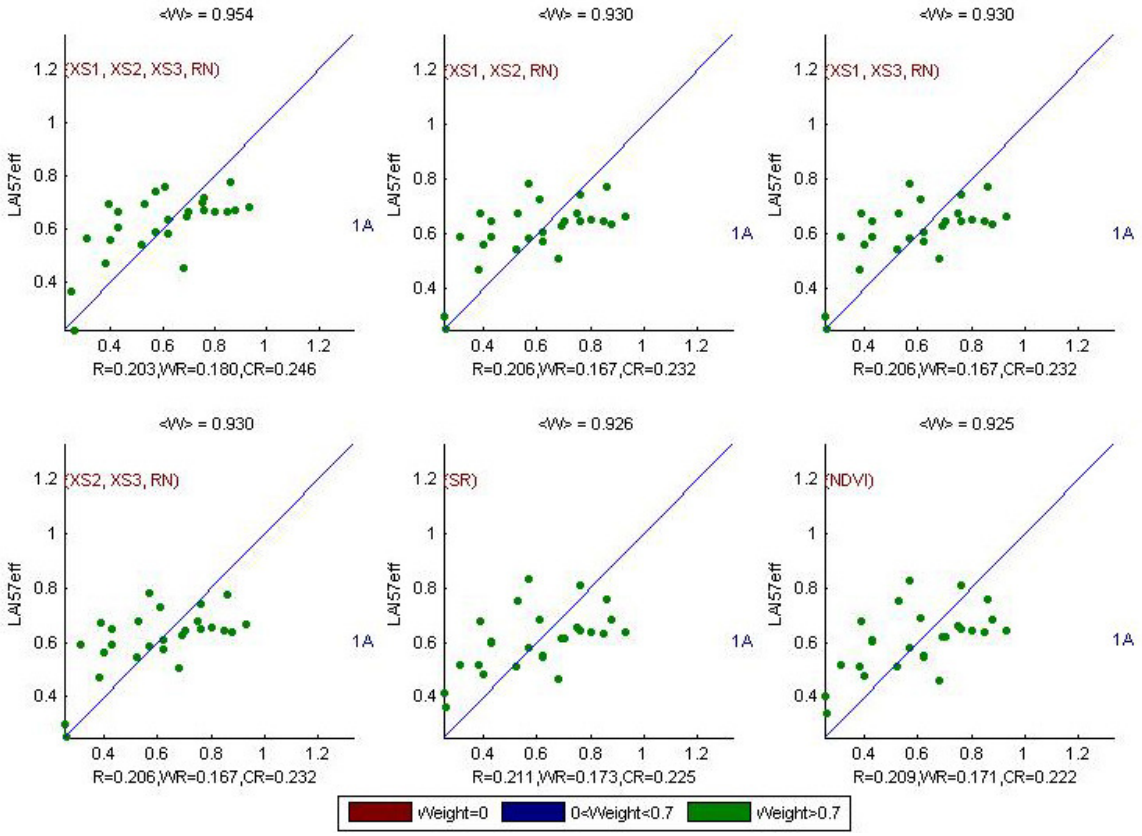


Figure 14. Effective LAI at 57.5°: results for regression on reflectance using different band combinations. R is the root mean square error computed between LAI57eff and estimated LAI57eff. WR is the weighted root mean square error and CR is the cross validation root mean square error.

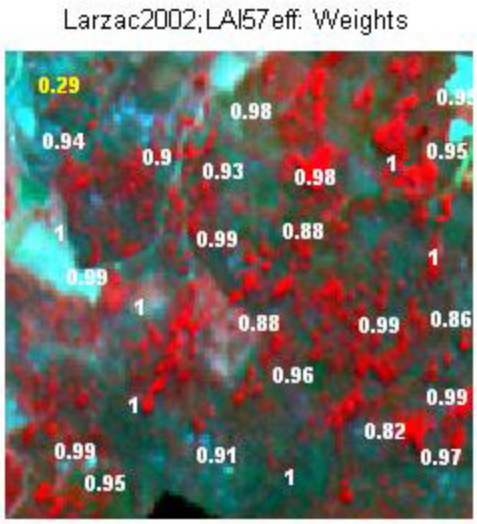


Figure 15. Weights associated to each ESU for the determination of LAI57eff transfer function.

