VALERI 2003 : Barrax site (cropland)

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

Marie Weiss

1 Introduction

This report describes the production of the high resolution, level 1, biophysical variable maps for the Barrax site in 2003 (see campaign report for more details about the site and the ground measurement campaign). 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 1.3) developed at INRA-CSE.

The derived biophysical variable maps are:
- Leaf Area Index: two LAI are considered, the first one corresponds to effective LAI derived from the description of the gap fraction as a function of the view zenith angle, the second one (LAI57) is derived from the gap fraction at 57.5°, which is independent on the leaf inclination and is also an effective LAI (does not take into account clumping effect).
- cover fraction (fCover) : it is the percentage of soil covered by vegetation between o° et 10° view zenith angle
- fAPAR: it is the fraction of Absorbed Photosynthetically Active Radiation (PAR=400-700nm). The fAPAR can be defined as instantaneous (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 mns and the daily fAPAR is the lowest for instantaneous fAPAR at 10h00 AM (local time, RMSE= 0.021). Therefore, the derivation of fAPAR from CAN-EYE corresponds to the instantaneous black sky fAPAR at 10h00 AM.

2 Available data

2.1 Sampling strategy

Figure 1 shows that the ESUs locations are well spatially distributed over the 5km x 3km site. The site has been extended to a 5kmx5km area in order to add bare soil points (Beatriz Martinez, personal communication) which are at the top and bottom limits of the image (ESUs 56,57,58).
The processing of the ground data has shown that:
- ESUs E32, E40, E26 (in black on Figure 1) were located on field borders. Those three ESUs were eliminated
- Considering that SPOT geo-location and GPS measurements are associated to errors, we found that processed LAI for ESUs E8, E11, E17, E34 and E38 did not correspond to the SPOT pixel in terms of reflectance: they have been shifted of 1 pixel, according to D. Béal and F. Baret who participated to the measurements.

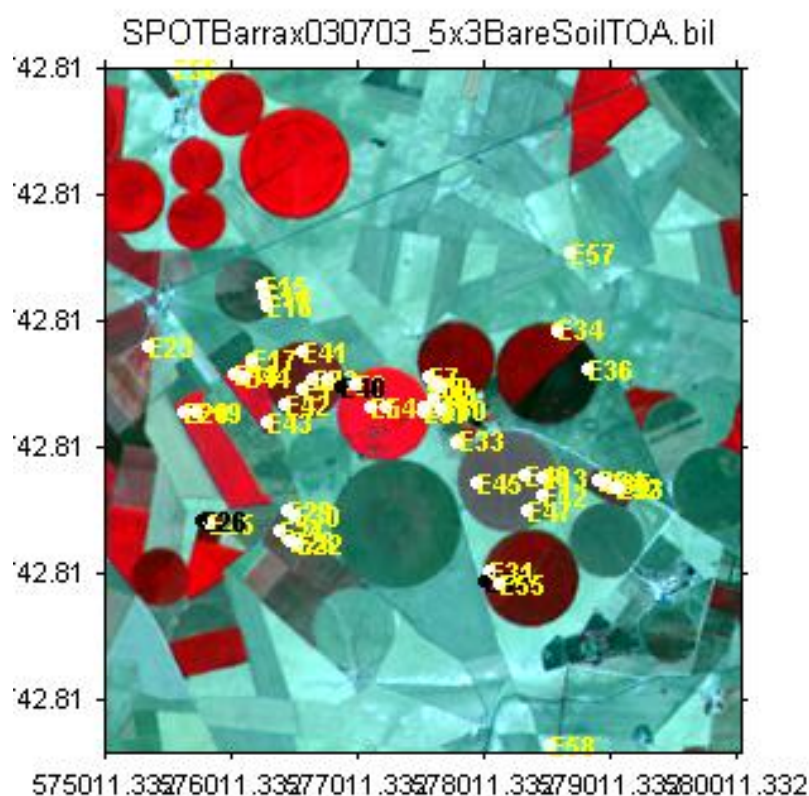Finally 48 ESUs have been kept for the computation of the transfer function.

Figure 1. Distribution of the ESUs around the Barrax site. ESUs in black were eliminated for the computation of the transfer function
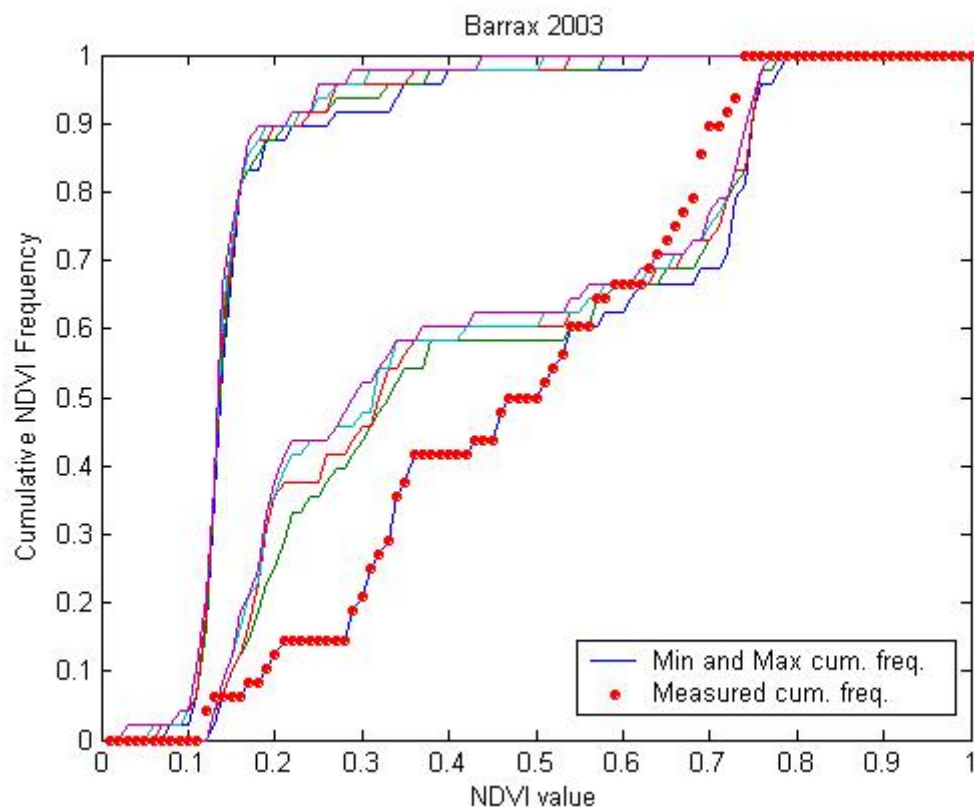

Figure 2. Comparison of the ESU NDVI distribution and the NDVI distribution over the whole image.

