VALERI 2003 : Turco 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 Turco site in 2003 (see campaign report for more details about the site and the ground measurement campaign, available at 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 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 3km x 3km site, which is completely flat. However, few points have been sampled in the upper left and lower right corners where differences in reflectance can be observed. The upper left area is a river (where water can be present or not) although the lower right corner corresponds to a place which is also better fed in water. Two crossing paths in the middle of the image can also be observed. Note also that the NDVI distribution is very tight between 0.1 and 0.2 which indicates very low vegetation as expected (Figure 2).

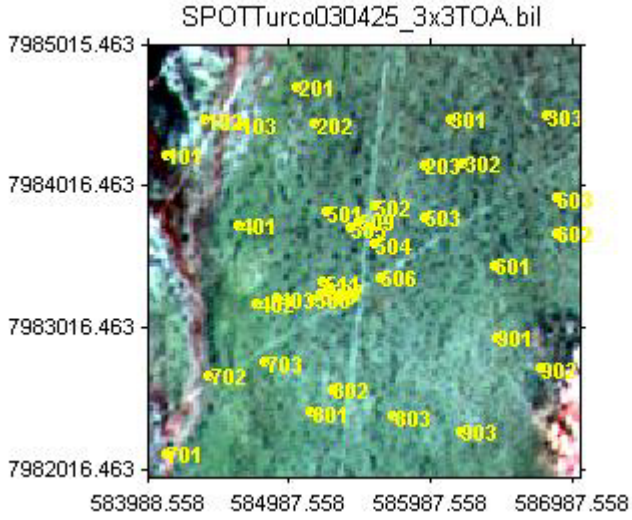

Figure 1. Distribution of the ESUs around the Turco site.

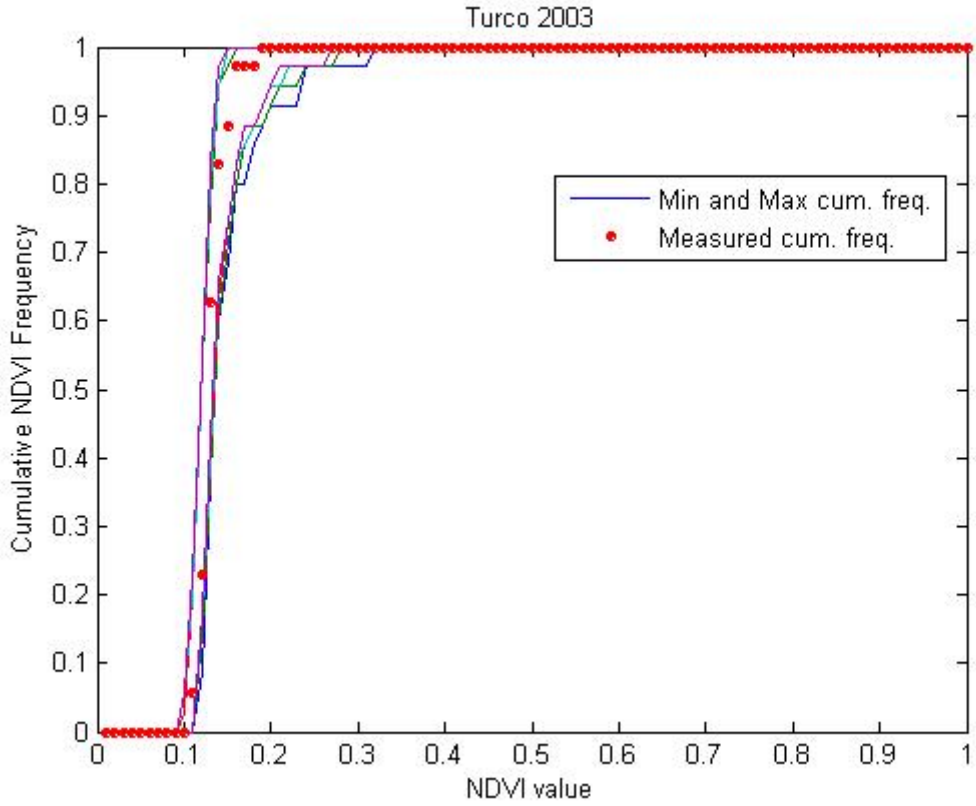

Figure 2. Comparison of the ESU NDVI distribution and the NDVI distribution over the whole image.

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 35 ESUs is well representing the NDVI distribution on the 3kmx3km image since the cumulative frequency curve is between the boundaries for high NDVI values. This implies that few extrapolations will be needed for the derivation of the biophysical variable maps.