For the LAI57true, the XS2, XS3, RN (Figure 14 and Figure 15) combination on reflectance was selected since it provides a good compromise between the cross-validation RMSE, the number of weights lower than 0.7 (one) and the weighted root mean square error (the lowest value). The following band combinations provide the same results: [XS1,XS2,RN]; [XS1,XS3,RN].

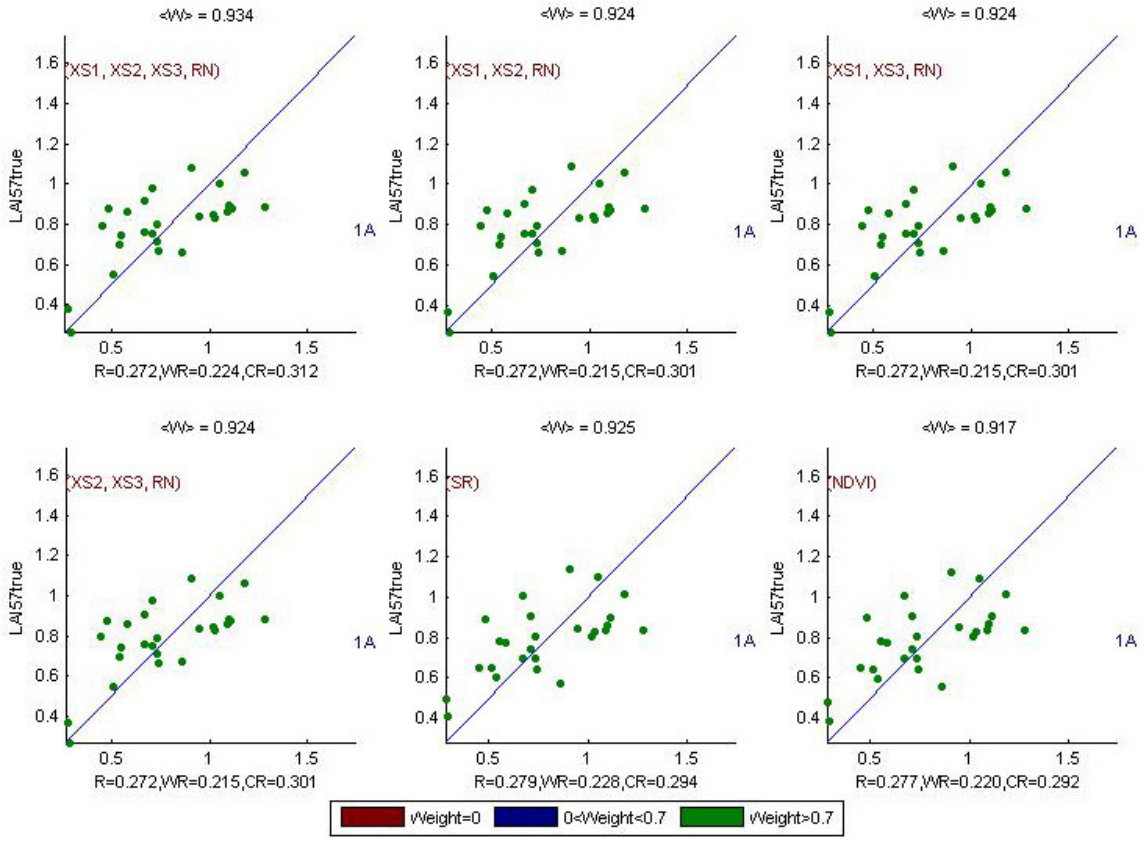


Figure 16. True Leaf Area Index at 57.5°: results for regression on reflectance using different band combinations. R is the root mean square error computed between LAI57true and estimated LAI57true. WR is the weighted root mean square error and CR is the cross validation root mean square error.