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 2). As the number of pixels is drastically different for the ESU and whole site (WS=22500 in case of a 3x3km 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 2 shows that the NDVI distribution of the 48 ESUs is not well representing the NDVI distribution on the 5kmx5km image. As a lot of bare soils area are observed on this site, and only 3 bare soil points are considered, low NDVI values are not well sampled. Therefore, as we consider cumulative frequency, the 'ESU' curve is outside the 'boundary curves', mainly because of the mis-representation of bare soils. However, the effect of the ESU NDVI distribution should be slight since bare soil points are taken into account in the regression. This can be checked by looking at the LAI map which should show very low LAI values for bare soil areas observed on the SPOT image.

2.2 SPOT image

The SPOT image was acquired the 3rd july 2003 by HRVIR2 on SPOT4. It was geo-located by SPOTimage (SPOTView basic). The projection is UTM 30N, WGS84, 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 5kmx3km site. Therefore, it will be taken into account everywhere in the same way. Figure 3 shows the relationship between RED and near infrared (NIR) SPOT channels: the soil line is well marked, and no saturated points are observed.
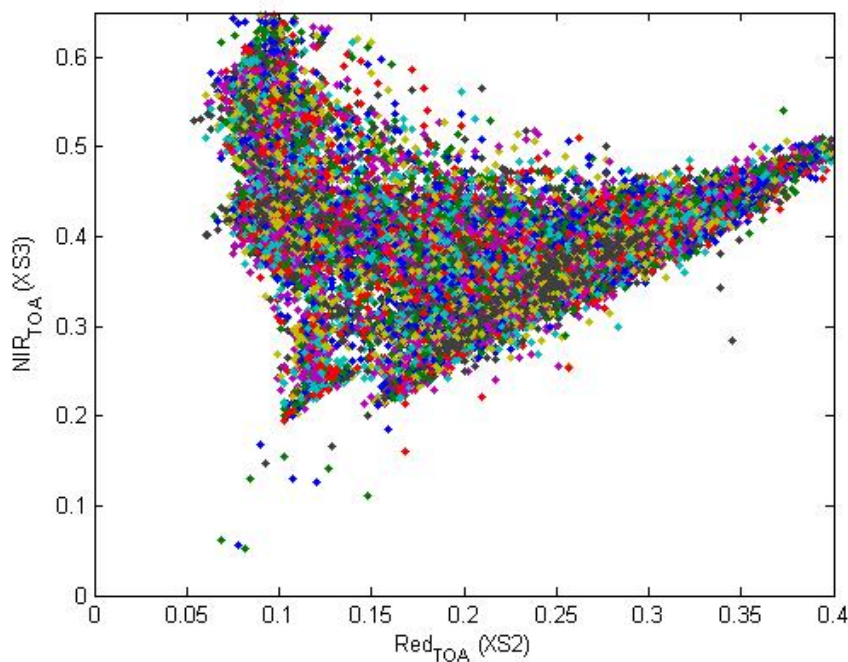

Figure 3. Red/NIR relationship on the SPOT image for Barrax, 2003.

A non supervised classification based on the k_means method (matlab statistics toolbox) was applied to the NDVI of the SPOT image to distinguish if different behaviours on the image for the biophysical variable-reflectance relationship exist. A number of 5 classes were chosen (Figure 4). The repartitions of the classes on the image and on the ESUs are quite different, since bare soils (class 2) are not sampled enough.
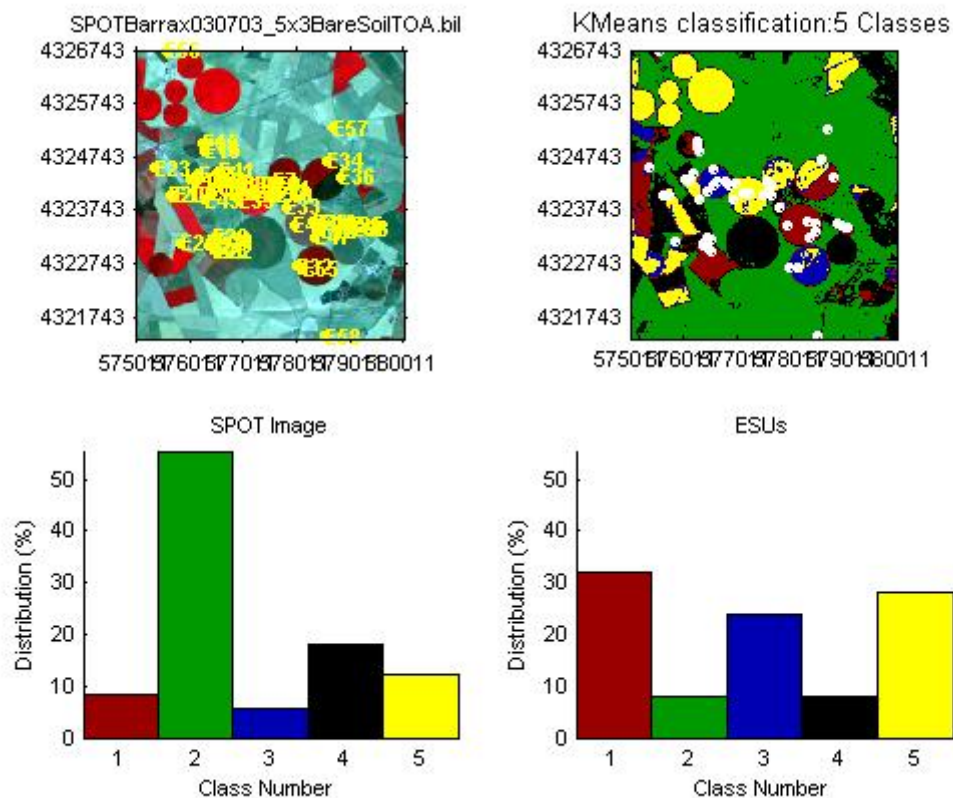

Figure 4. Classification of the SPOT image. Comparison of the class distribution between the SPOT image and sampled ESUs.