2.2 SPOT image

The SPOT image was acquired the 25th April 2003 by HRVIR1on SPOT4. They have been geo-located by SPOT image using maps. The projection is UTM19S, PSAM 56 and is fully described in the campaign report that can be downloaded at http://www.avignon.inra.fr. 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 3kmx3km 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.

A non supervised classification based on the k_means method (matlab statistics toolbox) was applied to the 4 bands 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 comparable.

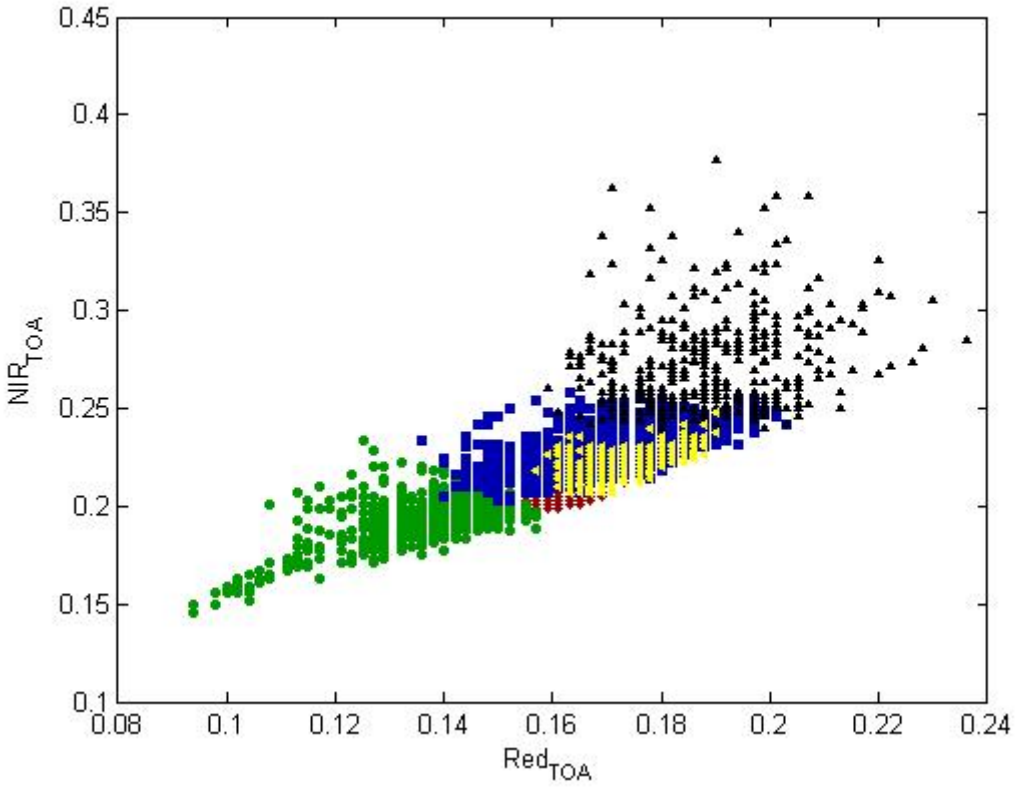

Figure 3. Red/NIR relationship on the SPOT image for Turco, 2003. Note that the colours correspond to the different classes that are described in the following.

