VALERI 2003 : Concepcion site (Mixed Forest)

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 Concepcion 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 3km x 3km site. The processing of the ground data has shown that:

- ESUs 12 and 27 (in black on Figure 1) were located on field borders. Those two ESUs were eliminated
- Considering that SPOT geo-location and GPS measurements are associated to errors, we found that processed LAI for ESUs 5, 15, 18 did not correspond to the SPOT pixel in terms of reflectance: they have been shifted of 1 pixel, according to S. Garrigues who participated to the measurements.

Finally 26 ESUs have been kept for the computation of the transfer function. The image can be divided in two main areas: the blue one that corresponds to an area which is re-forested (young eucalyptus and pine), and the red/dark red one which corresponds to pine forest.

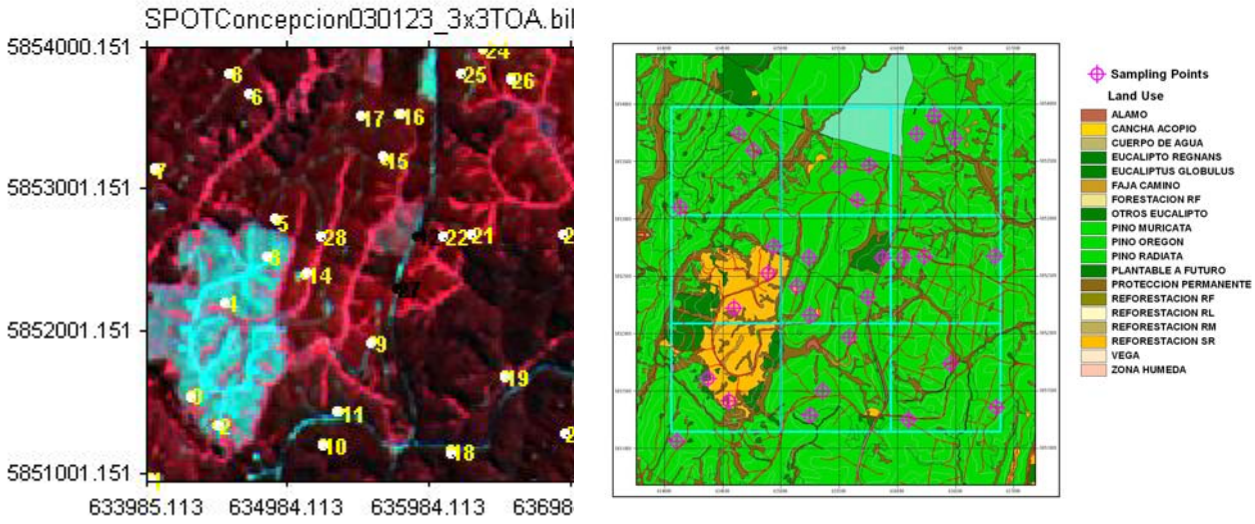

Figure 1. Distribution of the ESUs around the Concepcion site. SPOT image and land cover map (from Concepcion campaign report)

