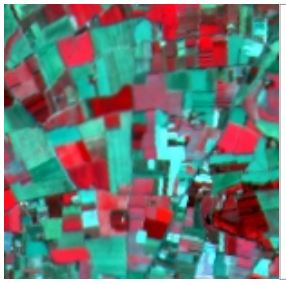
VALERI PROJECT
Website: http://www.avignon.inra.fr/valeri
CAMPAIGN 2002

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



VALERI 2002

ALPILLES site
(crops)

Philippe Rossello, Marie Weiss, Frédéric Baret

October 2005

CONTENTS

1. Introduction

This report describes the production of the high resolution, level 1, biophysical variable maps for the Alpilles site (Table 1 gives the coordinates) in 2002 (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 (hemispherical images). For each Elementary Sampling Unit (ESU), the hemispherical images were processed using the CAN-EYE software (Version 3.6) 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 crops: sunflowers, maize, tomatoes, alfalfa… The size of the fields which are often narrow fluctuates between 2 and 8 ha. The site is quite flat. It is at about 10-20 m altitude (for more information, see 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	4853753	636412	43.82431393	4.69611064
Lower right corner	4850753	639412	43.79657142	4.73286477
Center	4852253	637912	43.810472	4.714694

Table 1. Description of the site coordinates.

The ground measurements were carried out from 22/07/2002 to 23/07/2002, while the high spatial resolution image (SPOT4, HRV1, resolution: 20 m) was acquired on 20/07/2002. The characteristics of the SPOT image are specified in the campaign report.

2. Available data

2.1. SPOT Image

The SPOT image was acquired the 20th July 2002 by HRV1 on SPOT4. It was geo-located by SPOTimage (SPOTView basic). The projection is UTM 31 North, WGS-84 (please, refer to the campaign report for more details: http://www.avignon.inra.fr/valeri). The atmospheric correction[2] was performed by INRA CSE (Abadie, 2002).

[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.

[2] Aerosol optical thickness: AOT_550 (nm) = 0.347 (Avignon AERONET site); water vapor content (gcm-2): 3.032 (Avignon AERONET site); ozone content: 0.330 atm.cm (TOMS observations); air Pressure: 1013 hPa.

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.

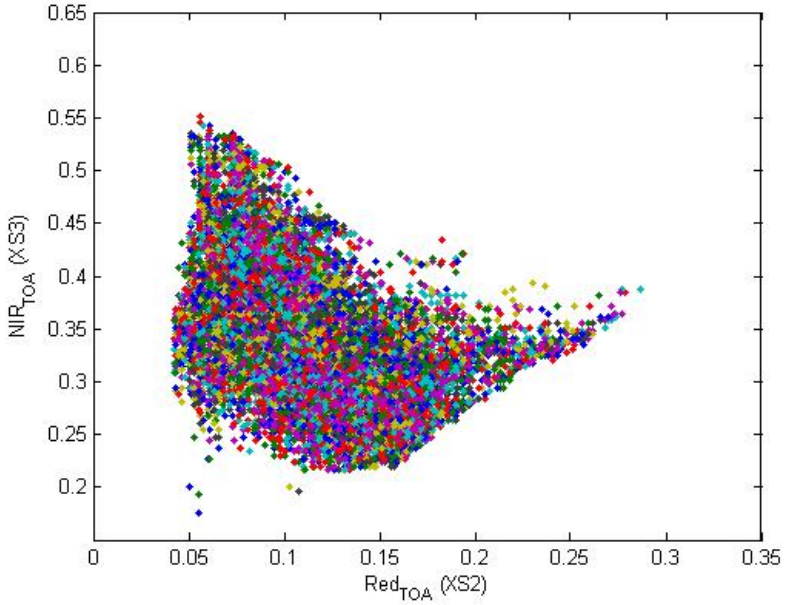


Figure 1. Red/NIR relationship on the SPOT image for Alpilles, 2002.

2.2. Hemispherical images

The hemispherical images were processed using the CAN-EYE software (Version 3.6) to derive the biophysical variables. Figure 2 and Figure 3 show the distribution of the several variables over the sampled ESUs. As there was understorey on three ESUs (orchards), hemispherical images were acquired from above the understorey and from below the canopy (trees). The two sets of acquisition were processed separately to derived 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: fCover(above)*fCover(below). This assumes that independency of the gaps inside the understorey and the gaps 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 to 4, while true LAI (LAItrue, LAI57true) varies from 0 to 5. This range shows a quite heterogeneous 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. There are no fAPAR values between 0.3 and 0.5.

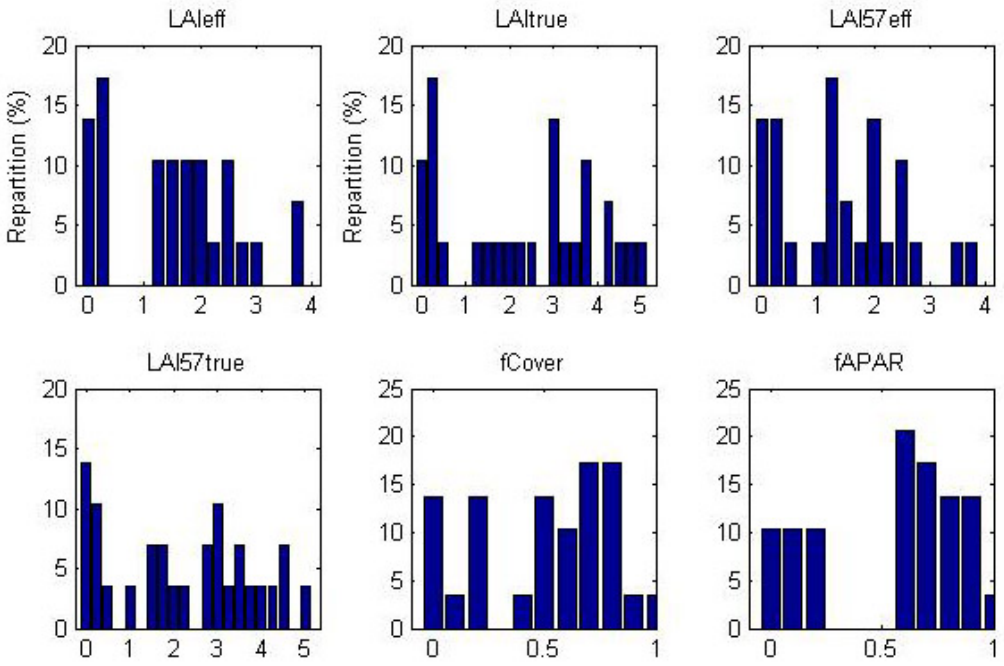


Figure 2. Distribution of the measured biophysical variables over the ESUs.

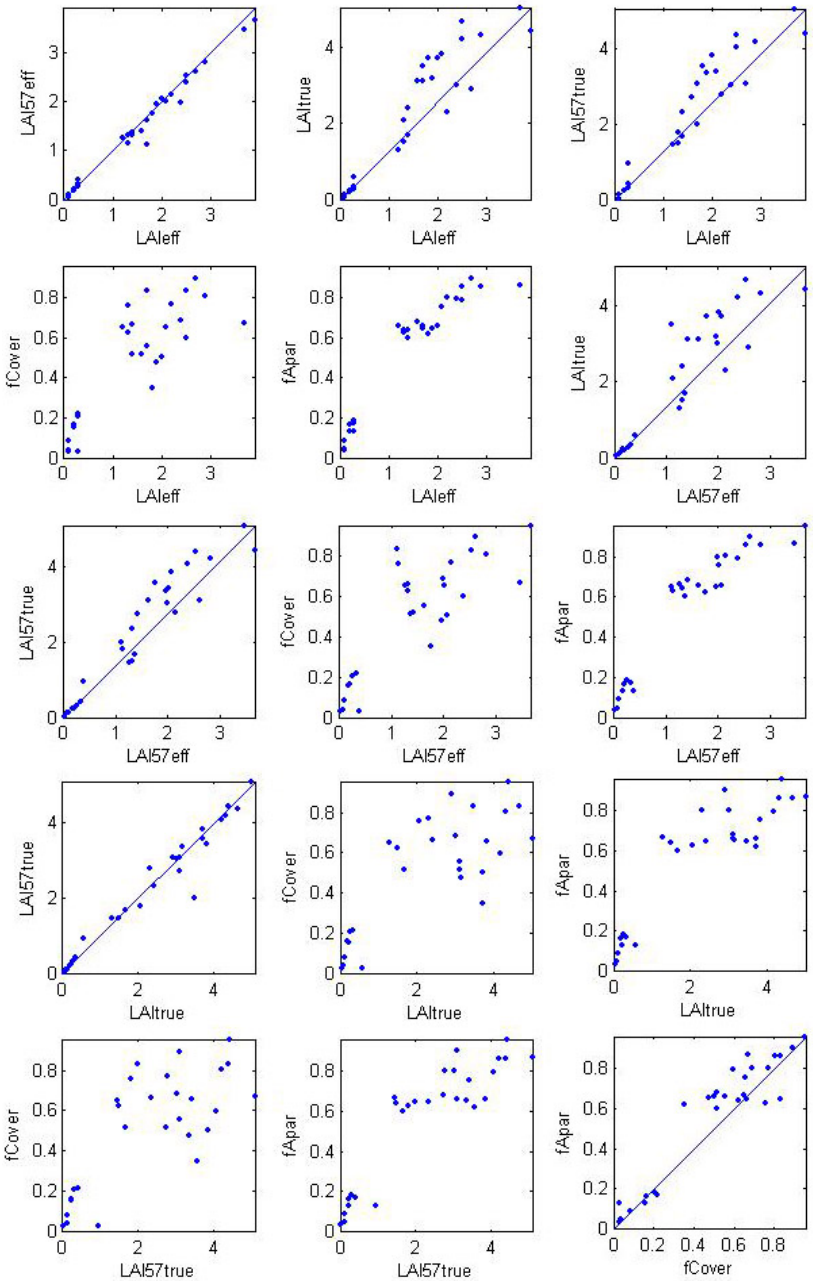


Figure 3. Relationships between the different biophysical variables

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 organized in a cross pattern.

Figure 4 shows that the 30 ESUs[3] are evenly distributed over the site (3 x 3 km). The processing of the ground data has shown that:

- ESU CA09 (in black on Figure 4) was located on a small plot with a strong heterogeneity on the borders. This ESU was eliminated;
- considering that SPOT geo-location and GPS measurements are associated to errors, we found that processed LAI for ESUs CA02, CA07, CA11, CA16 and CB11 did not correspond to the SPOT pixel in terms of reflectance as compared to the knowledge of the land use: according with the people who acquired the data, they have been shifted by 1 or 2 pixels.

[3] 6 sunflowers, 6 wheat, 4 tomatoes, 4 orchards, 4 grassland, 3 alfalfa, 1 maize, 1 salads, 1 fallow.

Finally 29 ESUs have been kept for the computation of the transfer function.