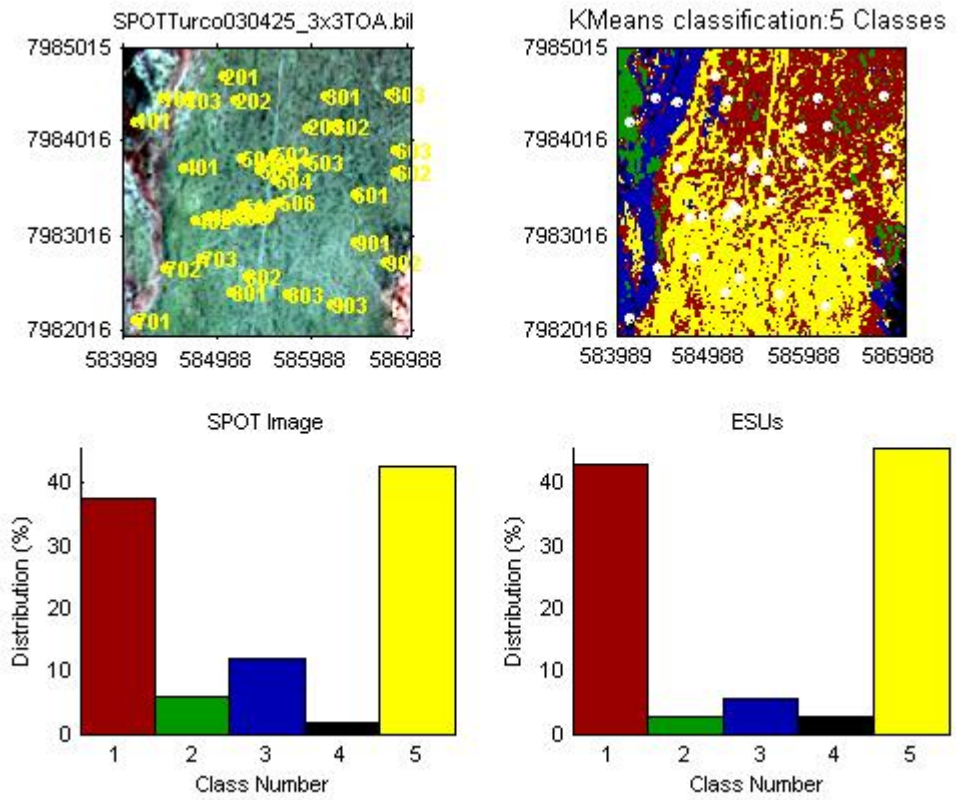

Figure 4. Classification of the SPOT image. Comparison of the class distribution between the SPOT image and sampled ESUs.

2.3 Hemispherical images

The hemispherical images were processed by the CAN-EYE software (Version 2.1, http://www.avignon.inra.fr/can_eye/) to derive the biophysical variables. Figure 5 shows the distribution of the different measured variables over the sampled ESUs. Note that the value of the biophysical variable is very low with LAI less than 0.2. As CAN_EYE V2.1, was providing the LAI at a resolution of 0.1, 3 values were available for all the ESUs. Therefore, we will use LAI57 for which the resolution was 0.01.

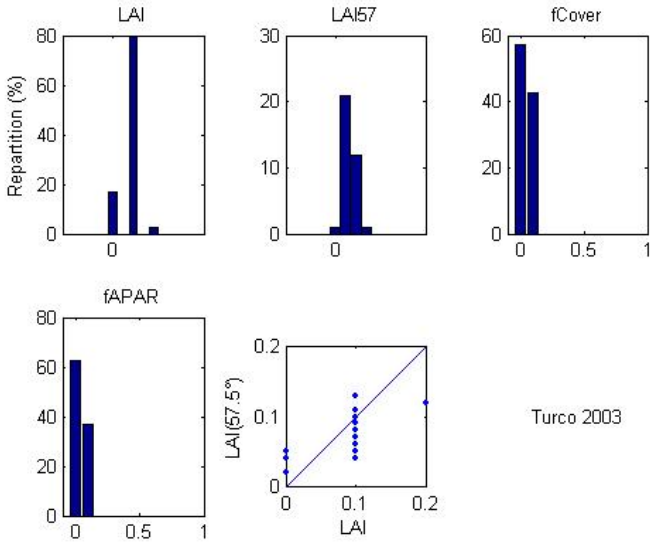

Figure 5. Turco 2003: distribution of the measured biophysical variables over the ESUs.

Figure 7 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 significant behaviours between the classes can be observed between classes 2 ,3, 4, for which a single transfer function per variable will be generated. Note that for class 1, only one ESU is available, showing low LAI value for a higher NDVI values as compared to the other classes. The hemispherical images available for this ESU (ESU 701) show that besides shrub, very low vegetation (very short grass, between green and dry) is present on the ground, which makes the distinction between soil and vegetation very difficult, and therefore CAN_EYE results questionable. Moreover, whereas the other ESUs are located in homogeneous areas (at SPOT resolution), the NDVI around ESU 701 varies between 0.15 and 0.19 (standard deviation= 0.04 although it is lower than 0.03 for the other ESUs). We have therefore decided not to use this ESU and to extrapolate the values for classes 1 and 5 by using the transfer function evaluated thanks to the other classes.