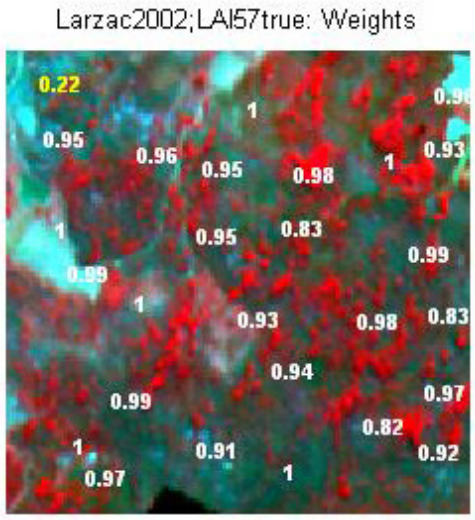


Figure 17. Weights associated to each ESU for the determination of LAI57true transfer function.

For the fCover, the XS1, XS2, XS3, RN (Figure 18 and Figure 19) combination on reflectance was selected since it provides a good compromise between the cross-validation RMSE, the number of weights lower than 0.7 (three) and the weighted root mean square error (the lowest value).

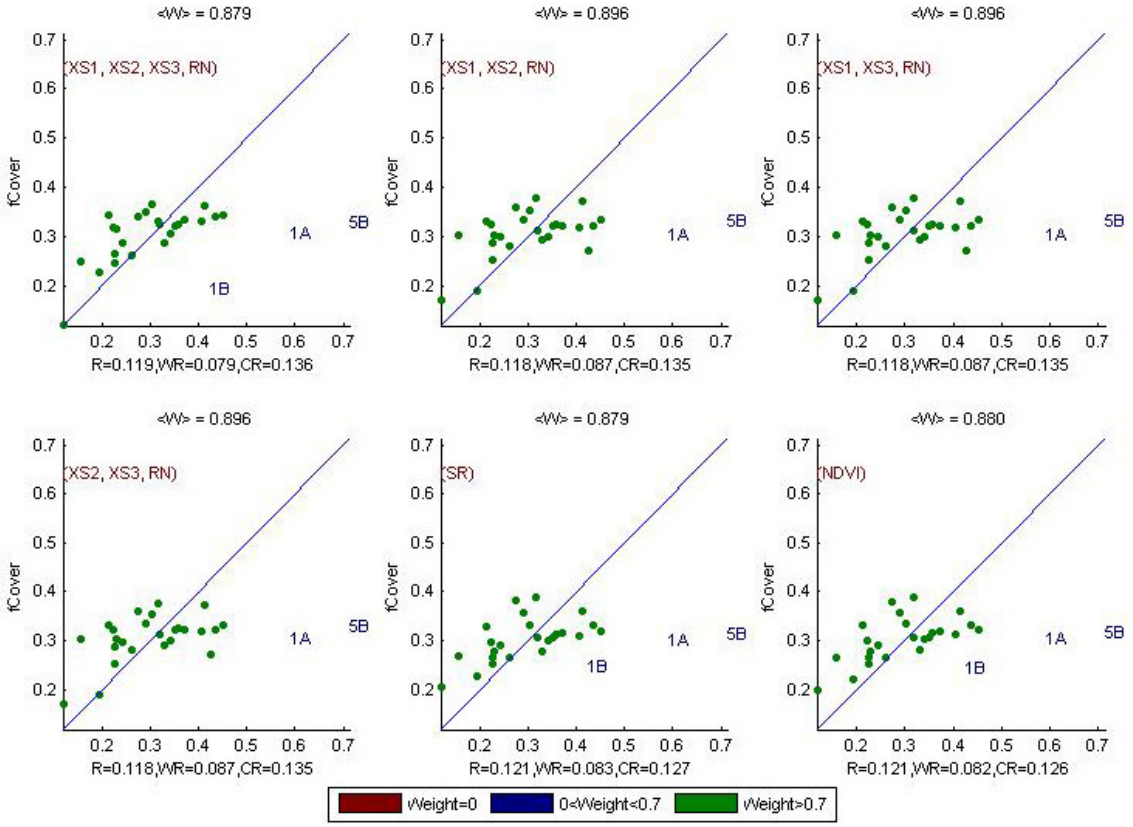


Figure 18. fCover: results for regression on reflectance using different band combinations. R is the root mean square error computed between fCover and estimated fCover. WR is the weighted root mean square error and CR is the cross validation root mean square error.