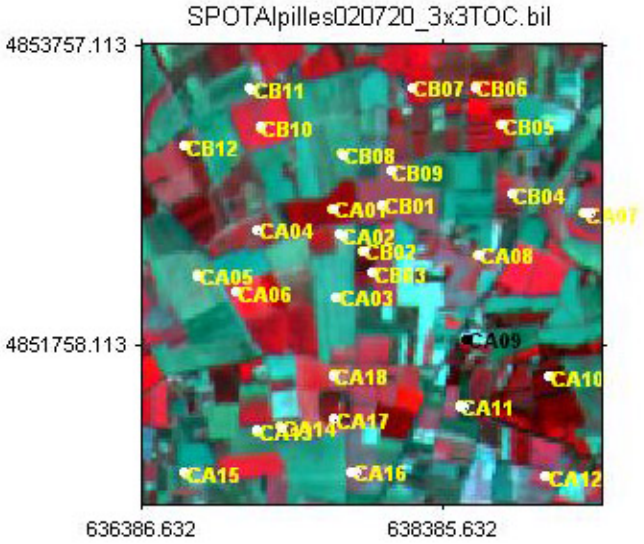


Figure 4. Distribution of the ESUs around the Alpilles site. ESU in black (CA09) was eliminated for the computation of the transfer function.

Figure 5 shows the land cover of the Alpilles site characterized by a mosaic of crops. The land cover map included here is approximative.

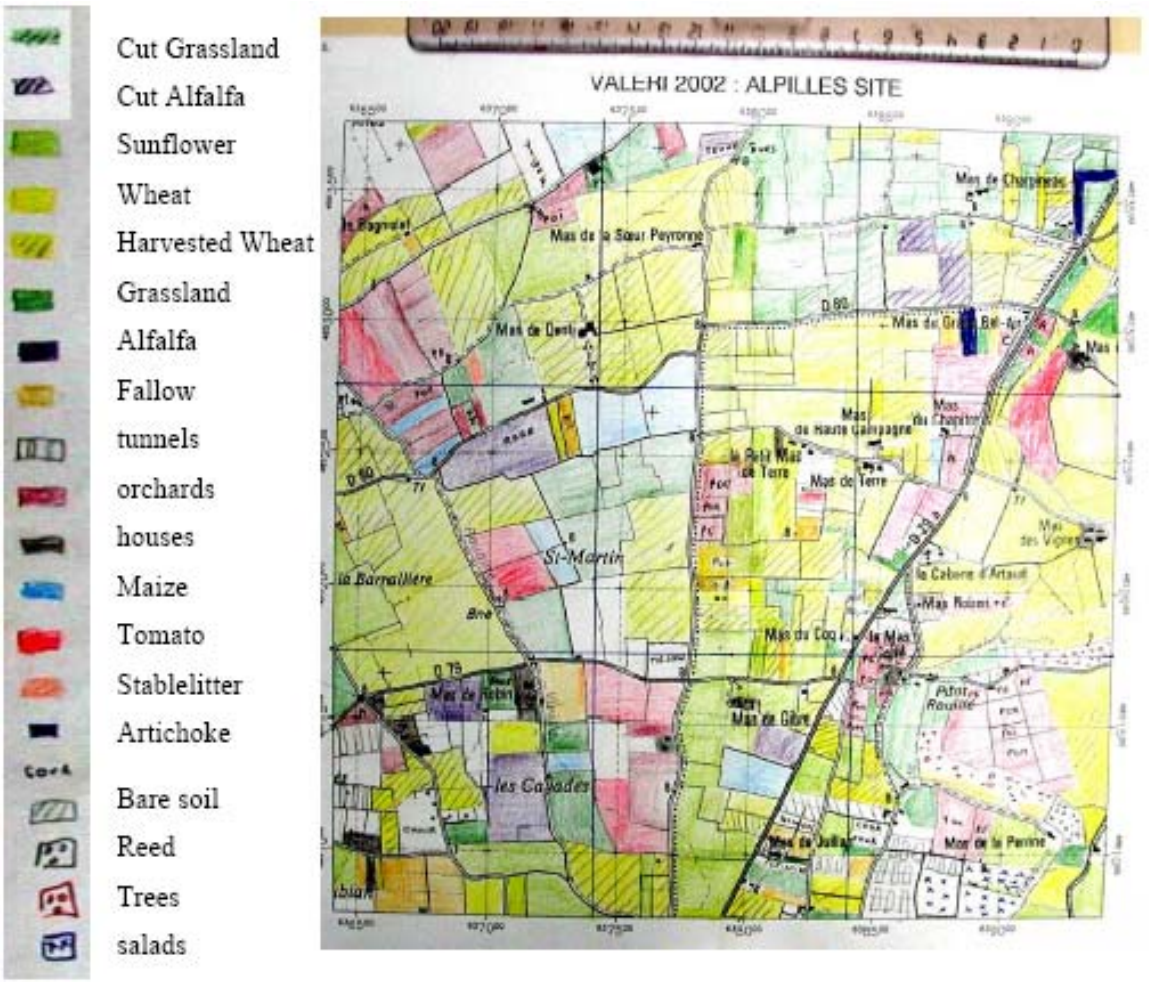


Figure 5. Land cover map of the Alpilles site (2002)

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 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.

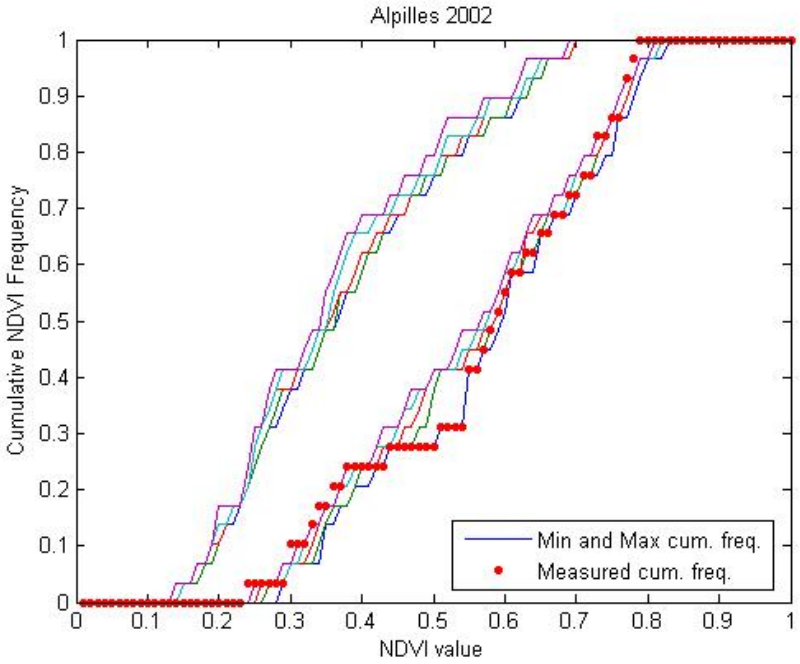


Figure 6. Comparison of the ESU NDVI distribution and the NDVI distribution over the whole image.

Figure 6 shows that the NDVI distribution of the 29 ESUs is quite good over the whole site (comprised between the 5 highest and lowest cumulative frequencies) even if the cumulative frequency curve is often close to the boundaries for high NDVI values. It reaches even the boundaries on several occasions since NDVIs have not been sampled between 0.23 and 0.29, 0.37 and 0.43… Note that NDVIs lower than 0.22 have not been sampled either although they are present in the image. Moreover, the site is quite homogeneous in terms of NDVI since the highest and lowest distributions are close.

2.3.3. 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 7). The distribution of the classes on the image and on the ESUs is rather different. Classes 1, 2 and 3 are under-represented while class 5 appears to be over-sampled. The five classes correspond to:

- class 1: fallow, harvested wheat fields (3 ESUs);
- class 2: tomatoes, alfalfa, grassland (4 ESUs);
- class 3: harvested wheat fields, grassland, alfalfa (5 ESUs);
- class 4: orchards, sunflower, grassland (4 ESUs);

- class 5: sunflower, grassland, alfalfa, orchards, maize, tomatoes, salads (13 ESUs).

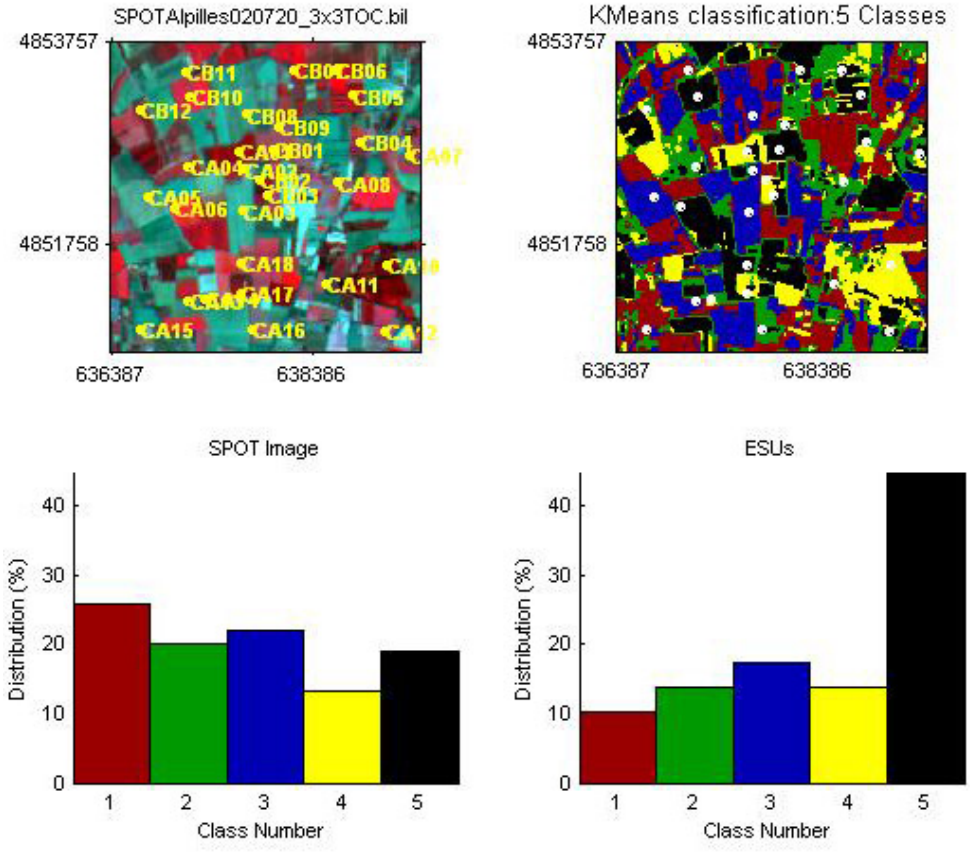


Figure 7. Classification of the SPOT image. Comparison of the class distribution between the SPOT image and sampled ESUs.

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.