Figure 6. Turco 2003 : hemispherical image of ESU 701

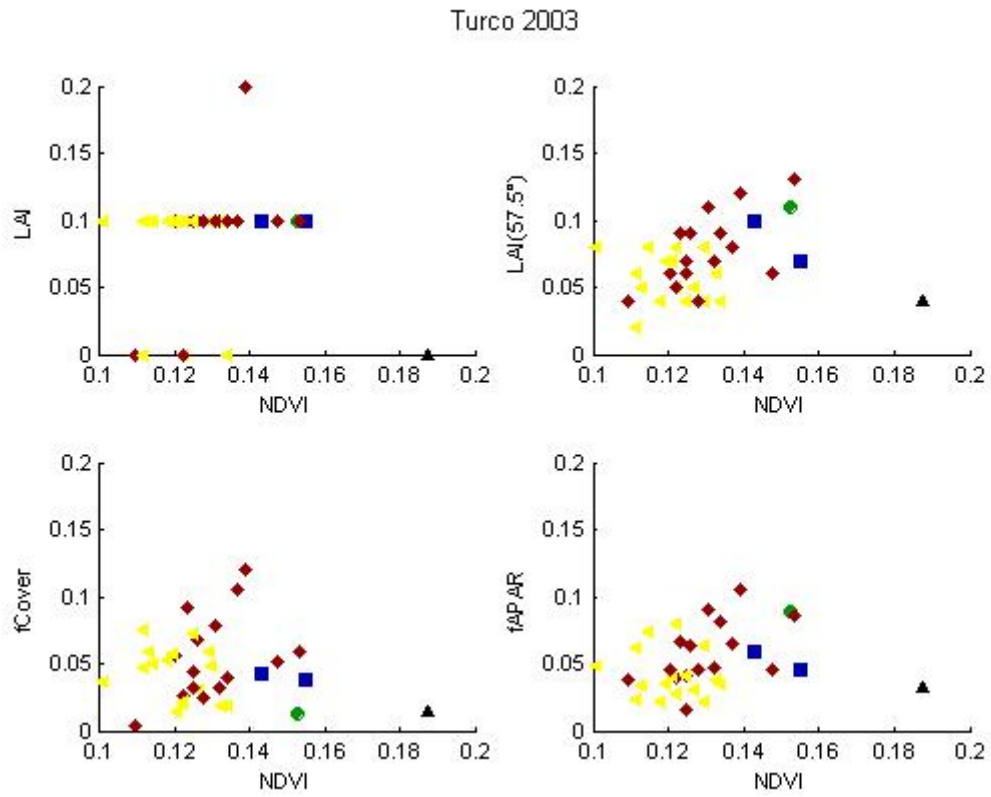

Figure 7. Turco 2003: NDVI-Biophysical Variable relationships as a function of SPOT classes

3 Determination of the transfer function for the 4 biophysical variables

3.1 Tested Transfer functions

As only one ESU was sampled for class 2 and none for class 4 (the ESU was eliminated because we were not confident in the CAN_EYE result for this particular ESU, we can consider only one type of transfer functions , REGRESSION (REG) among those generally tested for VALERI sites (Look-Up-Tables, class average): multiple robust regression between ESUs reflectance (or Single Ratio) and the considered biophysical variable is applied: 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).

Regressions are tested using either the reflectance or the logarithm of the reflectance for any band combination, plus the simple ratio, NDVI and WDVI. As the method may have poor extrapolation capacities, a flag image, based on the computation of convex hull over reflectances, is computed showing:

- Pixels inside the 'strict convex-hull': for each class, 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, 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 Turco site

3.2.1 Choice of the band combination

Erreur ! Source du renvoi introuvable. shows the results obtained for all the possible band combinations using either the reflectance or the logarithm of the reflectance. Both in terms of weighted or cross-validation, RMSE values are very close when comparing regression with the reflectance and regression with its logarithm, therefore REG on reflectance will be chosen for all the variables. Details are given hereafter for the choice of the band combination of each variable.

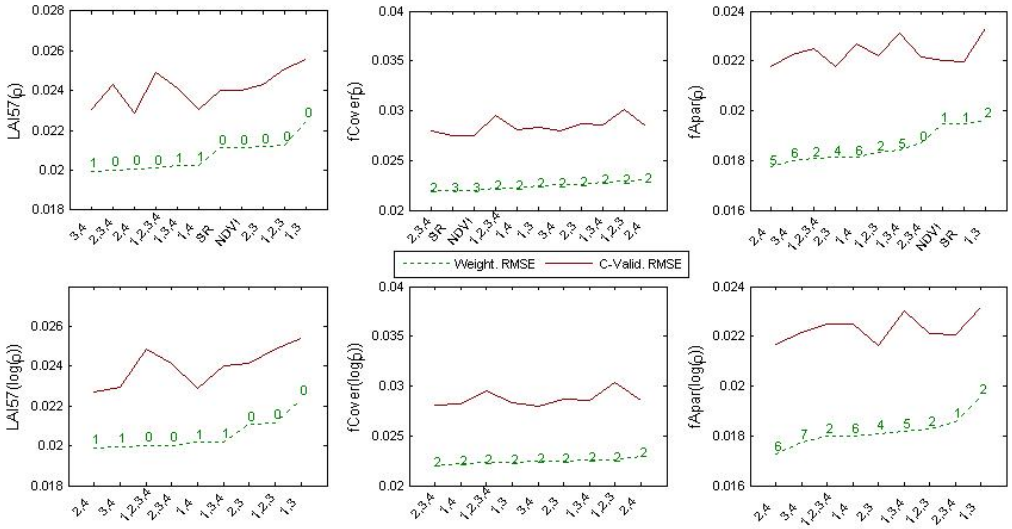

Figure 8. Transfer function: test of multiple regressions applied on different band combinations. Band combinations are given in abscissa. The estimated biophysical variable is given in ordinate. Top graphs correspond to regression made on reflectance (ρ): the weighted root mean square error (RMSE) is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

For the effective LAI at 57.5°, Figure 8 and Figure 10 show that the (XS2, XS4) band combination performs the best for LAI57 since there is no ESU with weight lower than 0.7 (Figure 9), and the cross-validation RMSE is the lowest among all the possible band combinations. The associated weighted RMSE is about 0.02 for a LAI varying between 0 and 0.13 and the cross-validation is 0.025.