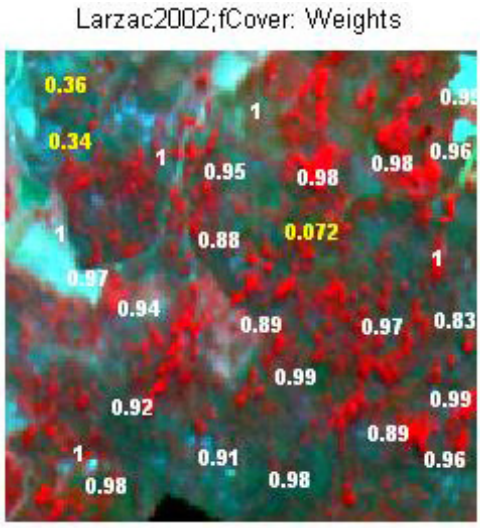


Figure 19. Weights associated to each ESU for the determination of fCover transfer function.

For the fAPAR, the XS2, XS3, RN (Figure 14 and Figure 15) combination on reflectance was selected since it provides a good compromise between the cross-validation RMSE, the number of weights lower than 0.7 (three) and the weighted root mean square error (among the lowest values). The following band combinations provide the same results: [XS1,XS2,RN]; [XS1,XS3,RN].

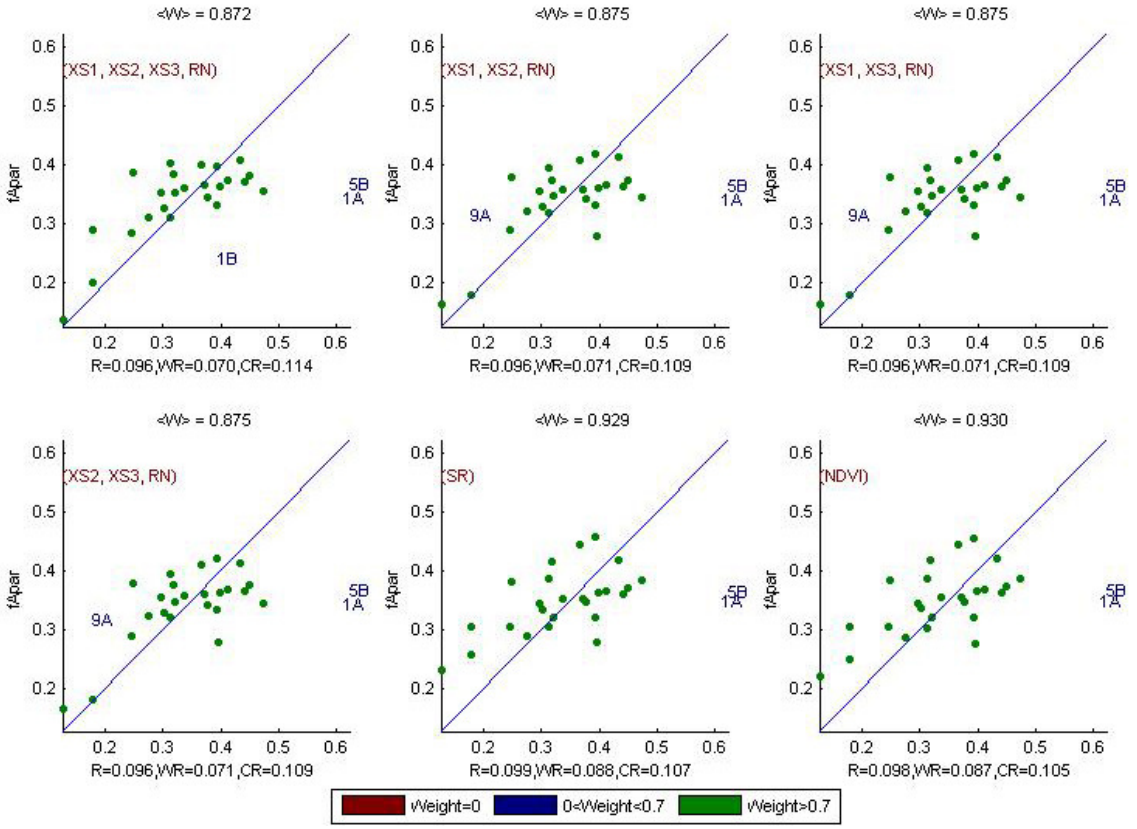


Figure 20. fAPAR: results for regression on reflectance using different band combinations. R is the root mean square error computed between fAPAR and estimated fAPAR. WR is the weighted root mean square error and CR is the cross validation root mean square error.