2.3 Hemispherical images

The hemispherical images were processed by the CAN-EYE software (Version 1.3) to derive the biophysical variables. Figure 5 shows the distribution of the different measured variables over the sampled ESUs. Green LAI derived from directional gap fraction and LAI derived from gap fraction at 57.5° are consistent and varies from 0 to 6 (alfalfa fields).

Figure 6 shows the different relationships observed between the biophysical variables and corresponding NDVI or the ESUs, as a function of the SPOT classes determined in §2.2. No different behaviours between the classes can be observed. However, Figure 4 shows that the bare soil class is under-represented in the ESU sampling. Therefore, for this class, for which we know that LAI is null, we use the one transfer function,, using a mean value of 0 and determine a single transfer function for all the other classes.
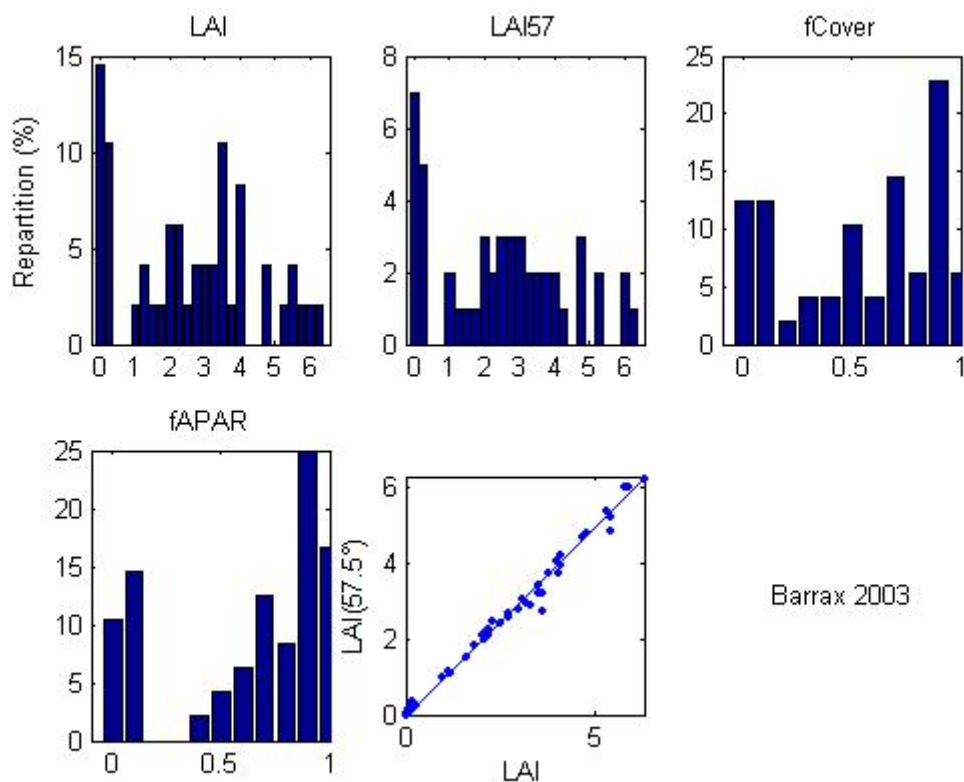
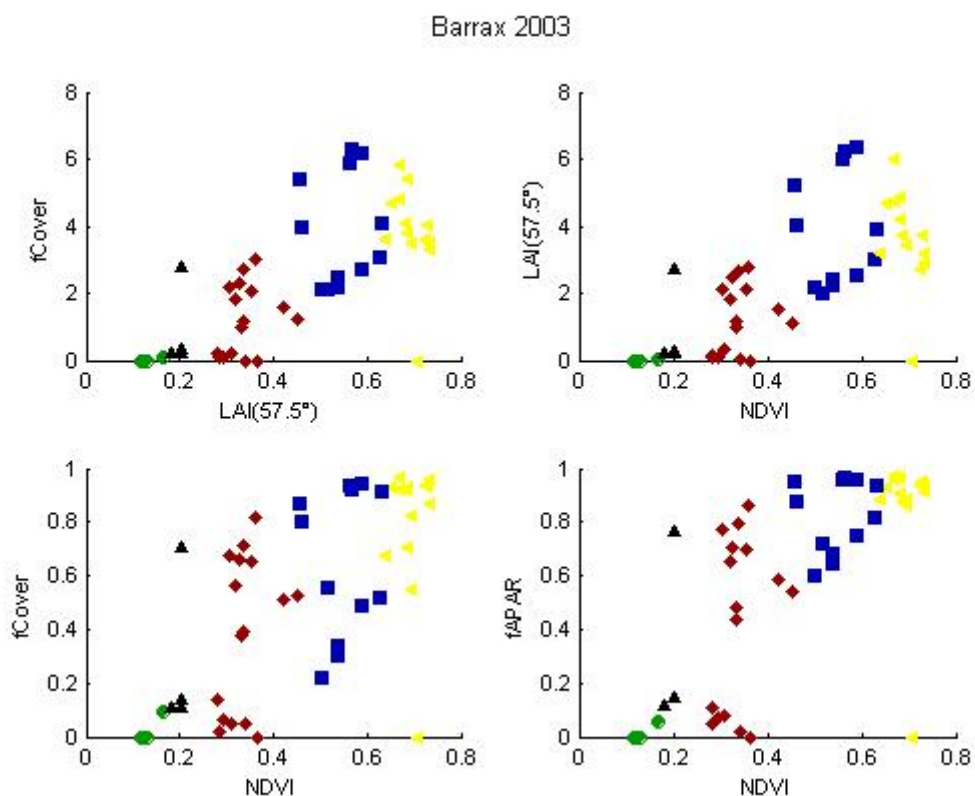**Figure 5. Distribution of the measured biophysical variables over the ESUs.**

Figure 6. NDVI-Biophysical Variable relationships as a function of SPOT classes

3   Determination of the transfer function for the 4 biophysical variables

3.1 Tested Transfer functions

For each class determined in §2.2, three types of transfer functions are tested:

- AVE: If the number of ESUs belonging to the class is too low, the transfer function consists only in attributing the average value of the biophysical variable measured on the class to each pixel of the SPOT image belonging to the class.