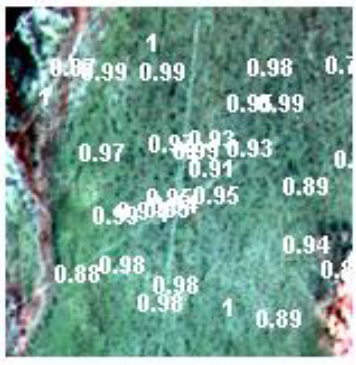

Figure 9. Weights associated to each ESU for the determination of LAI57.

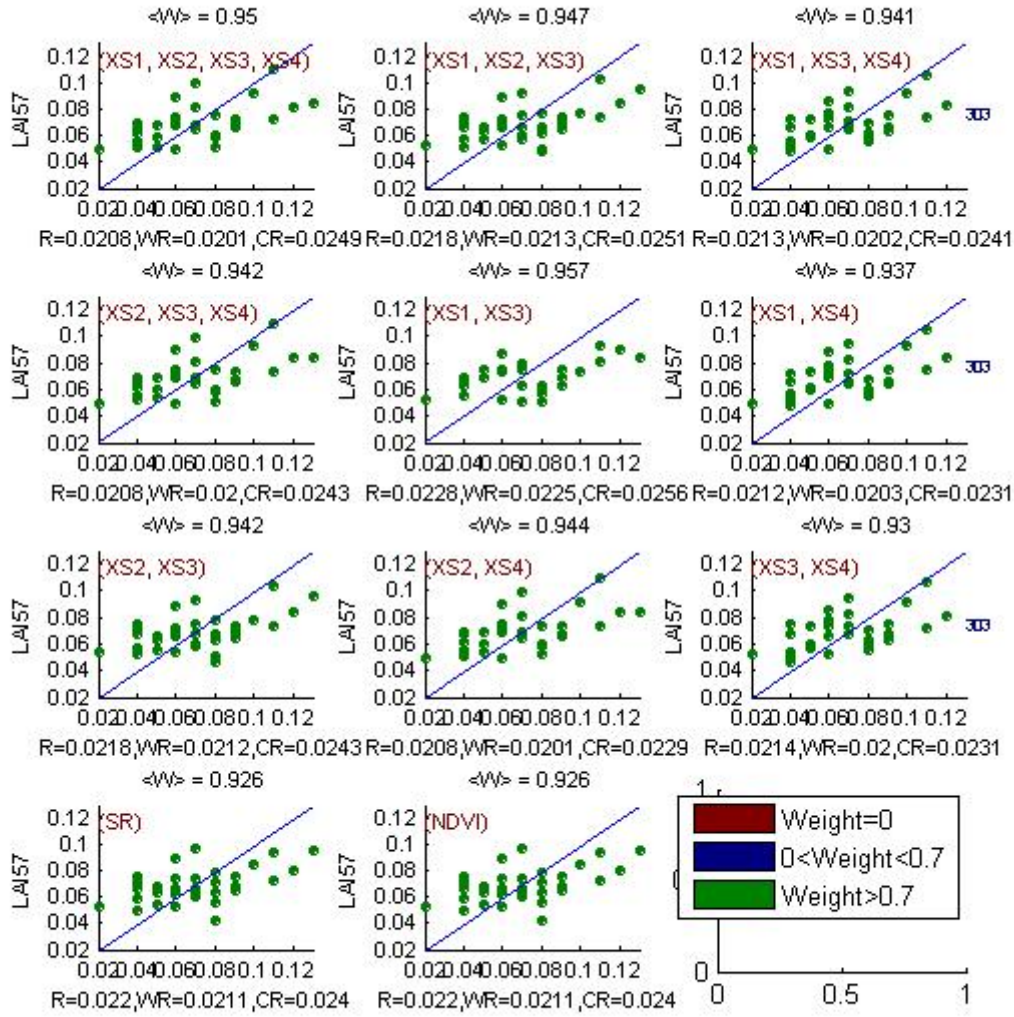

Figure 10. Effective Leaf Area Index at 57.5°: results for regression using different band combinations

For fCover (XS2, XS3, XS4), we have chosen the best combination (best weighted RMSE value) (Figure 11). However, regression results on this variable are bad, even for the best band combination, which was predictable when looking at the NDVI-fCover relationship where much scattering is observed as compared to the other variables (Figure 7).