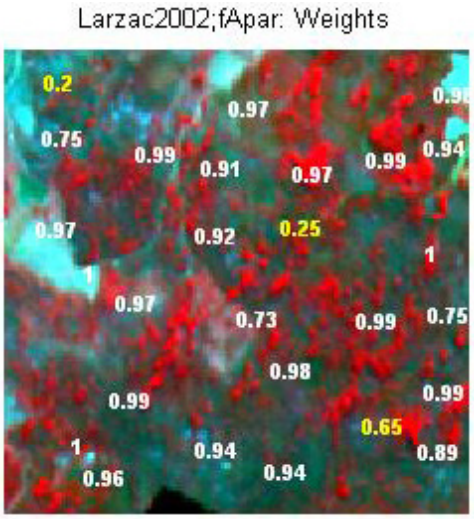


Figure 21. Weights associated to each ESU for the determination of fAPAR transfer function.

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

Variable	Band Combination	RMSE	Weighted RMSE	Cross-valid RMSE
LAIeff	1.2259 + 10.045(XS2) - 20.813(XS3) + 8.974(RN)	0.224	0.149	0.247
LAItrue	1.4147 + 14.48(XS2) - 31.263(XS3) + 25.343(RN)	0.299	0.206	0.324
LAI57eff	1.0164 + 9.0255(XS2) - 17.966(XS3) + 9.7724(RN)	0.206	0.167	0.232
LAI57true	0.51848 + 25.863(XS2) - 34.086(XS3) + 17.735(RN)	0.272	0.215	0.301
fCover	4.8193 - 1.5031(XS1) - 46.799(XS2) - 13.792(XS3) + 150.13(RN)	0.119	0.079	0.136
fAPAR	0.85838 - 0.17584(XS2) - 6.4284(XS3) + 5.0807(RN)	0.096	0.071	0.109

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 Larzac SPOT image extraction

Figure 22 presents the biophysical variable maps obtained with the transfer function described in Table 2. 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.