- REG: If the number of ESUs is sufficient, multiple robust regression between ESUs reflectance (or Single Ratio) and the considered biophysical variable can be considered: we used the 'robustfit' function from the matlab statistics toolbox. It uses an iteratively re-weighted least squares algorithm, with the weights calculated at each iteration by applying the bisquare function to the residuals from the previous iteration. This algorithm gives 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 build using ESUs reflectances and corresponding measured biophysical variable. For a given pixel, a cost function is computed as the sum 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. In §2.1, 2.2it has been shown that the two distribution are quite different because of the presence of bare soil areas. Therefore, the results of this method will be shown for Barrax site but will not be applied to derive the biophysical variable maps.

Both regression and Look-Up-Tables are tested using either the reflectance or the logarithm of the reflectance for any band combination, plus the simple ratio. As both methods have poor extrapolation capacities, a flag image, based on the computation of convex hull over reflectances, is computed showing:
- Pixels belonging to the class for which AVE method was applied
- Pixels inside the 'strict convex-hull': for each class (not AVE method), a convex-hull is computed using all the reflectance combination used for the transfer function, and corresponding to the ESUs belonging to the class. For those pixels, the transfer function is used as an interpolator, and the degree of confidence in the results obtained is quite high.
- Pixels inside the 'large convex-hull': for each class (not AVE method), a convex-hull is computed using all the reflectance combination (±5% in relative value) used for the transfer function, and corresponding to the ESUs belonging to the class. For those pixels, the transfer function is used as an extrapolator (but not far from interpolator), and the degree of confidence in the results obtained is quite good.
- Pixels outside the two convex-hulls: this means that for these pixels, the transfer function acted like 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.

3.2 Results on the Barrax site

Figure 4 shows that the bare soil class is under-represented in the ESU sampling. Therefore, for this class, for which we know that LAI is null, we use the AVE method, using a mean value of 0. For the other classes (1,2,4,5 = 44 ESUs), we have tested REG and LUT methods using all the classes together to keep a reasonable number of data for the regression.

3.2.1 Choice of the method for classes (1,2,4,5)

Figure 7 and Figure 8 show the results obtained for all the possible band combinations using either the reflectance or the logarithm of the reflectance:

- As, the NDVI distribution of the ESUs (Figure 2) does not well represent the whole site, which could represent a risk for establishing the map since the LUT method works only for interpolation purposes. Moreover, the REG method provides similar and even better results for all the variables and is therefore selected as the transfer function.

- Using either the logarithm of the reflectance or the reflectance itself provides very similar results in terms of cross-validation RMSE. However, the number of points with a weight lower than 0.7 in the robust regression is higher when using the logarithm of the reflectance. Therefore we choose to use the robust multiple regression using the ESUs reflectance value.
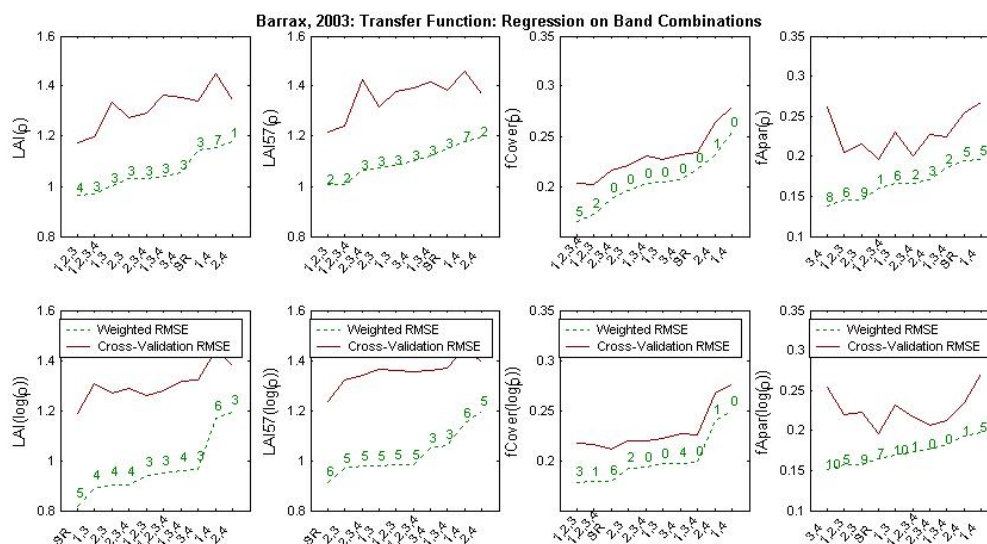

Figure 7. Transfer function: test of multiple regressions applied on different band combinations. Band combinations are given in abscissa. The estimated biophysical variable is given in ordinate. Top graphs correspond to regression made on reflectance (ρ): the weighted root mean square error (RMSE) is 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. Bottom graphs correspond to regression made on the logarithm of the reflectance.
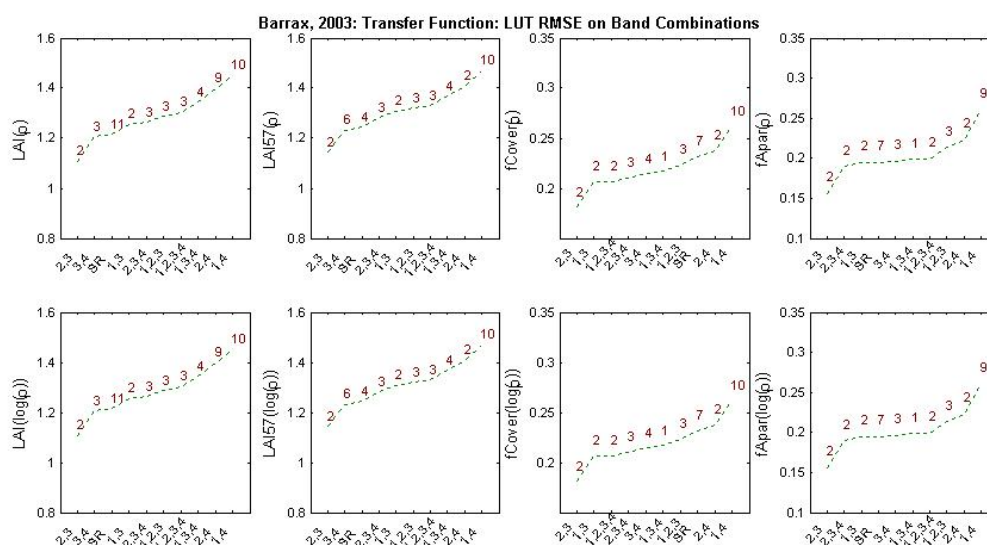

Figure 8. Transfer function: test of LUT applied on different band combinations. Band combinations are given in abscissa. The estimated biophysical variable is given in ordinate. Top graphs correspond to regression made on reflectance (ρ): the root mean square error is presented in green. The numbers

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.