For fAPAR (XS1, XS2, XS3, XS4), we made a compromise between RMSE values and the number of ESUs with weight lower than 0.7. Note that 2 ESUs have quite systematically a weight under 0.7 whereas this is not the case for the LAI57 and fCover: ESUs 505 (located in the centre of the image) and 902 (in the border of the area located in the south east of the image) which both belong to class 1.

Variable	Band Combination	RMSE	Weighted RMSE	C-valid RMSE
Effective LAI (57.5°)	0.0300 -1.2791XS2+ 0.9301XS4	0.021	0.020	0.023
fCover	0.0752 -1.2202 XS2 + 1.6579XS3-0.6779XS4	0.025	0.022	0.028
fAPAR	0.1022+4.0317XS1-2.4685XS2 -1.3067XS3+ 0.3242XS4	0.020	0.018	0.022

Table 1. Transfer function applied to the whole site for the different biophysical variables, and corresponding errors

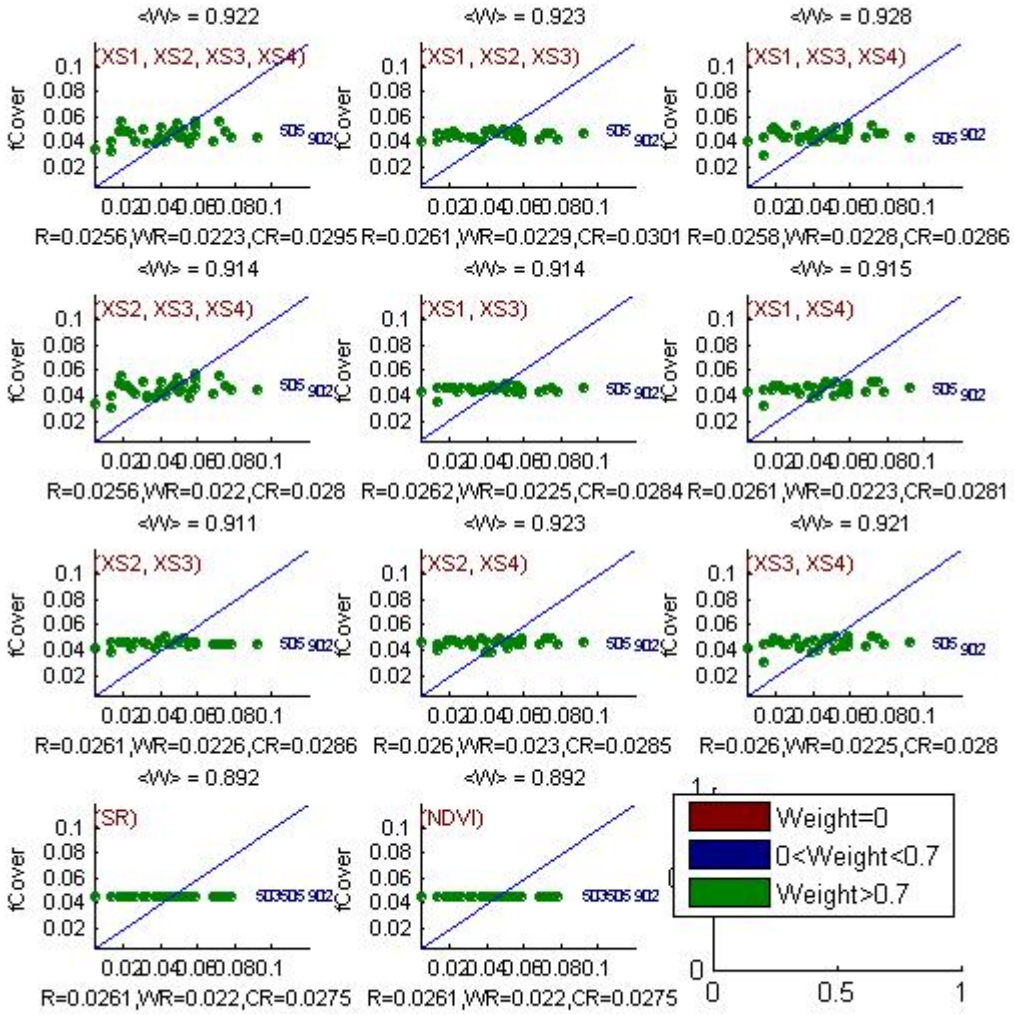

Figure 11. : fCover: results for regression using different band combinations.