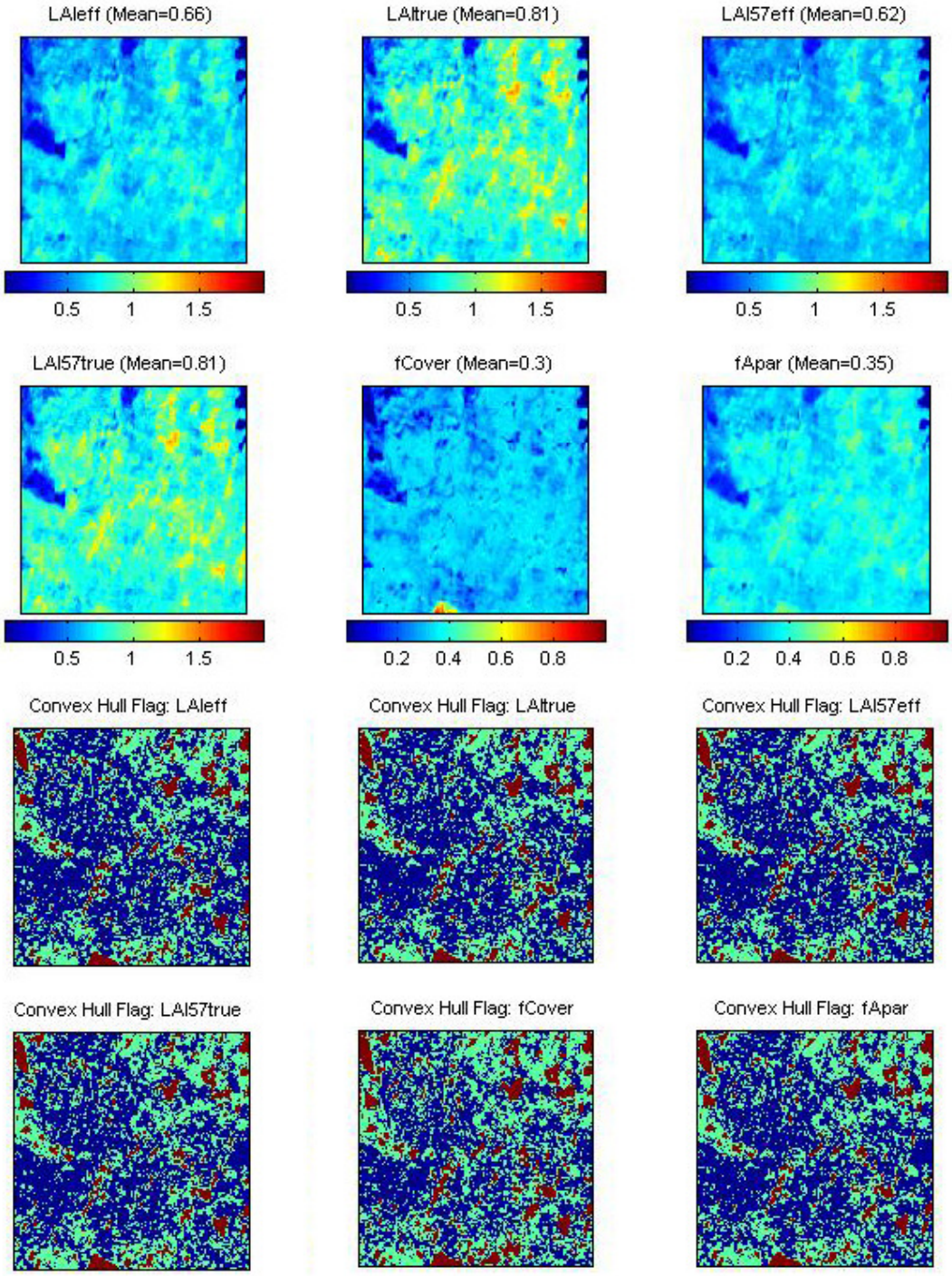


Figure 22. High resolution biophysical variable maps applied on the Larzac site (top). Associated Flags are shown at the bottom: blue and light blue corresponds to the pixels belonging to the 'strict' and 'large' convex hulls and red to the pixels for which the transfer function is extrapolating.

The flag maps are comparable between the different biophysical variables. Note that few pixels are outside the strict convex hull. They correspond mainly to woods or bare soil.

4. Conclusion

The transfer functions are obtained by using 27 ESUs. 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 wooded surfaces and bare soil. Note that the biophysical variable values are low over the whole site. For all the variables, the regression coefficients are computed by relating the variable itself to reflectance.

The biophysical variable maps are available in UTM, 31 North, projection coordinates (Datum: WGS-84) at 20m resolution.

5. Acknowledgements

We want to thank: **Simon Bancarel, Hervé Bohbot, Serge Rambal, Pascal Marty** (CEFE, Montpellier) for the organisation and participation to the campaign.