3.2.2 Choice of the band combination for classes (1,2,4,5)

For the effective LAI and LAI at 57.5°, Figure 9 and Figure 12 show that some ESUs have systematically a weight lower than 0.7 for all tested band combination: E41 and E42 (LAI and LAI57) correspond to dense alfalfa fields and E52 to grassland. For these canopies, hemispherical photographs are taken from above and the pixels corresponding to bare soil on the images may not be easy to be identified during the classification process of the CAN-EYE Software. Indeed, when comparing the processing of two different users (B. Martinez, university of Valencia, Spain and F. Baret, INRA-CSE, France) on the ESUs, the results are quite different (Figure 10). Note that the RMSE value is quite high (around 1).

The (XS1,XS2,XS3,XS4) combination for LAI and (XS1,XS2,XS3) for LAI at 57.5° were selected since it presents a good compromise between the RMSE values and the number of points with a weight lower than 0.7 (Figure 11).
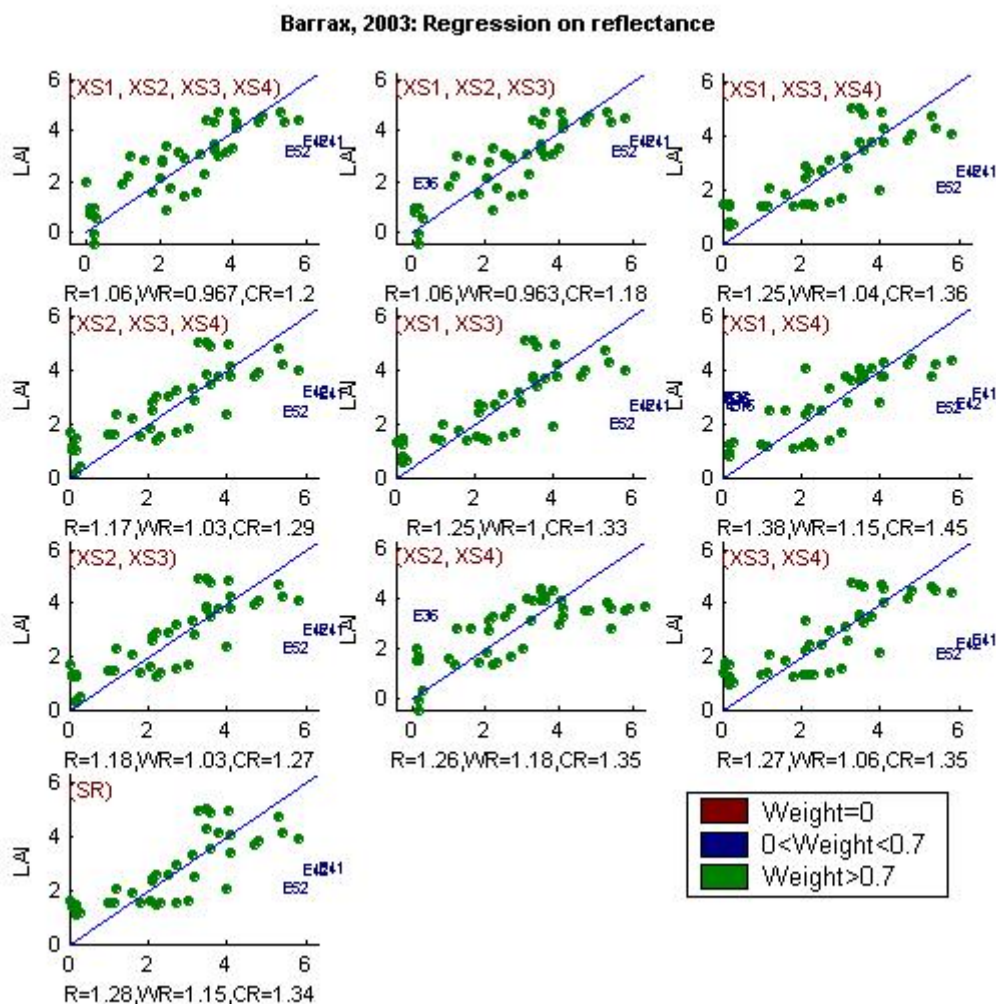