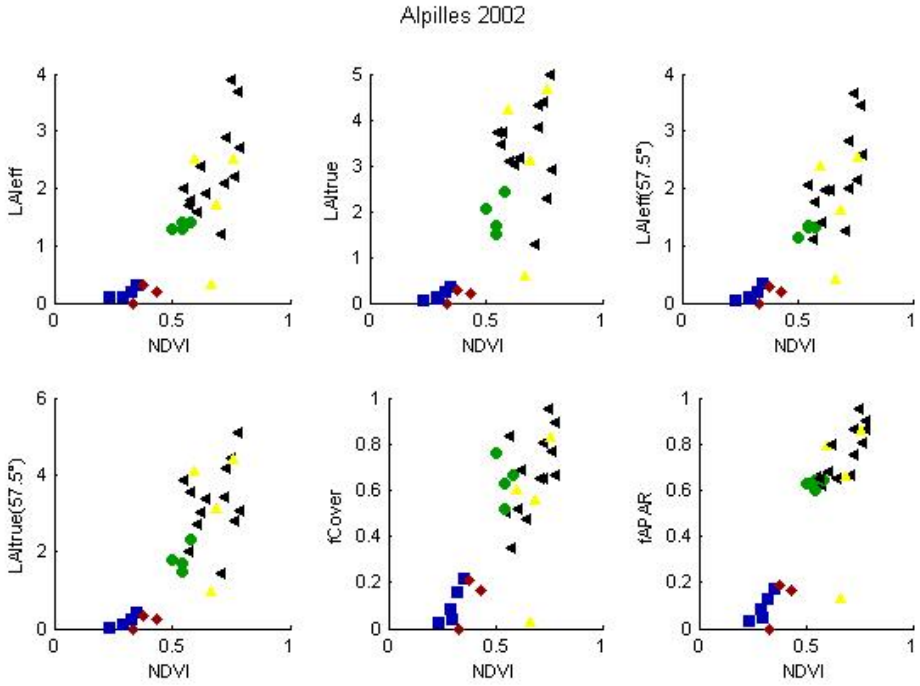


Figure 8. NDVI-Biophysical Variable relationships as a function of SPOT classes

Even if no different behaviour between the classes can be observed, one ESU (CB12) in class 4 (yellow) differs from the others: the NDVI value is high, while biophysical variable values are very low. It is located in a peach tree orchard. Its particular behaviour is probably due to the distance between the rows (four meters) and the lack of hemispherical images acquired from above the understorey. However, this ESU is kept. Therefore, 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.

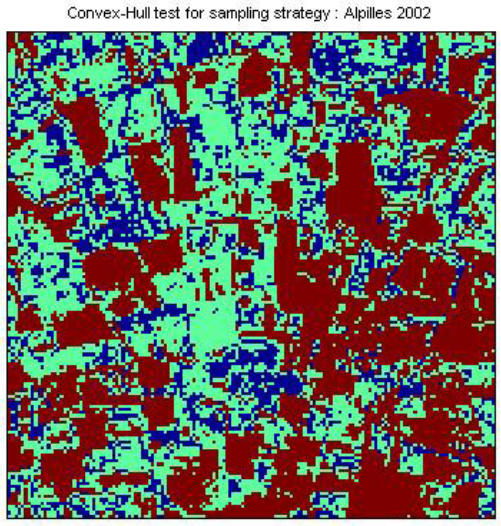


Figure 9. Evaluation of the sampling based on the convex hulls. The map is shown at the bottom: blue and light blue correspond to the pixels belonging to the 'strict' and 'large' convex hulls and red to the pixels for which the transfer function is extrapolating.

This map shows that the pixels outside the two convex-hulls are numerous. They correspond to a heterogeneous land cover (bare soil, wheat fields, tomatoes, salads…), mainly dominated by the bare soil surfaces. This is due to the fact that the distribution of bare soil reflectances is not well represented by the sampling (low NDVI values). This indicates that the sampling could have been better and that more low vegetation surfaces should have been measured. This was not made possible by the diversity of the landscape and therefore, the corresponding range of reflectance values is not taken into account to calculate the convex hulls.

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

3.1. The transfer functions considered

For each class determined in §2.3, two types of transfer functions were tested:

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

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

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

The results are detailed in §3.3. To improve the relation between the biophysical variables, the transfer functions use news band combinations.

3.2. A method to improve the relation between the biophysical variables

The transfer functions are applied over all band combinations (§3.1). The band combination giving the best results is selected to estimate the values of the biophysical variables over the whole site. This method is operational and the results of the multiple robust regression are pertinent, but the dependency between the estimated variables is questionable because of the linear nature of the individual transfer functions. For example, the relation between LAIeff and fAPAR is linear whereas it should a priori be exponential (as shown by the measured variables).

To improve the relation between the biophysical variables, three methods were tested using the reflectance to calculate LAIeff and fAPAR values with the transfer functions (Figure 10):

- method A: XS1, XS2, XS3, XS4. This method was used up till now;
- method B: XS1, XS2, XS3, XS4, Red*NIR;
- method C: XS1, XS2, XS3, XS4 with LAIeff = klog(fAPAR), where k is fitted using the measured values.

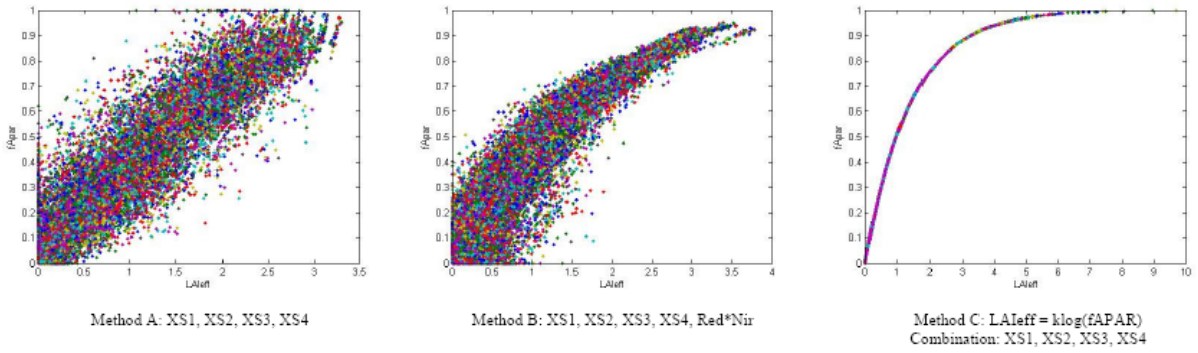


Figure 10. Relations between LAIeff and fAPAR over the whole site using different methods

The results show that the relation between LAIeff and fAPAR is linear with method A and exponential with methods B and C. The addition of the Red*NIR band to the initial combination (XS1, XS2, XS3, XS4) is enough to improve the relation between LAIeff and fAPAR. Method B produces a residual information while the relation resulting from method C is a perfect exponential curve (mathematical connection).

The VALERI project aims at validating biophysical products derived from medium or large swatch sensors. Therefore, the aggregation of the data (at 1 km resolution) resulting from the different methods was carried out to estimate the value of the biophysical variables (Figure 11 and Figure 12).

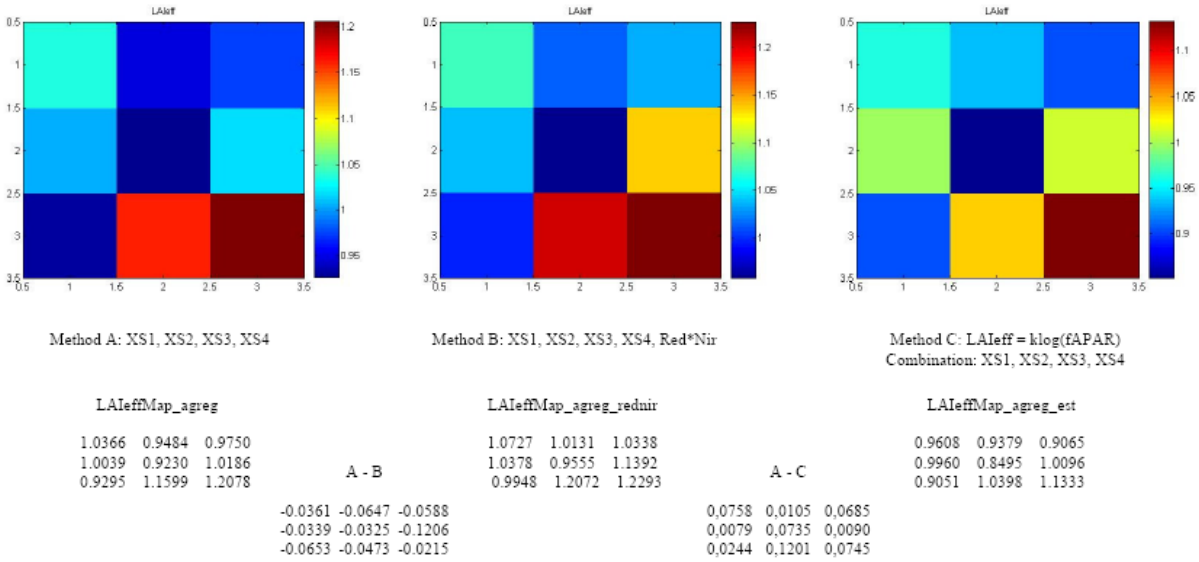


Figure 11. Estimation of the LAIeff value at 1 km resolution (Alpilles site, 2002)

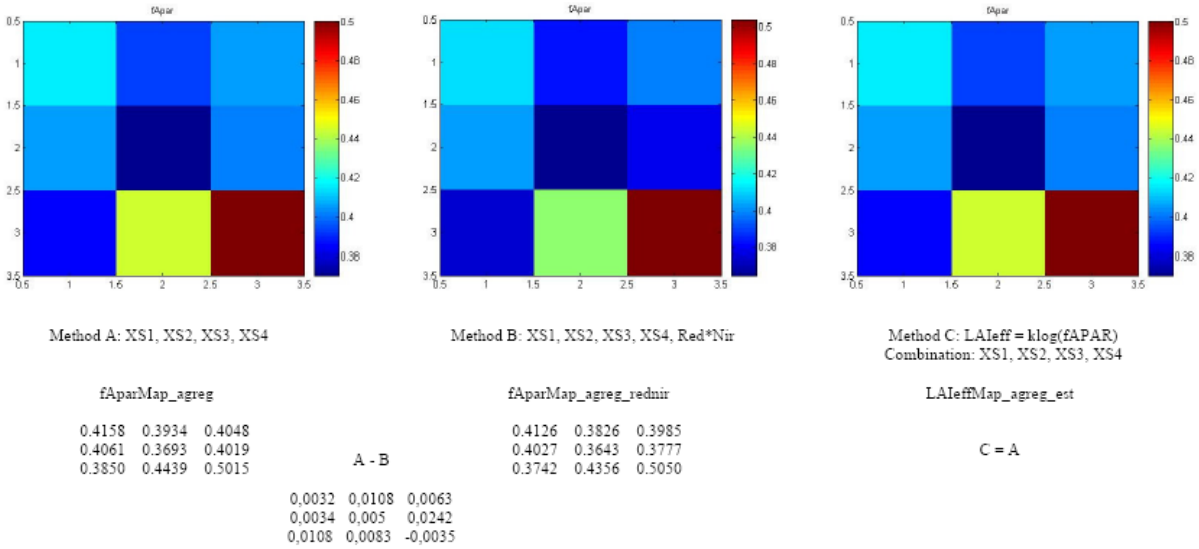


Figure 12. Estimation of the fAPAR value at 1 km resolution (Alpilles site, 2002)

Finally, the results show that while the incidence of the method is very significant at high spatial resolution, it is actually minimal at medium or large resolution. This is important since the previous processes within the framework of the VALERI project[4] is based on method A. Therefore, the previous processes stay pertinent at medium or large resolution, in agreement with the main VALERI objectives.