ANNEX

Field campaign report VALERI 2002
Larzac, France

July 1-3, 2002

CEFE-CNRS team:
Simon BANCAREL, Hervé BOHBOT, Pascal MARTY

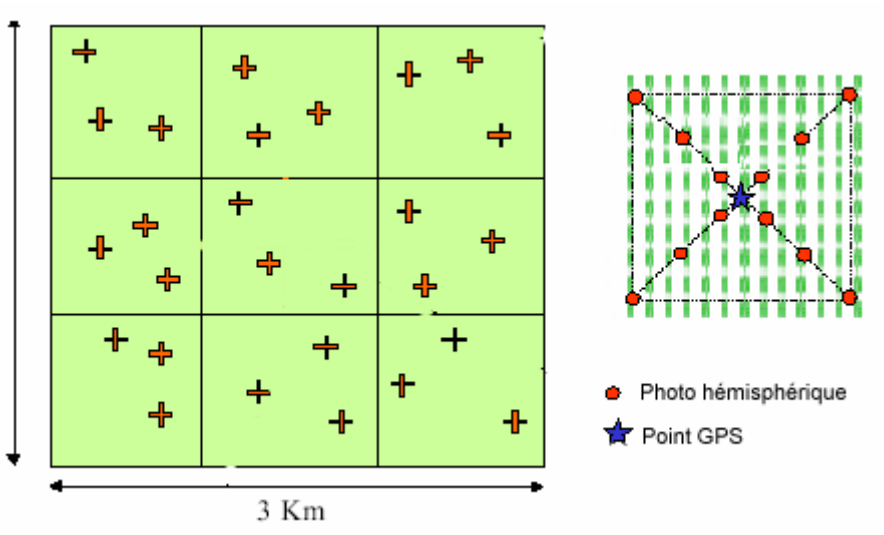
Site location (datum WGS84):

	Lat (DD)	Lon (DD)	Lat (UTM)	Lon (UTM)
UL	N43.9511278	E003.1040630	N4866450	E508350
UR	N43.9510876	E003.1414507	N4866450	E511350
LL	N43.9241170	E003.1040159	N4863450	E508350
LR	N43.9240769	E003.1413867	N4863450	E511350

Site surface is 3 x 3 km. Each square kilometre is referred to with a number, from 1 to 9, southward and eastward.

1	**2**	**3**
4	**5**	**6**
7	**8**	**9**

Sampling method: three spots have been randomly selected in every square *x* and named xA, xB and xC (see distribution next page). Coordinates of these sites have been transferred into 2 GPS receivers. Observers head towards the chosen points and take there a new position which is the centre of the cross that hemispherical photographs takes will design (see figure below).



Sampling sites (in lat/lon, DD):