Figure 9. Effective Leaf Area Index: results for regression using different band combinations. R is the root mean square error computed between LAI and estimated LAI. WR is the weighted root mean square error and CR is the cross validation root mean square error.
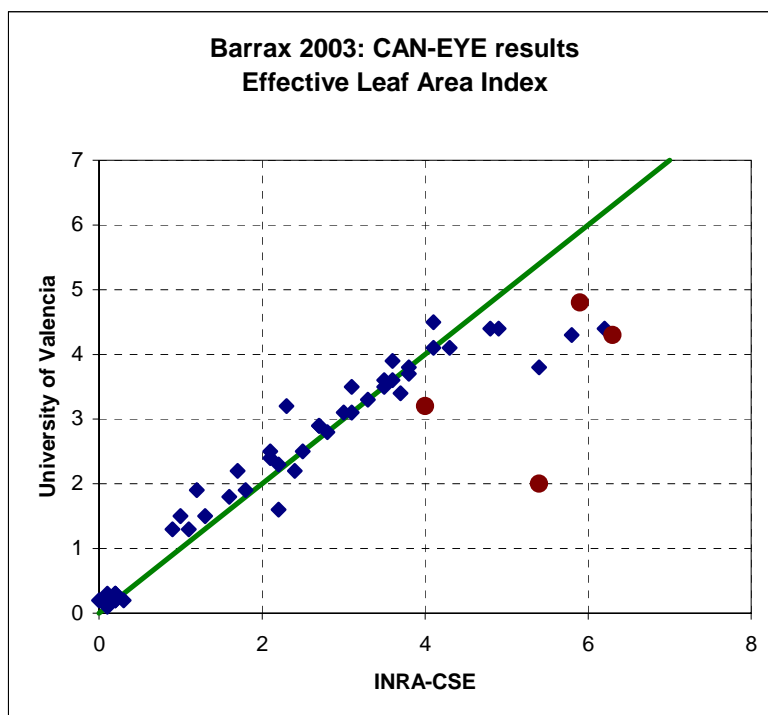
Figure 10. Comparison of CAN-EYE processing results from two different users. Points in red correspond to ESUs which have systematically a weight lower than 0.7 regarding the tested band combination.
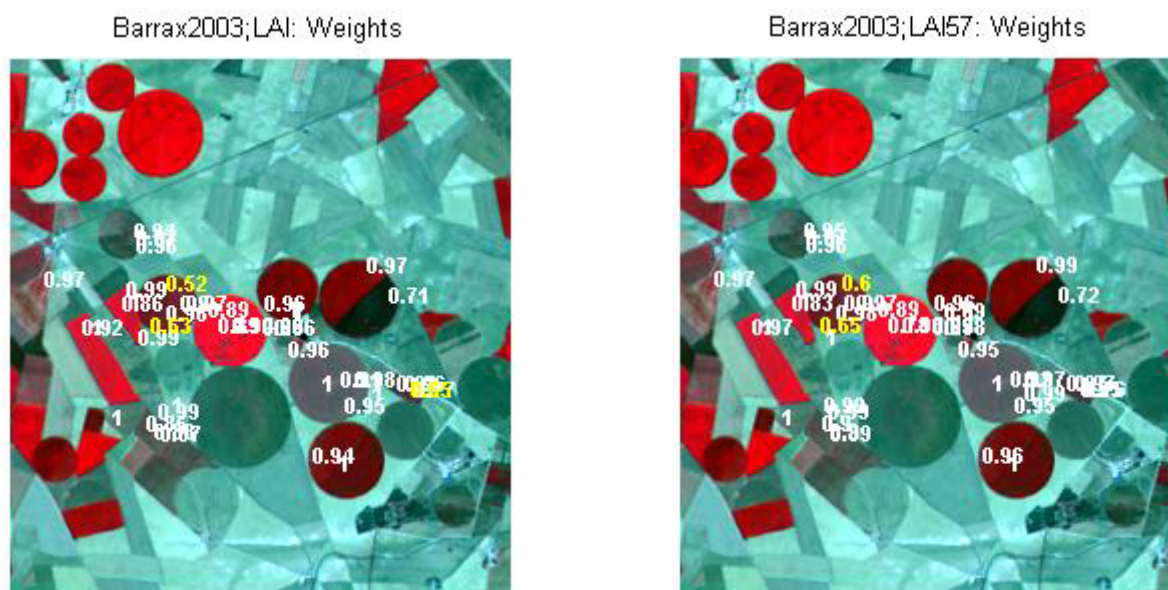

Figure 11. Weights associated to each ESU for the determination of LAI (left) and LAI57 (right) transfer function.
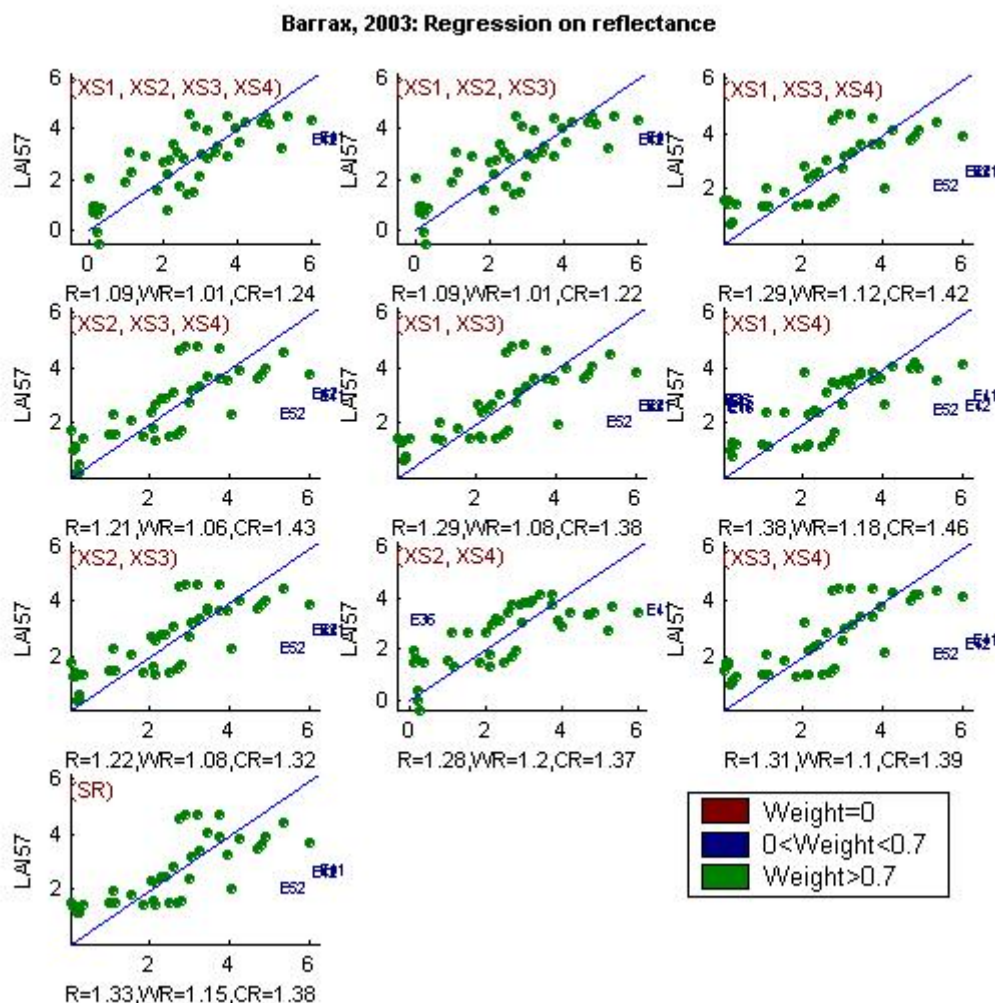
Figure 12. Effective Leaf Area Index at 57.5°: results for regression using different band combinations

For fCover (XS2,XS3,XS4), and fAPAR (XS1,XS2,XS3,XS4), compromise has to be made between RMSE values and the number of ESUs with weight lower than 0.7. Note that the ESUs causing problems for LAI are not the same for fCover and fAPAR. ESUs 21 and 22 correspond also to dense canopies (alfalfa and sugar beet) and the result may depend on the CAN-EYE processing. For fAPAR, ESU 14,15,16 and 36 correspond to fields which are mostly in the senescent phase, therefore the corresponding reflectance may be disturbed by the senescent vegetation. Moreover, during the classification phase in the CAN-EYE processing it may be sometimes difficult to make a separation between green parts and yellow ones.