However, an additional dimension is now added to the all the available band combinations. The multiple robust regression with the interaction of the Red*NIR band (method B) optimizes indeed the transfer functions.

3.3. Results

3.3.1. Choice of the method

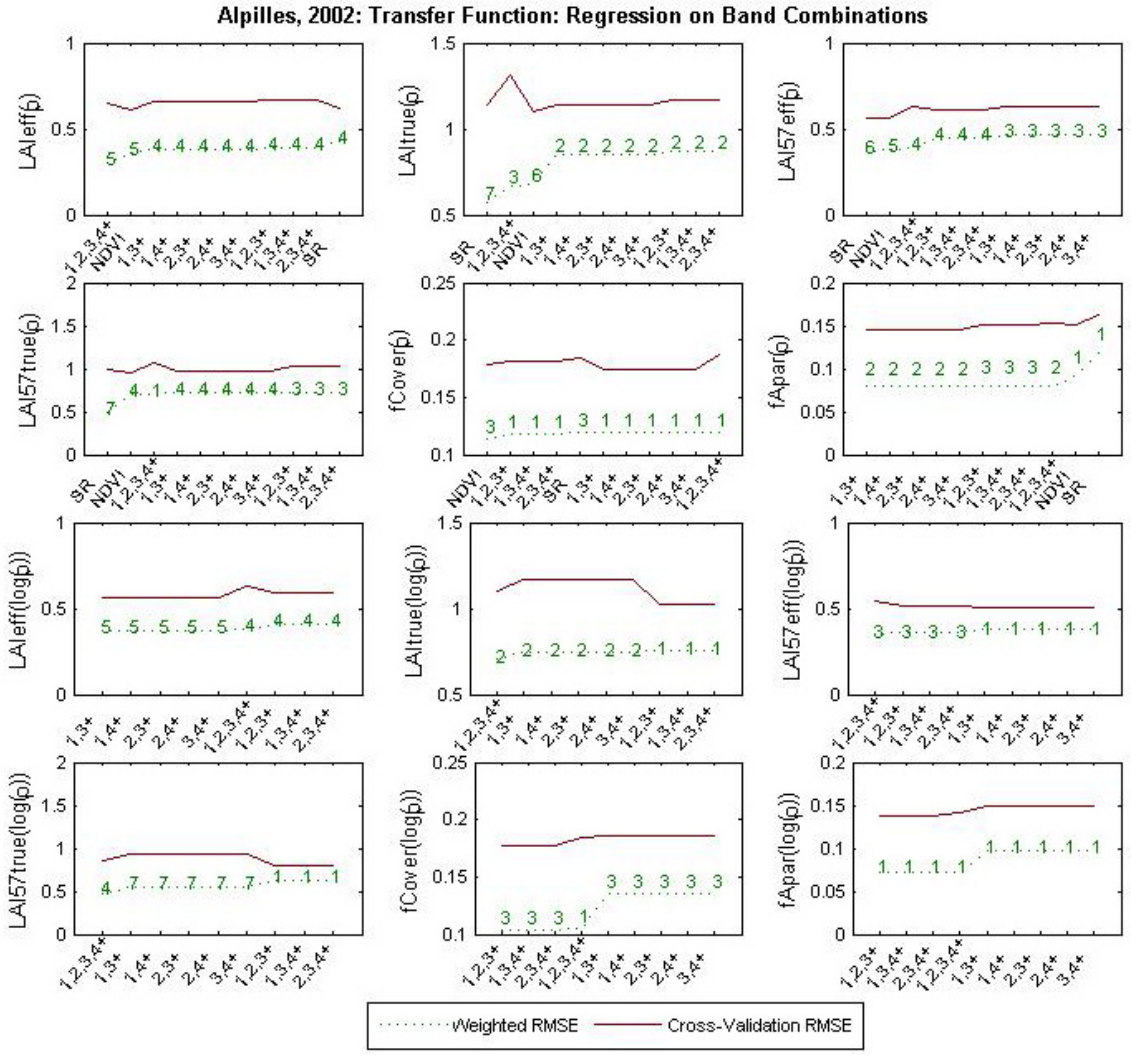
For the 5 classes, a unique transfer function was computed. Figure 13 and Figure 14 show the results obtained for all the possible band combinations using either the reflectance or the logarithm of the reflectance:

- The REG method provides better results in terms of cross-validation RMSE for all the variables and is therefore selected as the transfer function instead of the LUT;

[4] for more information: http://www.avignon.inra.fr/valeri

• For LAIeff, LAItrue, LAI57eff, LAI57true, fCover and fAPAR, the results using the logarithm of the reflectance are the best (Figure 13). However, the transfer function provides many LAI values of the order of -3 on the Alpilles site. Even if these values are put to 0 since they correspond to bare soils, the biophysical variable maps are not pertinent since unrealistic extrapolations were performed due to the use of the logarithm in the transfer function. Note that the estimated fAPAR and fCover values are also negative. The bare soil sampling is in question. Moreover, the addition of the Red*NIR dimension to the band combination is problematic to compute the convex hulls: the transfer function using the logarithm of the reflectance creates coplanar points which do not allow the determination of the 'strict' and 'large' convex hulls (§3.4). Different options were tested but the results are not conclusive. Therefore, the results using the reflectance which are satisfactory were selected.

• Depending on the biophysical variable, the choice of the method proves to be difficult because the results are close.



*+ : the Red*Nir combination is added to all the band combinations (except for NDVI and SR). Please, read §3.2.*

Figure 13. 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.

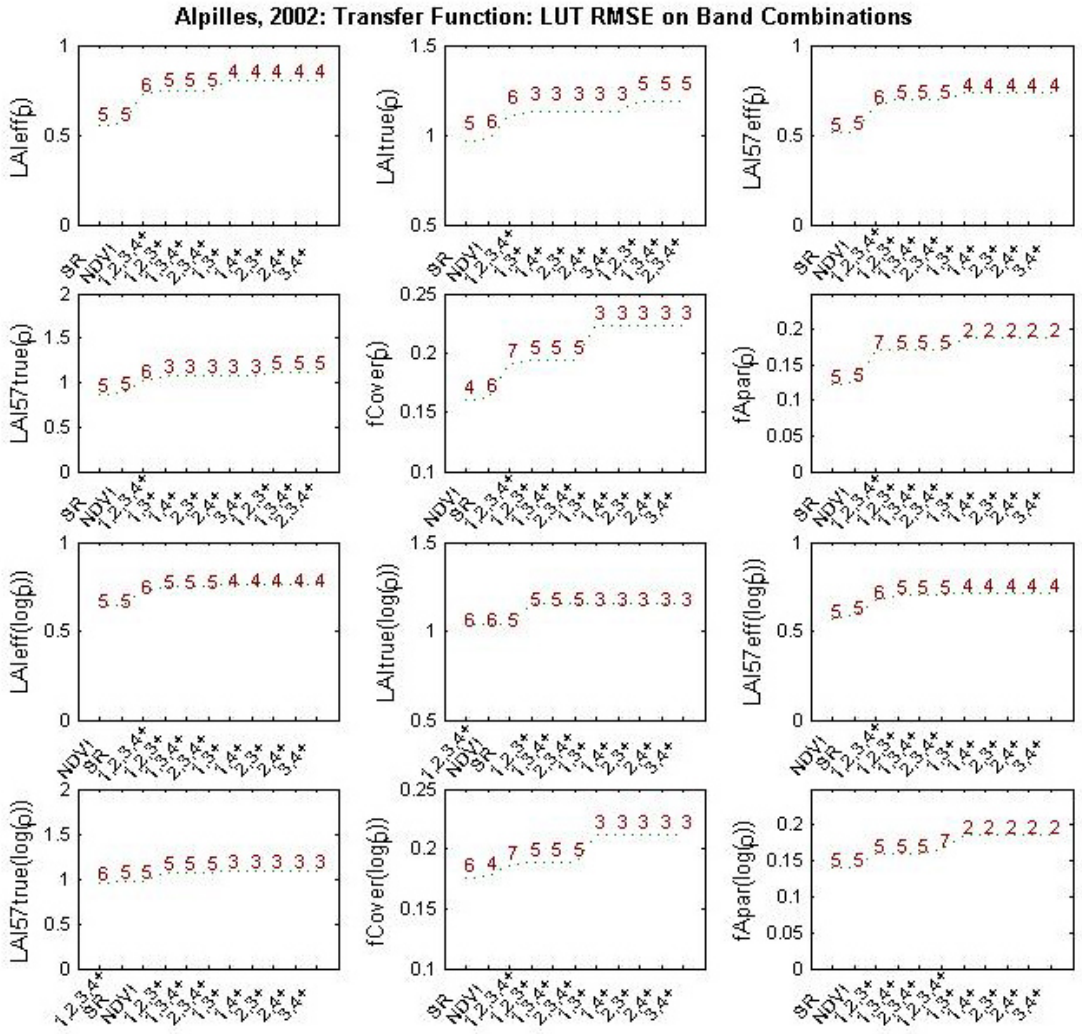


Figure 14. Transfer function: test of LUT applied on different band combinations. Band combinations are given in abscissa. The estimated biophysical variable is given in ordinate. Top graphs correspond to regression made on reflectance (ρ): the root mean square error is presented in green. The numbers indicate the number of elements selected in the LUT to compute the resulting biophysical variables. Bottom graphs correspond to LUT using the logarithm of the reflectance.

3.3.2. Choice of the method for the future processing of the VALERI sites

Note that, up to now, for the main part of the sites that have been processed, the transfer function was based on multiple robust regression (REG) on reflectance (Table 2). The LUT method is never selected since the number of ESUs is generally too small and the REG method on logarithm of the reflectance is not much used. In fact, the results using the logarithm of the reflectance are often similar to those using the reflectance.