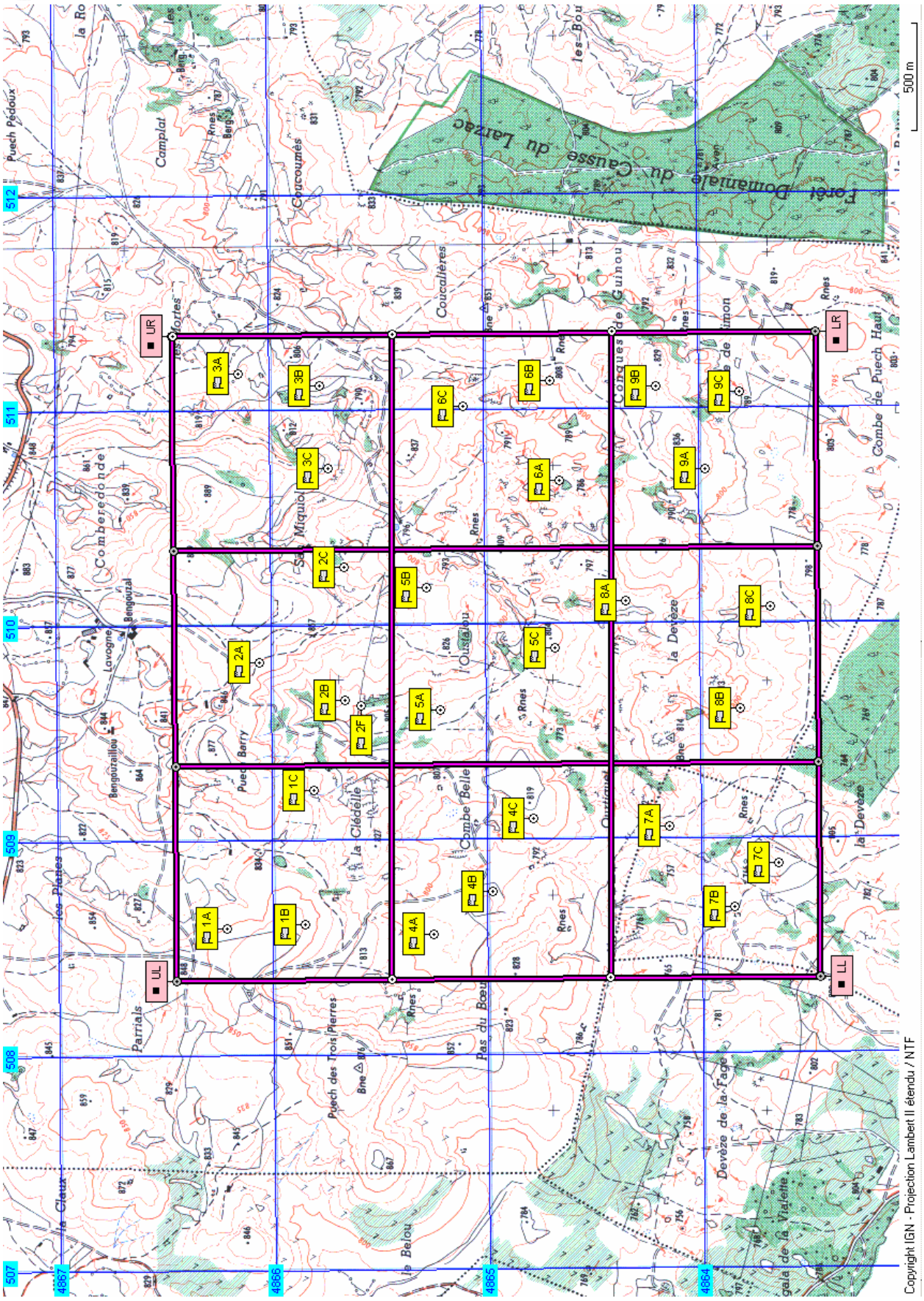
ID	LATITUDE	LONGITUDE	ALT
1A	N43.9489981	E003.1070535	00842
1B	N43.9457315	E003.1072846	00842
1C	N43.9452960	E003.1150904	00825
2A	N43.9475478	E003.1225013	00863
2B	N43.9439646	E003.1202388	00813
2C	N43.9439630	E003.1279853	00808
3A	N43.9483285	E003.1392375	00843
3B	N43.9449092	E003.1385383	00820
3C	N43.9445968	E003.1337266	00861
4A	N43.9403189	E003.1066572	00780
4B	N43.9378633	E003.1091321	00776
4C	N43.9361000	E003.1133226	00805
5A	N43.9399506	E003.1196303	00774
5B	N43.9404868	E003.1267412	00792
5C	N43.9351516	E003.1231889	00788
6A	N43.9348964	E003.1328936	00803
6B	N43.9352555	E003.1386806	00787
6C	N43.9388786	E003.1372550	00796
7A	N43.9304045	E003.1127815	00773
7B	N43.9276864	E003.1081244	00761
7C	N43.9258247	E003.1106197	00768
8A	N43.9321514	E003.1258800	00771
8B	N43.9273824	E003.1196056	00788
8C	N43.9260825	E003.1254748	00765
9A	N43.9287850	E003.1334875	00807
9B	N43.9307891	E003.1383129	00827
9C	N43.9272841	E003.1379271	00789

For all sites, hemispherical photographs were taken above vegetation cover. File names created by the camera were suffixed with the site name (ex: Dscn0001_9B.jpg). Stills representing the sampling sites (taken from one of the crosses extremities) were also taken (except for 9B) and named according the same scheme: Dscn0020_6A_site.jpg. In 6A series, one photo has been taken erroneously below canopy, whilst the others were taken from above. An extra sampling site (a group of trees) has been chosen to have an example of below canopy photographs (unfortunately, 3 of them failed).

2F	N43.9432992	E003.1199435	00805

SPOT Programmation:

A SPOT XS satellite image was acquired the 12th Of July (2 046 261 020712103150 1 X), level 1B. Despite an important acquisition angle (+26°), the relative flatness of the study zone will probably allow a good geometrical correction, which will be performed by the CEFE. The area of interest is cloud-free and presents a good radiometrical quality.


UR
LR
UL
LL
1A
1B
1C
2A
2B
2C
2F
3A
3B
3C
4A
4B
4C
5A
5B
5C
6A
6B
6C
7A
7B
7C
8A
8B
8C
9A
9B
9C
500 m
Copyright IGN - Projection Lambert II étendu / NTF

Courageous scientists…

… facing wild beasts hordes.