Variable	Band Combination	RMSE	Weighted RMSE	Cross-valid RMSE
Effective LAI	0.24 +99.21XS1-76.84XS2-2.31XS3+ 0.53XS4	1.06	0.97	1.20
Effective LAI (57.5°)	0.26+112.54XS1-83.51XS2-4.38XS3	1.09	1.01	1.22
fCover	-0.11-4.97XS2+1.88XS3+2.04XS4	0.20	0.19	0.22
fAPAR	0.14+12.35XS1-13.61XS2+0.15XS3+ 2.18XS4	0.17	0.16	0.20

Table 1. Transfer function applied to the whole site for the different biophysical variables, and corresponding errors
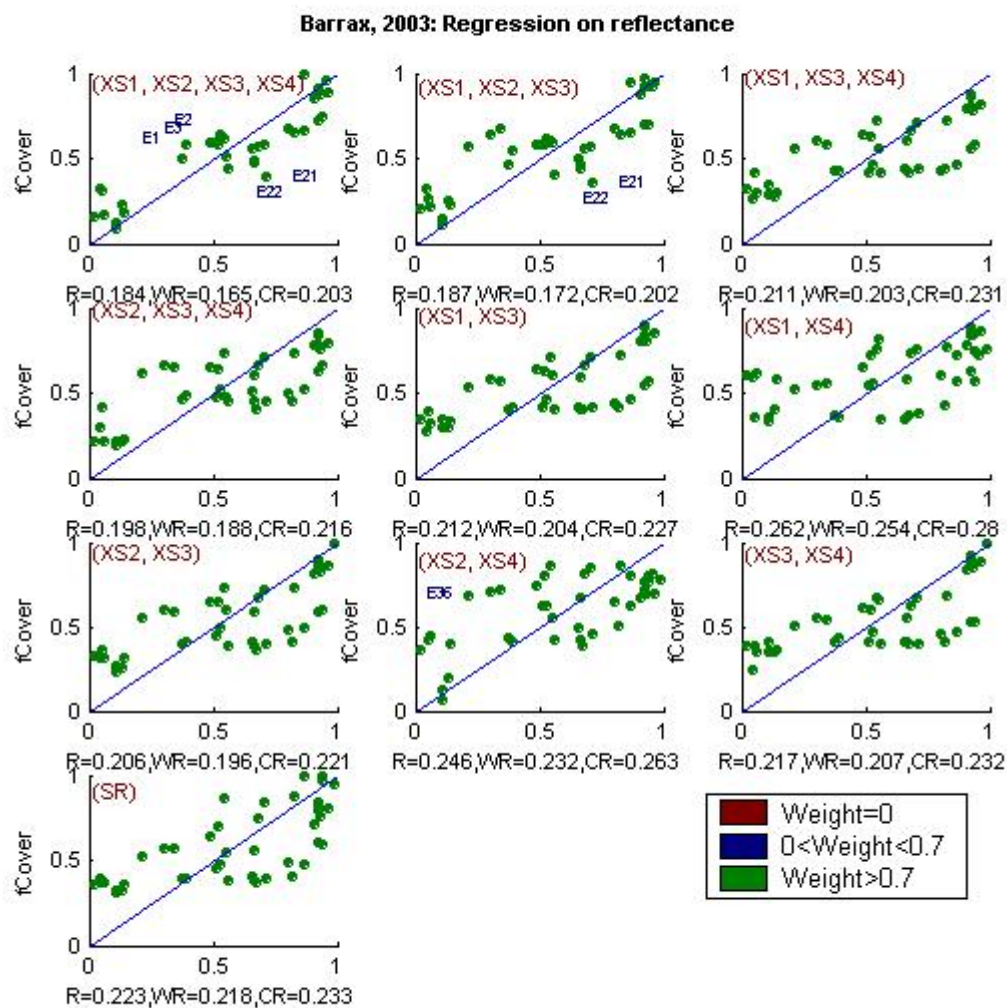
Figure 13. : fCover: results for regression using different band combinations.
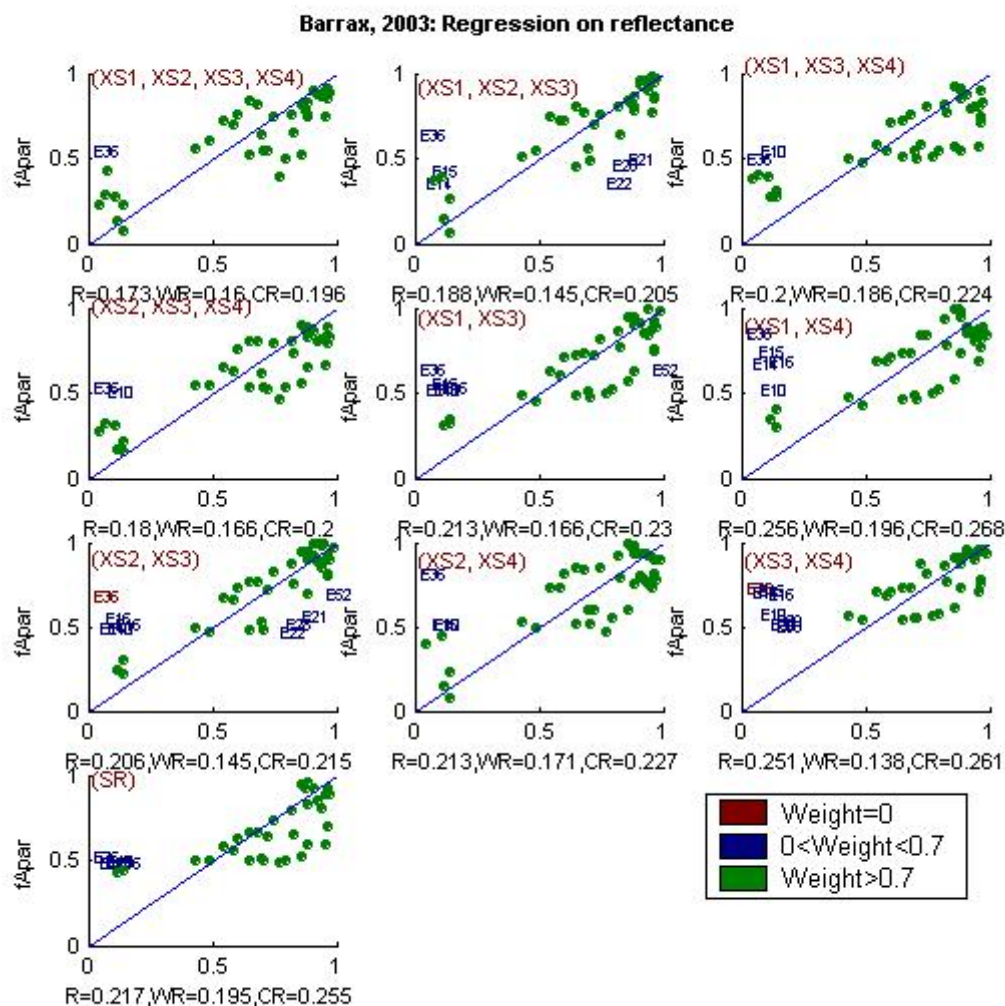
Figure 14. fAPAR : Results for regression using different band combinations