Table 2 indicates the selected methods for the different VALERI sites that have been processed:

				methods			
site	country	date	landcover	AVE*	REG** (ρ=reflectance ; log=logarithm)	LUT ***	comments
Aek Loba	Indonesia	05/2001	plam tree plantation	-	ρ: LAI	-	LAI map retrieved using the linear NIR-LAI relationship
Alpilles	France	03/2001	cropland	-	ρ: LAI	-	only the multiple regression was applied
Alpilles	France	07/2002	crops and grassland	-	ρ: LAIeff,LAItrue,LAI57eff, LAI57true,fCover,fAPAR	-	the transfer function using the log(ρ) creates coplanar points and estimates negative LAI, fCover and fAPAR values; the REG method provides better results in terms of cross-validation RMSE
Barrax	Spain	07/2003	cropland	class 2	ρ: LAI,LAI57,fCover,fAPAR	-	very similar results between REG on ρ and REG on log(ρ) in terms of cross-validation RMSE, but the number of ESUs with weights < 0.7 is higher when using the log(ρ)
Concepción	Chile	01/2003	mixed forest	-	ρ: LAI,LAI57,fAPAR	-	very similar results between REG on ρ and REG on log(ρ), but the number of ESUs with weights < 0.7 is higher when using the log(ρ)
Counami	French Guyana	10/2002	tropical forest	classes 1,3	-	-	no relation between NDVI and biophysical variables; the average value of the ESUs is representative since the Counami site is very homogeneous (class 2=clouds)
Fundulea	Romania	03/2001	crops	-	ρ: LAI	-	only the multiple regression was applied
Fundulea	Romania	06/2003	crops	class 1	ρ: LAI,LAI57,fCover,fAPAR	-	similar results between REG on (ρ) and REG on log(ρ) in terms of cross-validation RMSE
Gilching	Germany	07/2002	crops and forests	-	ρ: LAIeff,LAItrue,LAI57eff, LAI57true,fCover,fAPAR	-	the REG method provides better results in terms of cross-validation RMSE for all the variables; close results between REG on (ρ) and REG on log(ρ)
Haouz	Morocco	03/2003	cropland	-	log(ρ): LAI,LAI57,fCover,fAPAR	-	the number of ESUs with weights < 0.7 is lower using the log(ρ); the LUT method provides systematically higher RMSE value than for REG (weighted RMSE for REG)
Hisikangas	Finland	08/2003	forests	class 5	log(ρ): LAI,fCover	-	the LUT method provides systematically higher RMSE value than for REG
Järvselja	Estonia	06/2003	boreal forest	-	log(ρ): LAI; ρ: fCover	-	the REG method provides better results in terms of cross-validation RMSE; for LAI, the results using the log(ρ) are slightly better; a simple multiple regression was applied for fCover
Romilly-sur-Seine	France	07/2000	cropland	-	-	-	LAI map retrieved using collocated kriging
Sud-Ouest	France	07/2002	crops and grassland	-	ρ: LAIeff,LAItrue,LAI57eff, LAI57true,fCover,fAPAR	-	the REG method provides better results in terms of cross-validation RMSE for all the biophysical variables
Turco	Bolivia	04/2003	cropland	-	ρ: LAI,LAI57,fCover,fAPAR	-	very close results between REG on ρ and REG on log(ρ)

*The method AVE is used 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 = multiple robust regression. ***LUT = Look-Up-Tables.

Table 2. The methods used by the transfer functions to estimate the biophysical variable values

Therefore, for the sites that have not been processed up to now, the LUT method will not be tested anymore. The transfer functions will use the REG on reflectance while the REG on logarithm of the reflectance will be the subject of new tests in particular on the forest sites.

3.3.3. Choice of the band combination

For the LAIeff, the XS1,XS2,XS3,XS4,RN (Figure 15 and Figure 16) combination on reflectance was selected since it provides a good compromise between the cross-validation RMSE (among the lowest values), the number of weights lower than 0.7 (five) and the weighted root mean square error (the lowest value).

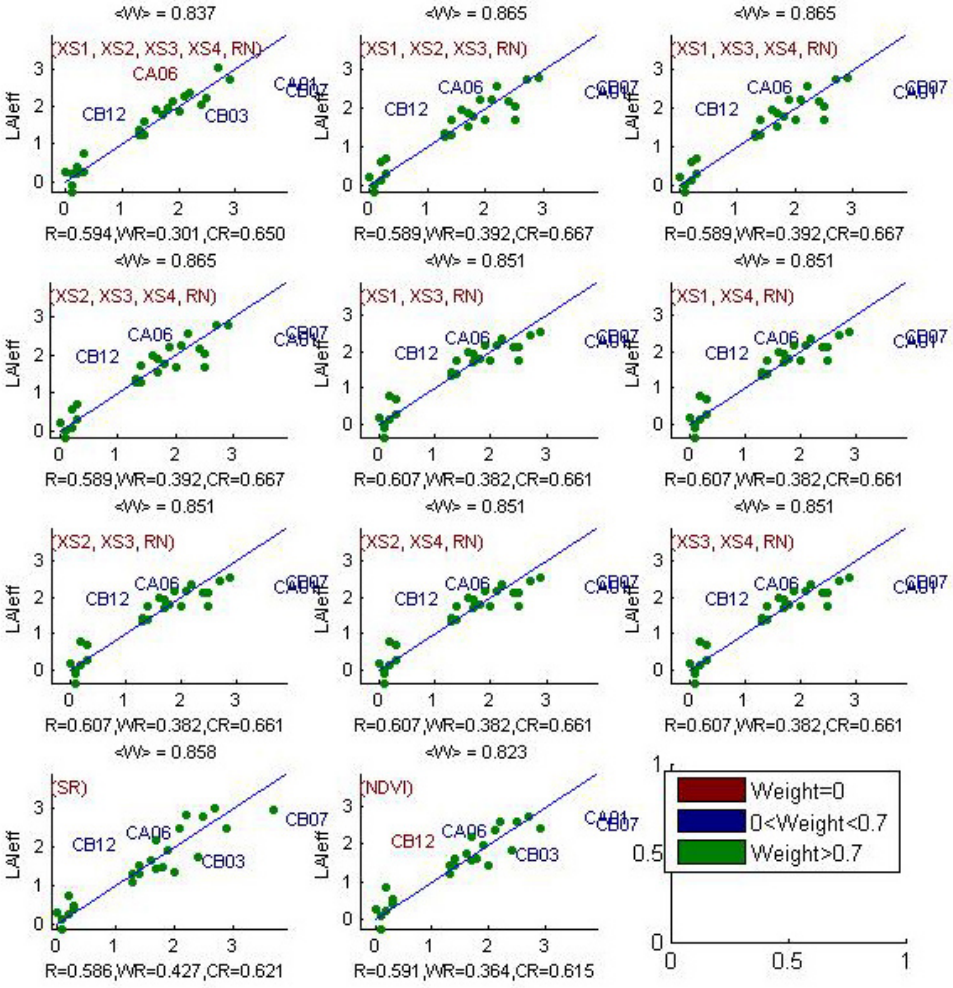


Figure 15. Effective Leaf Area Index: results for regression on reflectance using different band combinations. R is the root mean square error computed between LAIeff and estimated LAIeff. WR is the weighted root mean square error and CR is the cross validation root mean square error.

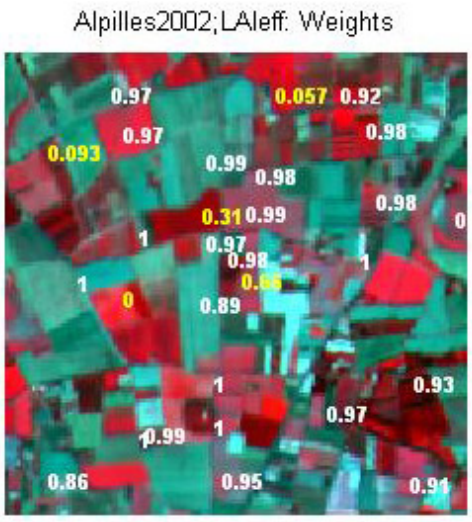


Figure 16. Weights associated to each ESU for the determination of LAIeff transfer function.

For the LAItrue, the NDVI combination on reflectance was selected since it provides the lowest cross-validation RMSE value and a low weighted root mean square error value. However, six ESUs have weights lower than 0.7 (Figure 17 and Figure 18).

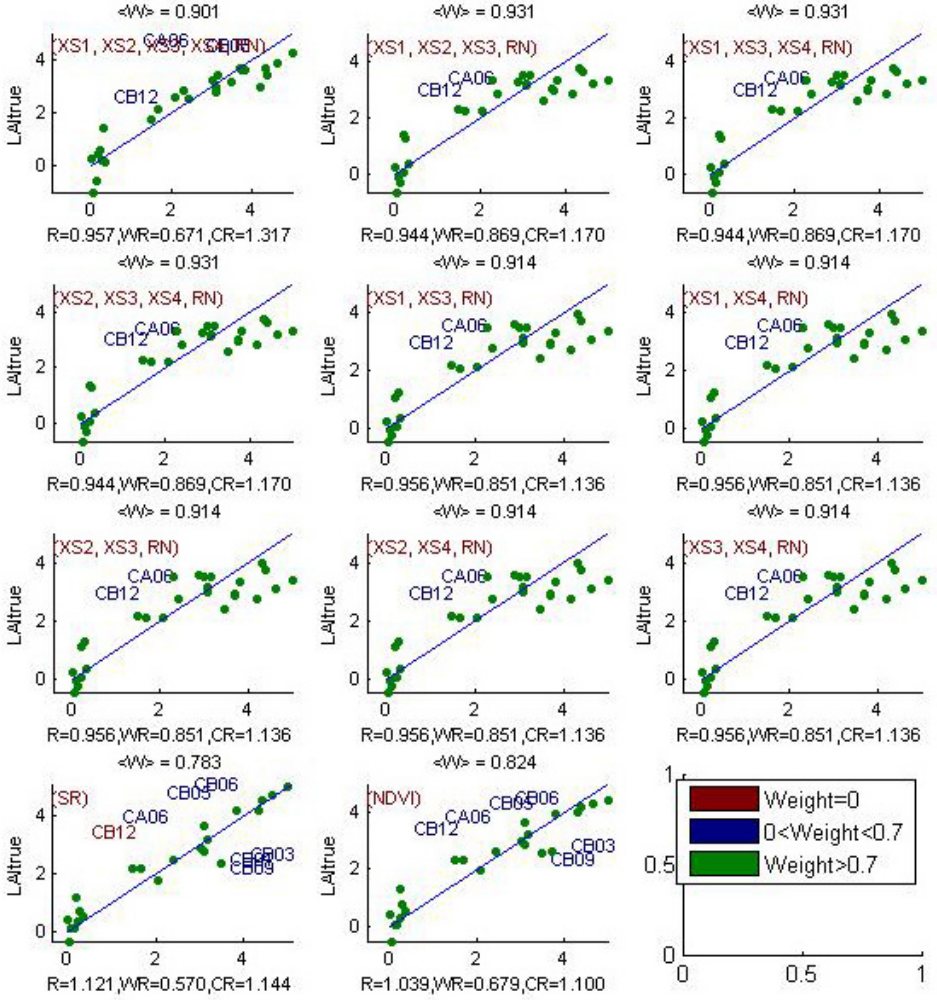


Figure 17. True Leaf Area Index: results for regression on reflectance using different band combinations. R is the root mean square error computed between LAItrue and estimated LAItrue. WR is the weighted root mean square error and CR is the cross validation root mean square error.

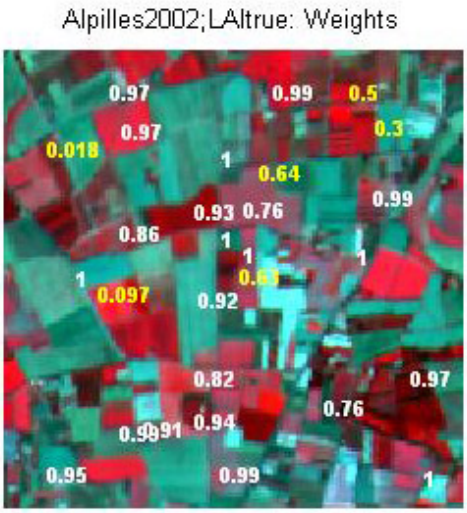


Figure 18. Weights associated to each ESU for the determination of LAItrue transfer function.

For the LAI57eff, the NDVI combination on reflectance (Figure 19 and Figure 20) was selected since it provides a good compromise between the cross-validation RMSE (the lowest value), the number of weights lower than 0.7 (five) and the weighted root mean square error (among the lowest values).