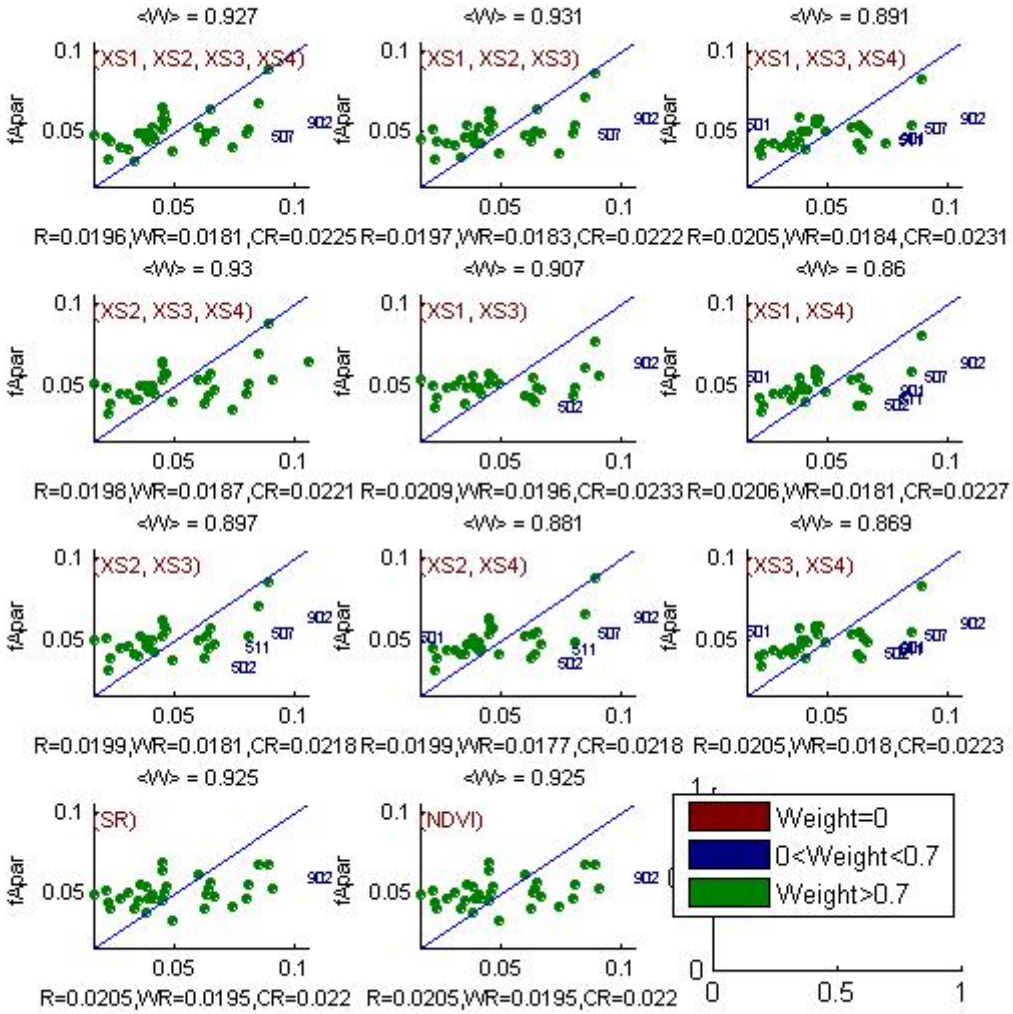

Figure 12. fAPAR : Results for regression using different band combinations

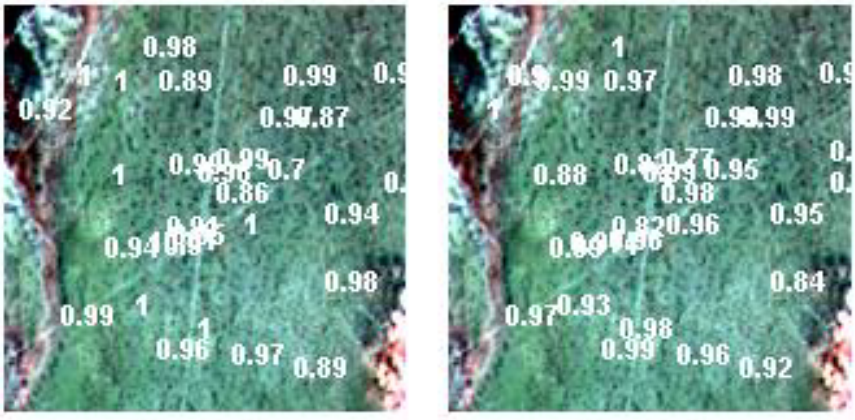

Figure 13. Weights associated to each ESU for the determination of fCover (left) and fAPAR (right) transfer function