Figure 15. Weights associated to each ESU for the determination of fCover (left) and fAPAR (right) transfer function

3.3 Applying the transfer function to the Barrax SPOT image extraction

Figure 16 presents the biophysical variable maps obtained with the transfer function (REG) described in Table 1 for classes (1,2,4,5) and the AVE method for class 3. The maps obtained for the different variables are consistent, showing similar patterns, low LAI values where low fAPAR and fCover are observed and conversely. Note that the average value for LAI and LAI57 are consistent (around 0.8 for both).

The flag maps show also that a large part of the image correspond to bare soil (55%) and that the transfer function is interpolating for 15% for LAI57, 21% for fCover and 23% for both LAI and fAPAR. These pixels correspond for a big part to two round fields located at the centre and the upper left corner of the image, with a red flag (Figure 16) for LAI, fCover and fAPAR. The field in the center presents a sort of dark green colour on the composite SPOT image (Figure 4) which corresponds to null LAI. The field at the upper left corner is represented in red in the composite SPOT image showing that dense vegetation can be observed. Figure 16 show that the LAI estimated for this field is quite high. Note that for LAI57, the centre of the field appears in blue and green flags, which means that the transfer function is used as an extrapolator for the other variables, but the extrapolation is closed to the boundaries of the convex hull.
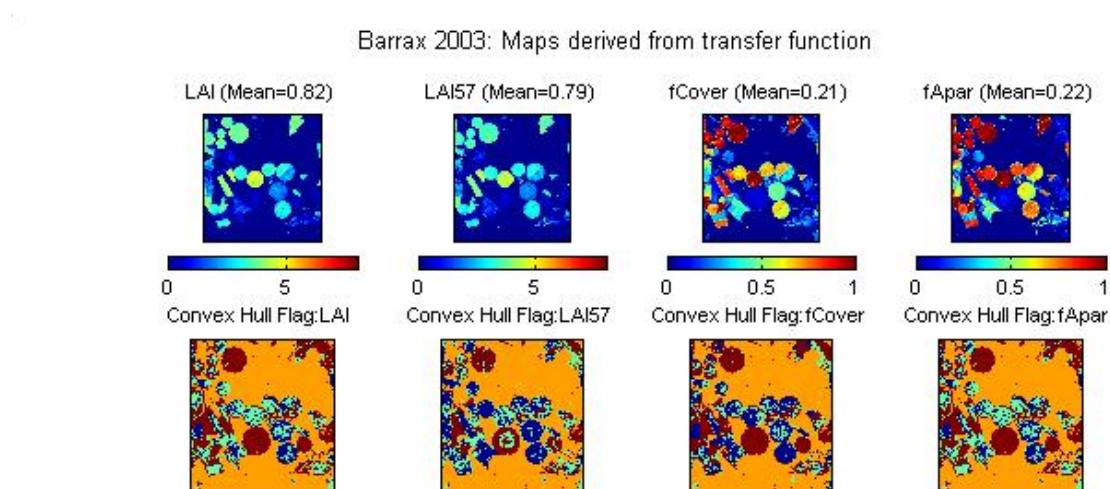

Figure 16. High resolution biophysical variable maps applied on the Barrax site (top). Associated Flags are shown at the bottom: orange corresponds to the bare soil class for which we applied the averaging method. dark blue and green corresponds to the pixels belonging to the 'strict' and 'large' convex hulls, and red to the pixels for which the transfer function is extrapolating.

4 Conclusion

The transfer functions are finally obtained by using Reg and 44 ESUs together for the pixel which do not belong to the bare soil class. A value of 0 was applied to bare soil pixels. For all the variables, the regression coefficients are computed by relating the variable itself to the reflectance. The band combinations are different from one regression to another. Results show good consistency between the variables and the flag associated to each map show that the transfer function is used as an extrapolator (red) in little areas or close to the boundaries of the convex-hull composed of the ESU pixel reflectance.

The biophysical variable maps are available in two projections:
- Plate carrée: latitude/longitude in WGS84 at 1.7857142857e-004° resolution
- UTM, 30N, WGS84 at 20m resolution.

The transfer function was applied on the SPOT image which was geo-located in UTM30N, WGS84. The resampling to obtain the plate carrée projection was performed using the nearest neighbour convolution to get consistent flags.

5 Acknowledgements

We want to thank:

- Beatriz Martinez (University of Valencia), David Béal, Fréderic and Benjamin Baret for the organisation and participation to the campaign, as well for all the documentation they provided on the site.
- Frédéric Baret (INRA-CSE) for the processing of CAN-EYE data
- Beatriz Martinez for the numerous comments and exchange on the results obtained on this site
- José Moreno who let us make a VALERI site during the 2003 SPARC campaign (http://gpds.uv.es/sparc/)