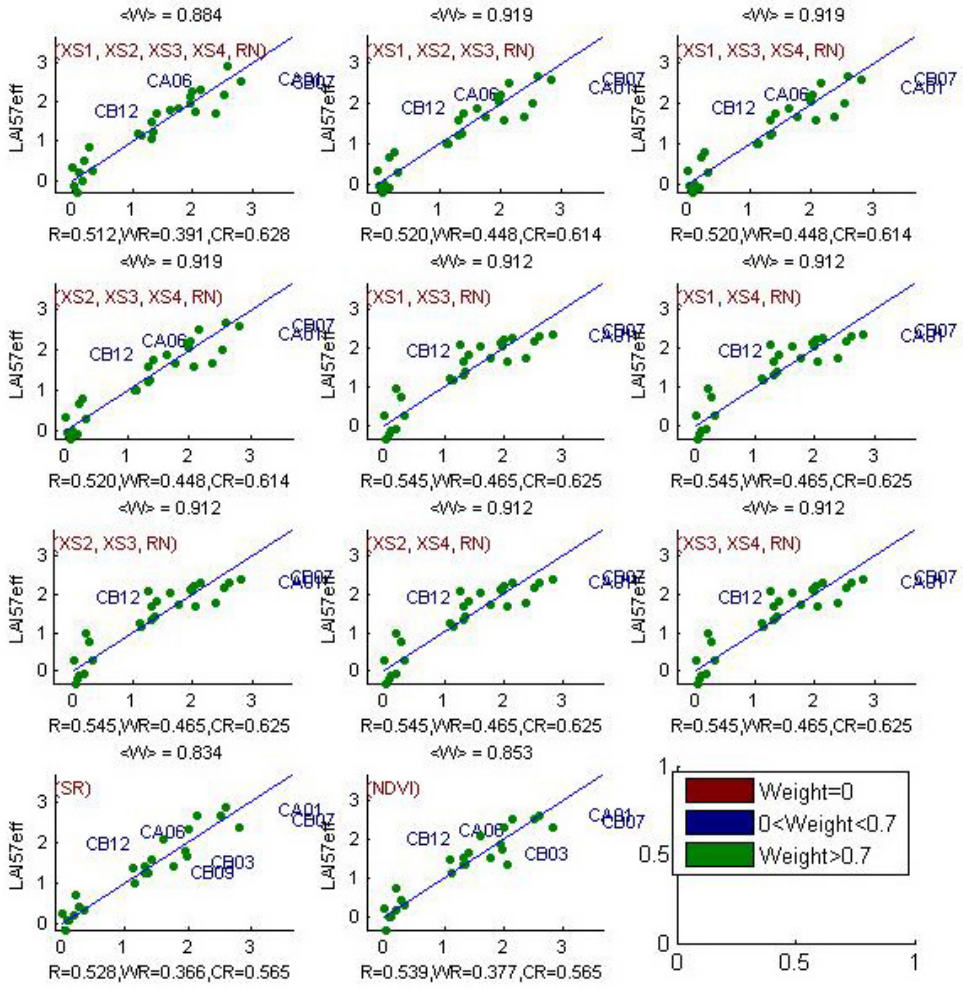


Figure 19. Effective LAI at 57.5°: results for regression on reflectance using different band combinations. R is the root mean square error computed between LAI57eff and estimated LAI57eff. WR is the weighted root mean square error and CR is the cross validation root mean square error.

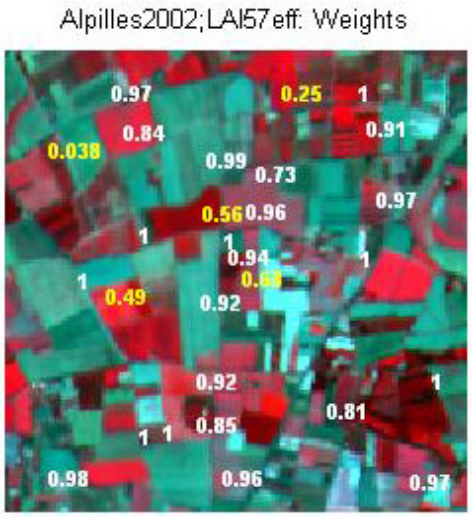


Figure 20. Weights associated to each ESU for the determination of LAI57eff transfer function.

For the LAI57true, the XS1,XS2,XS3,XS4,RN combination on reflectance (Figure 21 and Figure 22) was selected since it provides a good compromise between the number of weights lower than 0.7 (one), the cross-validation RMSE and the weighted root mean square error (among the lowest values).

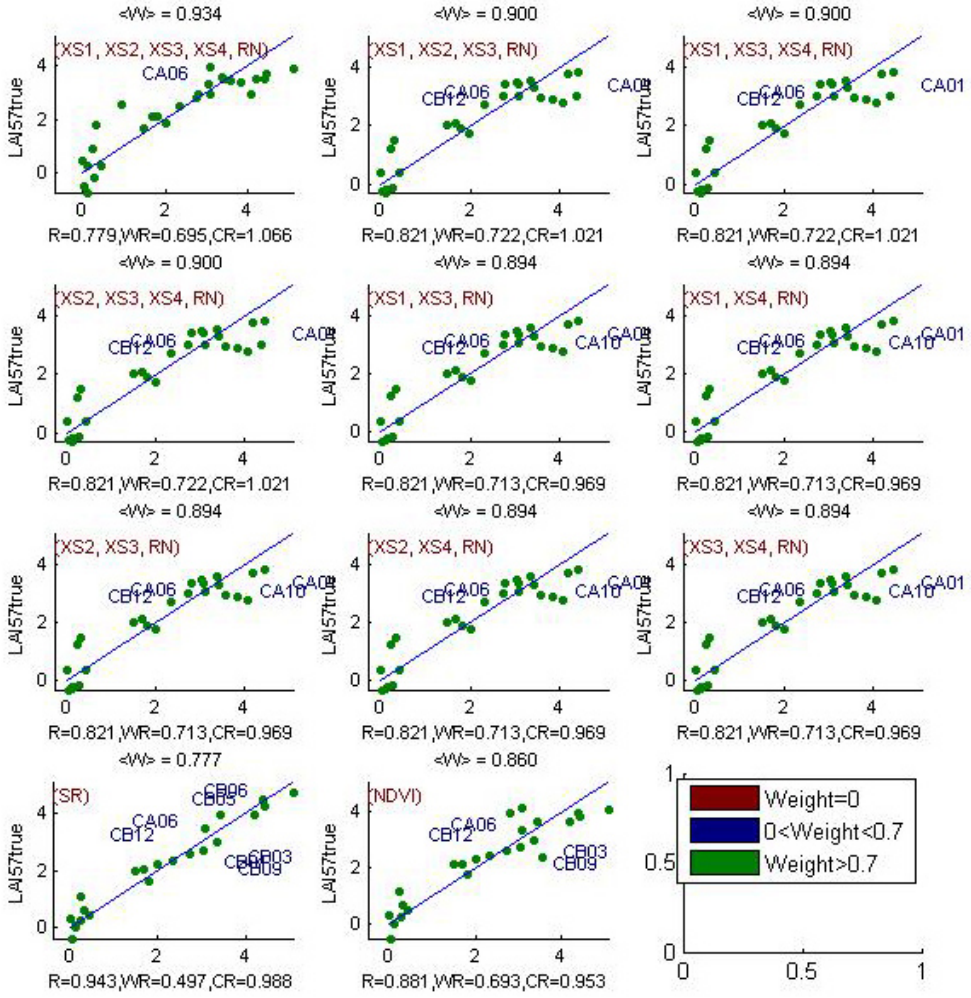


Figure 21. True Leaf Area Index at 57.5°: results for regression on reflectance using different band combinations. R is the root mean square error computed between LAI57true and estimated LAI57true. WR is the weighted root mean square error and CR is the cross validation root mean square error.

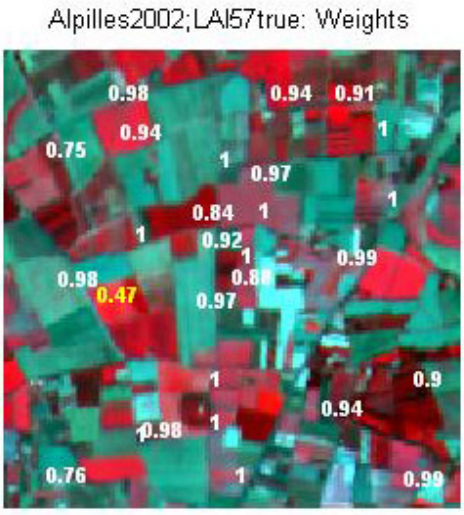


Figure 22. Weights associated to each ESU for the determination of LAI57true transfer function.

For the fCover, the XS3,XS4,RN combination on reflectance was selected since it provides a good compromise between the number of weights lower than 0.7 (one), the cross-validation RMSE (the lowest value) and the weighted root mean square error (Figure 23 and Figure 24). The following band combinations provide the same results: [XS1,XS3,RN]; [XS1,XS4,RN]; [XS2,XS3,RN]; [XS2,XS4,RN].

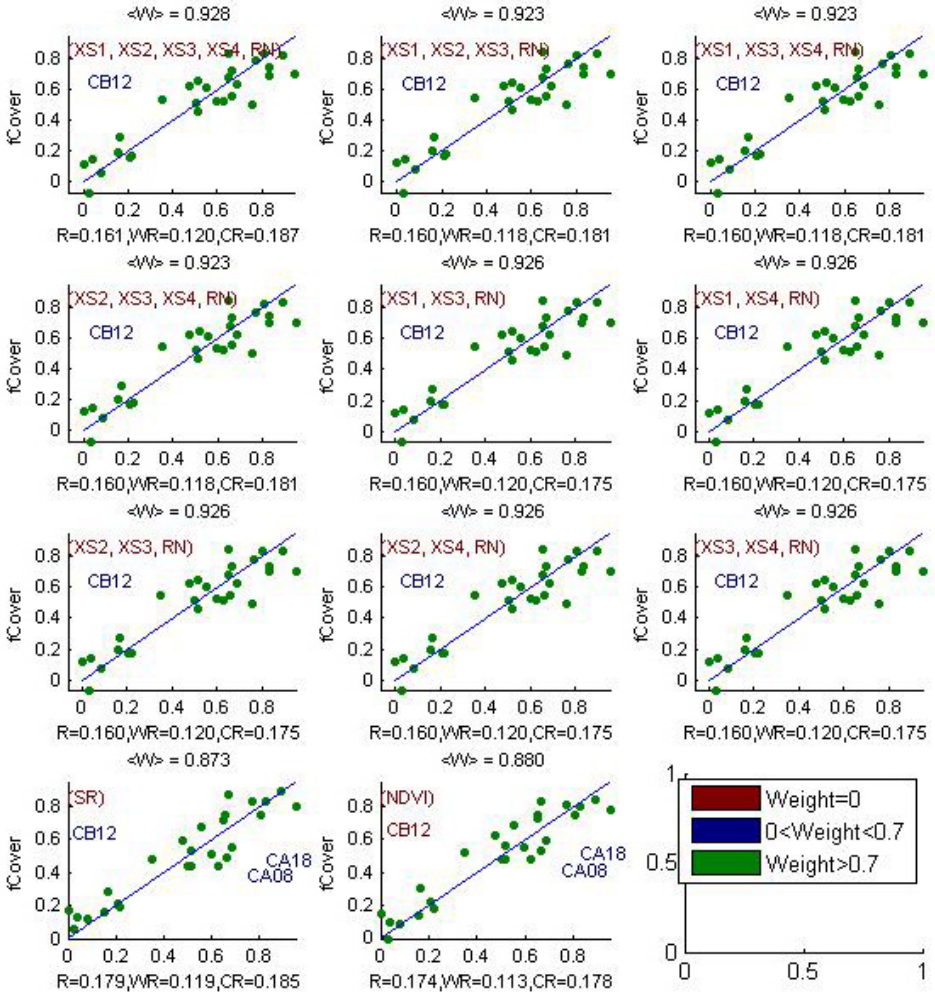


Figure 23. fCover: results for regression on reflectance using different band combinations. R is the root mean square error computed between fCover and estimated fCover. WR is the weighted root mean square error and CR is the cross validation root mean square error.