3.3 Applying the transfer function to the Turco SPOT image extraction

Figure 14 presents the biophysical variable maps obtained with the transfer function (REG on the reflectance) described in Table 1 for all the classes. LAI57 and fCover maps obtained for the different variables are quite consistent, showing similar patterns whereas fApar map is different, showing very low values where fCover and LAI are quite high and conversely (lower right corner of the image). The most significant differences are observed in areas where flags show extrapolation (red pixels). This is confirmed when comparing the average fAPAR over the site and over the ESU. However, the values obtained for all the biophysical variables are very low and correspond to very low vegetation. Slight differences are observed between the average value of the biophysical variable over the whole site, and the average value over the ESU:

	ESU averaging (level 0)	Site Averaging (level 1)
LAI	0.05	0.07
fCover	0.05	0.05
fAPAR	0.09	0.05

Table 2. Comparison of average biophysical variable values issued from the regression and from the ESU

The flag maps are quite different between LAI57, fCover and fAPAR. This is explained by the fact that the summits of the convex-hulls are defined using 2 (LAI57) or 4 (fAPAR) bands. Therefore, the use of four bands in the regression induces a convex-hull which is stricter than when using two bands.

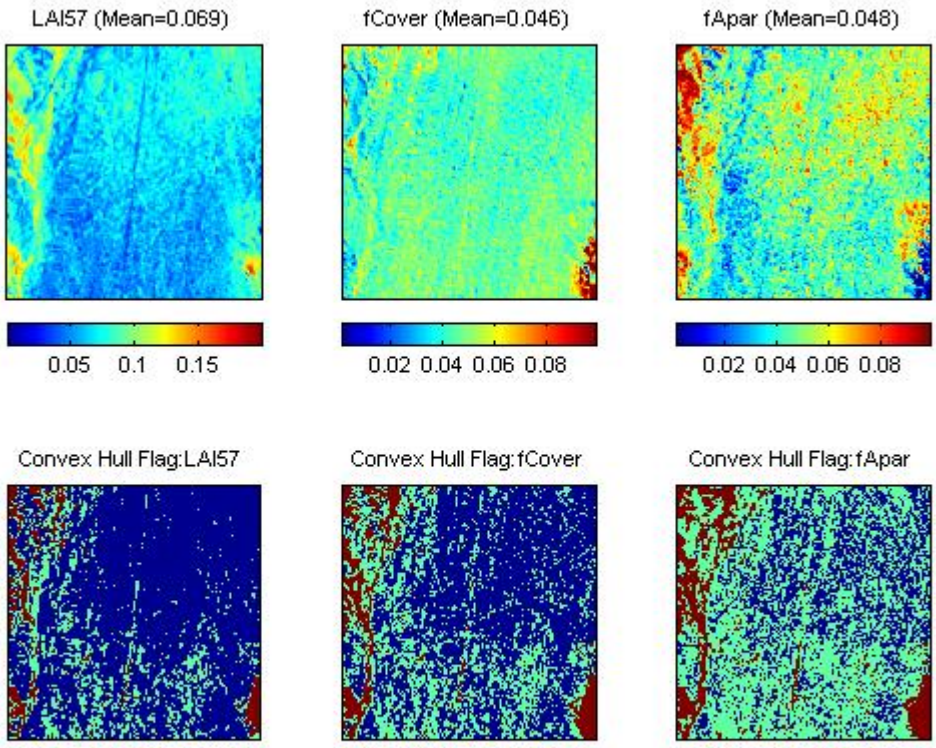

Figure 14. High resolution biophysical variable maps applied on the Turco site (top). Associated Flags are shown at the bottom: dark blue and green corresponds to the pixels belonging to the 'strict' and 'large' convex hulls, and red to the pixels for which the transfer function is extrapolating.

4 Conclusion

The transfer functions are finally obtained by using Reg on the reflectance and 34 ESUs together. The band combinations are different from one regression to another. LAI map was not produced due the use of CAN_EYE V2.1 which provides LAI at a 0.1 resolution, which is not enough for Turco (very low vegetation). Results show

quite good consistency between LAI57 and fCover variables and the flag associated to each map show that the transfer function is used as an extrapolator in reduced areas.

The biophysical variable maps are available in UTM 19S projection (WGS84)

5 Acknowledgements

We want to thank:

- Roland Bosséno (INRA-CSE) for the processing of CAN-EYE data
- All the people who have participated to the campaign: Roland Bosséno, Freddy Loza de la Cruz from ABTEMA, Justina Mollo de Villca and Rufian Villca Carex.
- The CNES who has supported the VALERI project and ISIS for the SPOT image acquisition.