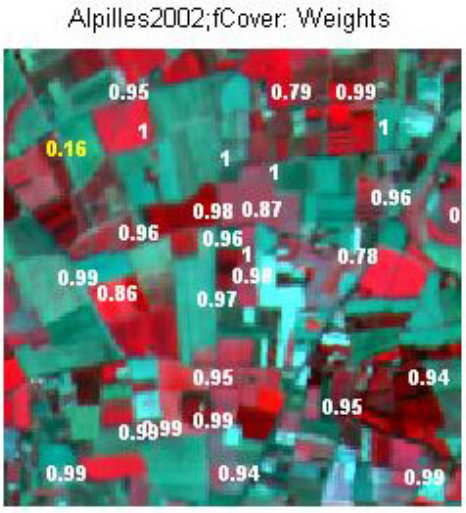


Figure 24. Weights associated to each ESU for the determination of fCover transfer function.

For the fAPAR, the XS3,XS4,RN combination on reflectance was selected since it provides the lowest cross-validation RMSE value, the lowest weighted root mean square error value and one weight lower than 0.7

(Figure 25 and Figure 26). The following band combinations provide the same results: [XS1,XS3,RN]; [XS1,XS4,RN]; [XS2,XS3,RN]; [XS2,XS4,RN].

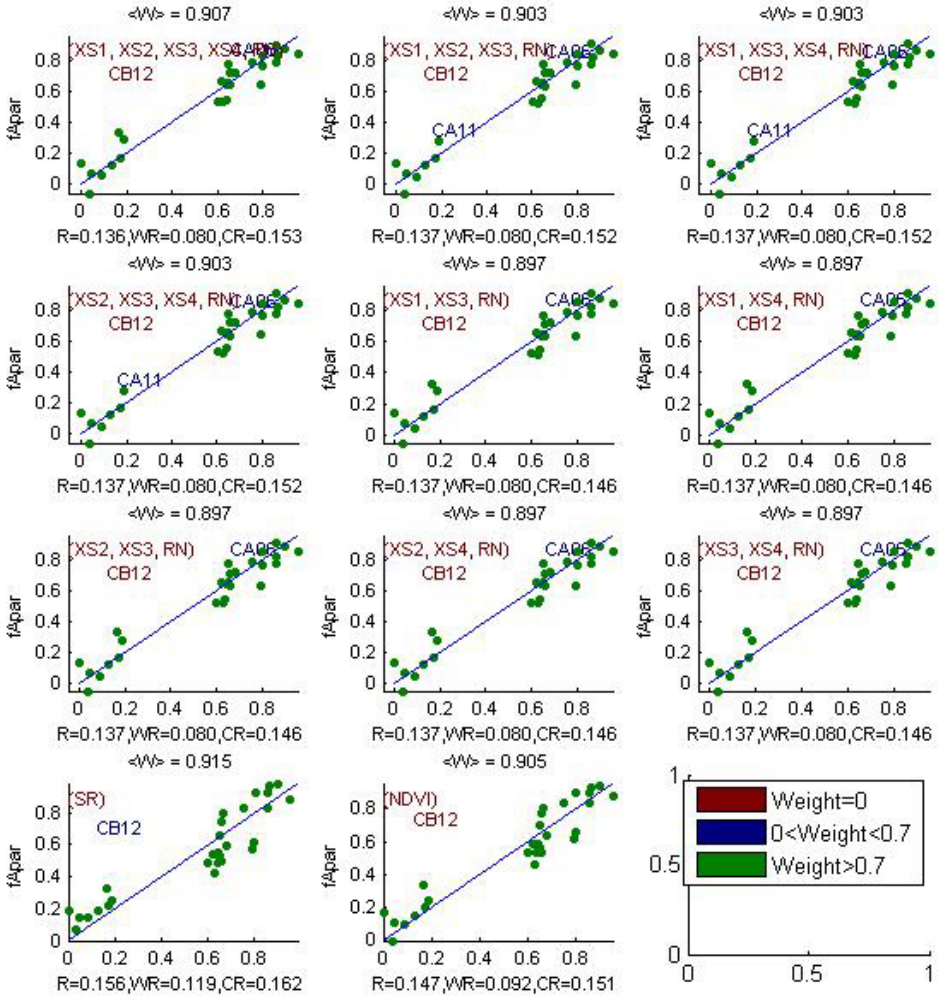


Figure 25. fAPAR: results for regression on reflectance using different band combinations. R is the root mean square error computed between fAPAR and estimated fAPAR. WR is the weighted root mean square error and CR is the cross validation root mean square error.

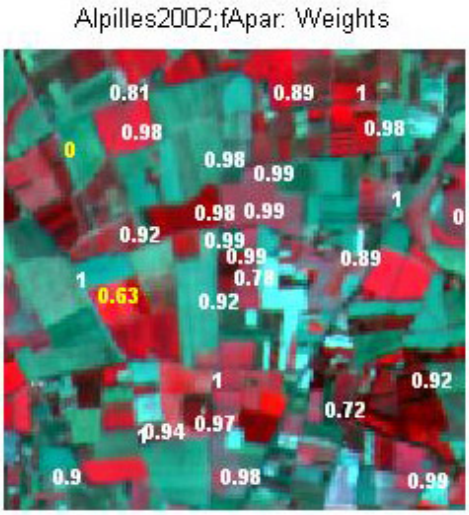


Figure 26. Weights associated to each ESU for the determination of fAPAR transfer function.

Following, the results of the transfer function (Table 3):

Variable	Band Combination	RMSE	Weighted RMSE	Cross-valid RMSE
LAIeff	-0.40176 - 11.5(XS1) + 22.793(XS2) + 12.354(XS3) - 8.607(XS4) -47.242(RN)	0.594	0.301	0.650
LAItrue	-2.6613 + 9.1391(NDVI)	1.039	0.679	1.100
LAI57eff	-1.5413 + 5.3112(NDVI)	0.539	0.377	0.565
LAI57true	0.74817 + 28.779(XS1) + 26.461(XS2) + 14.955(XS3) - 20.428(XS4) - 125.66(RN)	0.779	0.695	1.066
fCover	1.0061 - 6.5545(XS3) - 8.9584(XS4) + 29.141(RN)	0.160	0.120	0.175
fAPAR	0.8325 + 6.9631(XS3) - 15.139(XS4) + 15.705(RN)	0.137	0.080	0.146

RN = Red*NIR

Table 3. Transfer function applied to the whole site for the different biophysical variables, and corresponding errors

3.4. Applying the transfer function to the Alpilles SPOT image extraction

Figure 27 presents the biophysical variable maps obtained with the transfer function described in Table 3. 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 27). This was expected when looking the LAIeff/LAItrue relationship in Figure 27, showing that for high LAI the difference between the two can be significant.

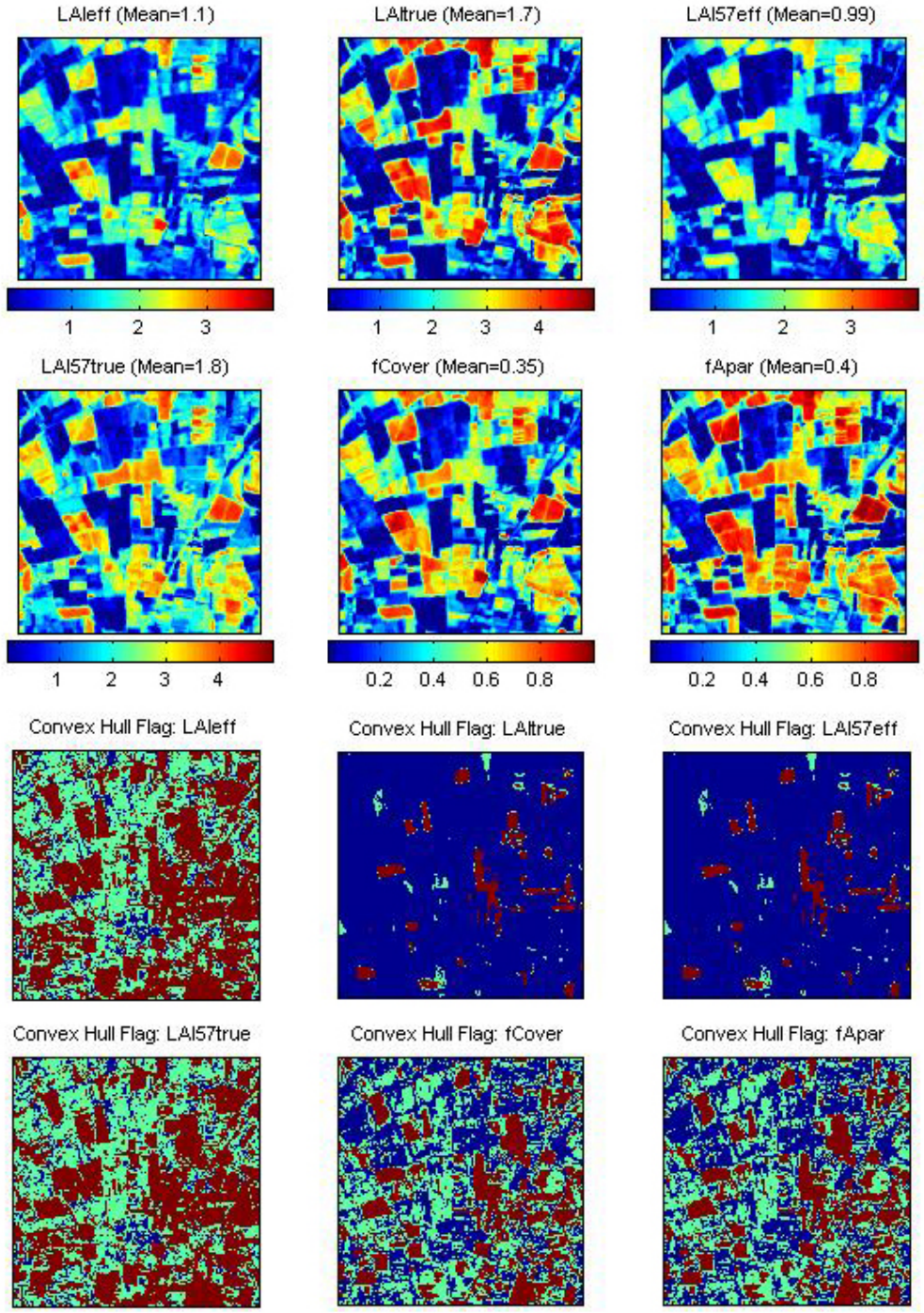


Figure 27. High resolution biophysical variable maps applied on the Alpilles site (top). Associated Flags are shown at the bottom: blue and light blue corresponds to the pixels belonging to the 'strict' and 'large' convex hulls and red to the pixels for which the transfer function is extrapolating.

The flag maps are different between the biophysical variables since the number and the bands used for the regression are different. The results are comparable between LAIeff and LAI57true, between LAItrue and LAI57eff and between fCover and fAPAR. The pixels outside the two convex hulls are quite numerous for LAIeff and LAI57true. For these variables, the extrapolation of the transfer function is large all over the site. Note that few pixels are outside the strict convex hull for LAItrue and LAI57eff. This is due to the choice of the combinations. In theory, the more the number of bands increases, the larger the extrapolation is.

4. Conclusion

The transfer functions are obtained by using 29 ESUs. The representativeness of the land cover of the different ESUs is not optimal. However, the results of the robust regression are good and the maps obtained for the biophysical variables are consistent. The flag associated to each map show also that the transfer function is mainly used as an extrapolator for LAIeff and LAI57true but the choice of the band combinations (§3.4) is decisive.

For all the variables, the regression coefficients are computed by relating the variable itself to reflectance even if the results on logarithm of the reflectance are better. The transfer function using the logarithm of the reflectance creates coplanar points which do not allow to determine the 'strict' and 'large' convex hulls (§3.3.1) and it provides many negative LAI, fCover and fAPAR values on the scale of the Alpilles site. The multiple robust regression on logarithm of the reflectance will be the subject of new tests in order to know if it is really useful to compute it. Note that the next processes will not take into account the LUT to estimate the biophysical variable values since the LUT method is never used by the transfer function (§3.3.1).

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

5. Acknowledgements

We thank people who participated to the field experiment: **Cédric Bacour, Sébastien Garrigues, Franck Oro, Nadine Bruguier** (INRA, Avignon), **Roland Bosseno** (IRD/INRA, Avignon), **Eleanor Stevens** and **Marie Weiss** (Noveltis, Toulouse).

ANNEX

Ground measurement acquisition report for the VALERI site **Alpilles**

sampled from 22/07/2002 to 23/07/2002

Sébastien Garrigues
Organization: INRA
email: sebastien.garrigues@avignon.inra.fr

Date of report 15/10/2002

People participating to the field experiment:

Fistname & Name	0rganization
Cédric Bacour	Inra, Avignon
Roland Bosseno	IRD/INRA CSE, Avignon, France
Sébastien Garrigues	Inra, Avignon
Franck Oro	Inra, Avignon
Eleanor Stevens	
Marie Weiss	Noveltis, Toulouse
Nadine Bruguier	Inra, Avignon

Site coordinates

	Lat-Long WGS84 (Deg min.00)		UTM / WGS84 UTM 31	
	Lat.	Long.	Easting	Northing
Upper left corner	43°49'27"	4°41'47"	636412	4853753
Lower right corner	43°47'48"	4°43'58"	639412	4850753

Ground control points

GPSAlpilles2002.xls contains different waypoint taken on the site:

% GCP1 638799 4852376
% GCP2 637800 4852575
% GCP3 637791 4852890

GPS system used: Garmin12CX device and Garmin e-Trex devices.
Typical uncertainty of GPS position: 6-7 m.

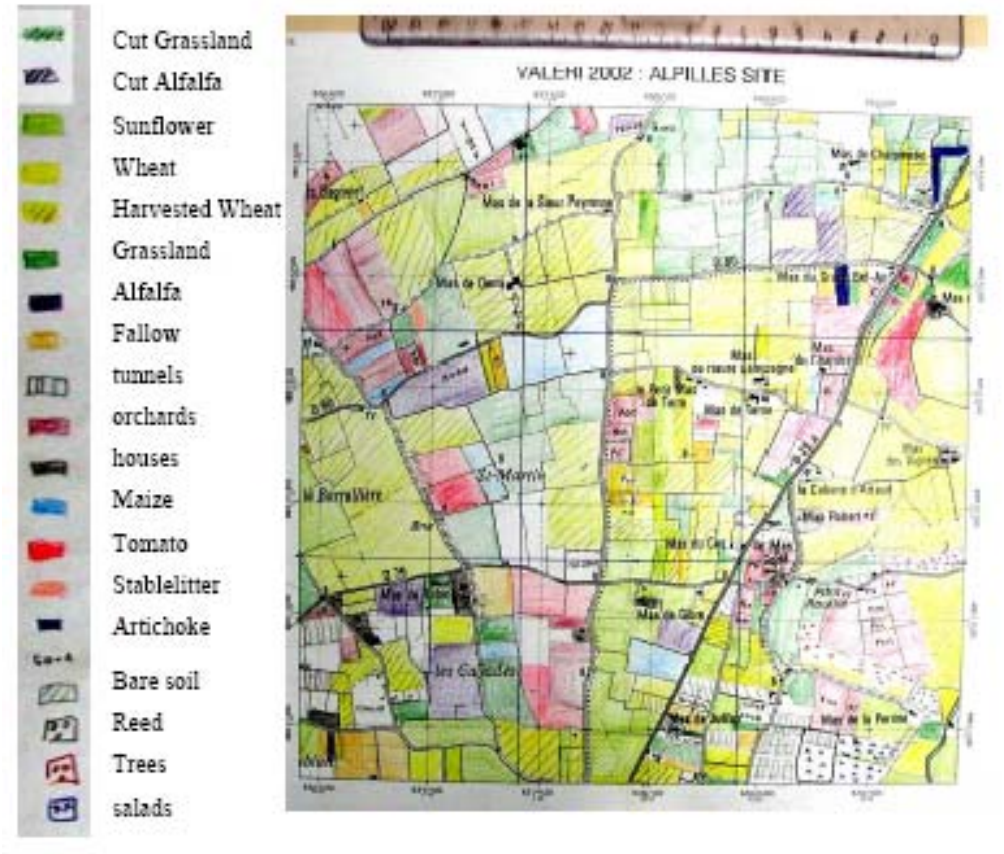
Description of the site and land cover

Category according to IGBP classification
Croplands.

Comments on the land cover
Narrow fields, main crops: sunflower,maize, tomato,Alfalfa.
Field size : between 2 ha and 8 ha.

Topography
The site is at about 10-20m altitude. It is generally quite flat.

Land cover map

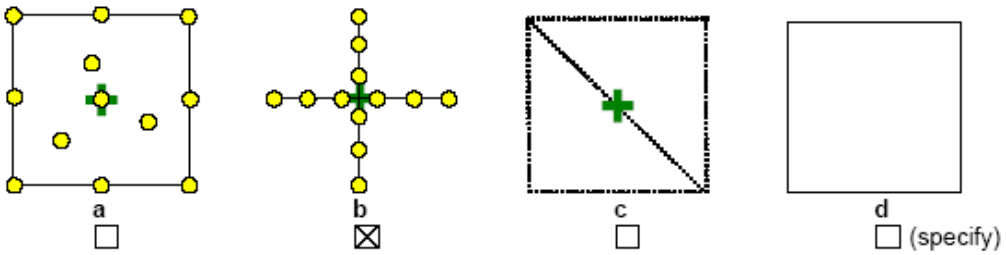


Spatial sampling scheme

Sensors used for sampling the ESUs

	Method	Comments
☒	Hemispherical photographs	
☐	LAI2000	
☐	TRAC	
☐	Ceptometer	
☐	Direct measurements	
☐	Other	

Sampling strategy for the ESU



☐ ☒ ☐ ☐ (specify)

Distribution of the Elementary sampling units

The crosses distribution was used : 12 hemispherical photos are taken over each ESU.

The high spatial resolution image

Satellite

Satellite used	SPOT4 HRV1
Level of processing	1A
Projection type	UTM 31/ WGS 84
Acquisition date	20/07/2002

The image was kindly provided by the Chrysalide project (CHRIS), funded by the Programme National de Teledetection Spatiale. In the frame of this project, the image was geo-referenced, as well as corrected from atmospheric effect (Abadi, 2002). The geoferencing accuracy is very good since a image coregistration was applied.

List of the ESUs

The GPSAlpilles2002.xls file contains the information for each ESU:

%Name	Easting (m)	Northing (m)	Note
% GCP1	638799	4852376	crossing roads
% GCP2	637800	4852575	crossing roads
%GCP3	637791	4852890	crossing roads
ULC	636412	4853753	
LRC	639412	4850753	
CB01	637993	4852687	sunflower
CB02	637864	4852383	apple tree orchard, above and below acquisitions (understorey)
CB03	637931	4852238	sunflower
CB04	638841	4852760	tomato
CB05	638776	4853212	alfalfa, height 40cm, homogeneous
CB06	638604	4853468	grassland, homogeneous, different species
CB07	638186	4853461	grassland, homogeneous, different species
CB08	637722	4853026	harvested wheat, green regrowth, heterogeneous
CB09	638045	4852910	sunflower, same as CB01
CB10	637183	4853195	sunflower, height: 1.70m
CB11	637134	4853447	harvested wheat, LAI=0, no photo
CB12	636691	4853076	peach tree orchard, below photos, distance between rows = 4m, height = 3m
CA01	637658	4852663	maize
CA02	637708	4852489	grassland, very low and sparse vegetation
CA03	637695	4852080	harvested wheat + green regrowth
CA04	637164	4852521	alfalfa, heterogeneous
CA05	636771	4852223	harvested wheat + green regrowth
CA06	637031	4852111	tomato
CA07	639317	4852633	sunflower, it stings!
CA08	638621	4852359	tomato
CA09	638552	4851799	harvested wheat + green regrowth
CA10	639094	4851554	apple tree orchard, above and below acquisitions (dense green understorey), high
CA11	638500	4851363	harvested wheat + green regrowth
CA12	639075	4850895	salads
CA13	637164	4851192	alfalfa (sparse)
CA14	637317	4851225	grassland
CA15	636690	4850916	fallow
CA16	637788	4850925	sunflower
CA17	637668	4851279	peach tree orchard, peaches not mature, hedges at the middle of the field
CA18	637667	4851566	tomato

Acknowledgements

Many thanks to the farmers who let us walk into their fields (J. GRANGIER, Yvonne CAVALIER, Henriette PALOT, E. ESPIGUE, G. GONTIER, A. MISTRAL, N. CLARION, C. VIVARES, L. GRANGIER, T. CHASSEFIERE…). Thanks also to the Chrysalide project for providing the SPOT image. Thanks to all the people who participated to the field campaign and to the field campaign preparation (thanks to Nadine and Marie !!)

Photo gallery

The photos illustrating the campaign are to be stored in the directory “photo gallery” and the labels should be indicated in the table above. For each ESU, a panoramic photo was taken, the photo name is the ESU number

#	File name	Comments
1	CYXX.jpg	ESU Panoramic photo for Team Y and photo number XX
2	CHRIS&VALERITeamsatLunch.jpg	
3	HappyEnd.jpg	
4	HappyWork.jpg	
5	HemiTeamAB.jpg	
6	HowToUseAcar.jpg	
7	MixedTeams.jpg	
8	SebIsAnExcellentDriver.jpg	
9	TeamsAB.jpg	

Additional comments

Very friendly campaign.