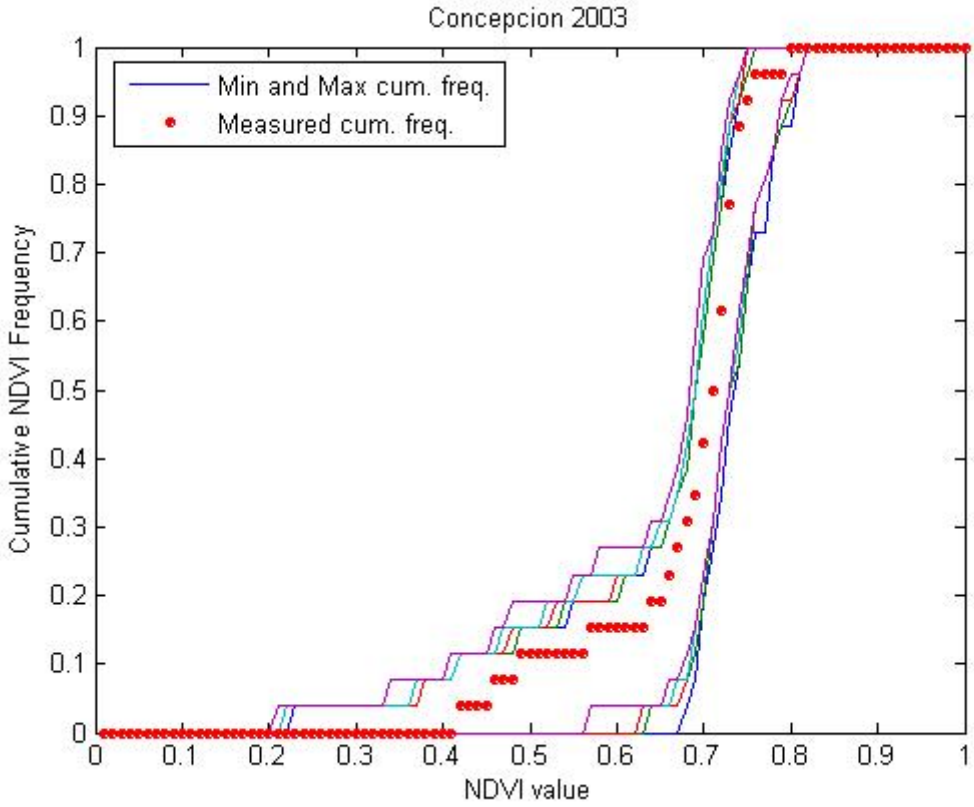

Figure 2. Comparison of the ESU NDVI distribution and the NDVI distribution over the whole image.

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 26 ESUs is very good as compared to the NDVI distribution over the whole site since the 'ESU' curve is inside the 'boundary curves'. This result induces that the biophysical variable map should be accurate since little extrapolation will be required to extend the ESU measurement to the whole site.

2.2 SPOT image

The SPOT image was acquired the 23rd january 2003 by HRVIR1 on SPOT4. It was geo-located by SPOTimage (SPOTView basic). The projection is UTM 18S, PSAD56, 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. Few points show saturation in the Near_Infrared domain for low values in the red band.

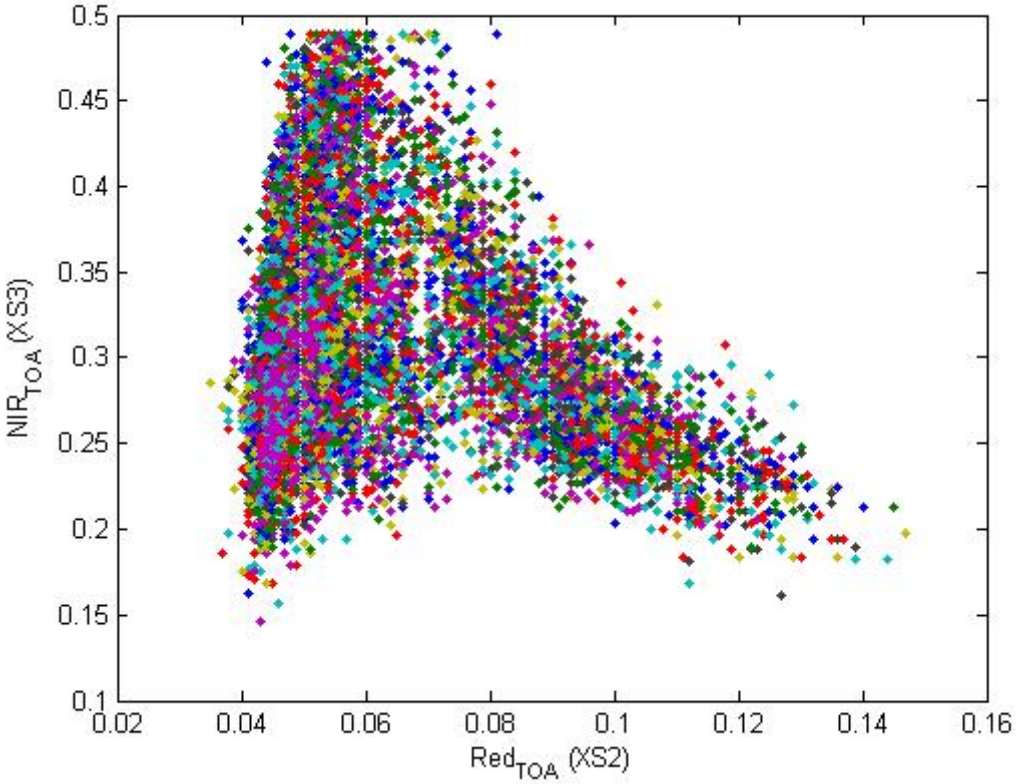

Figure 3. Red/NIR relationship on the SPOT image for Concepcion, 2003.

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. Regarding the land cover map (Figure 1), a number of 4 classes was chosen (Figure 4). The repartitions of the classes on the image and on the ESUs are very similar.

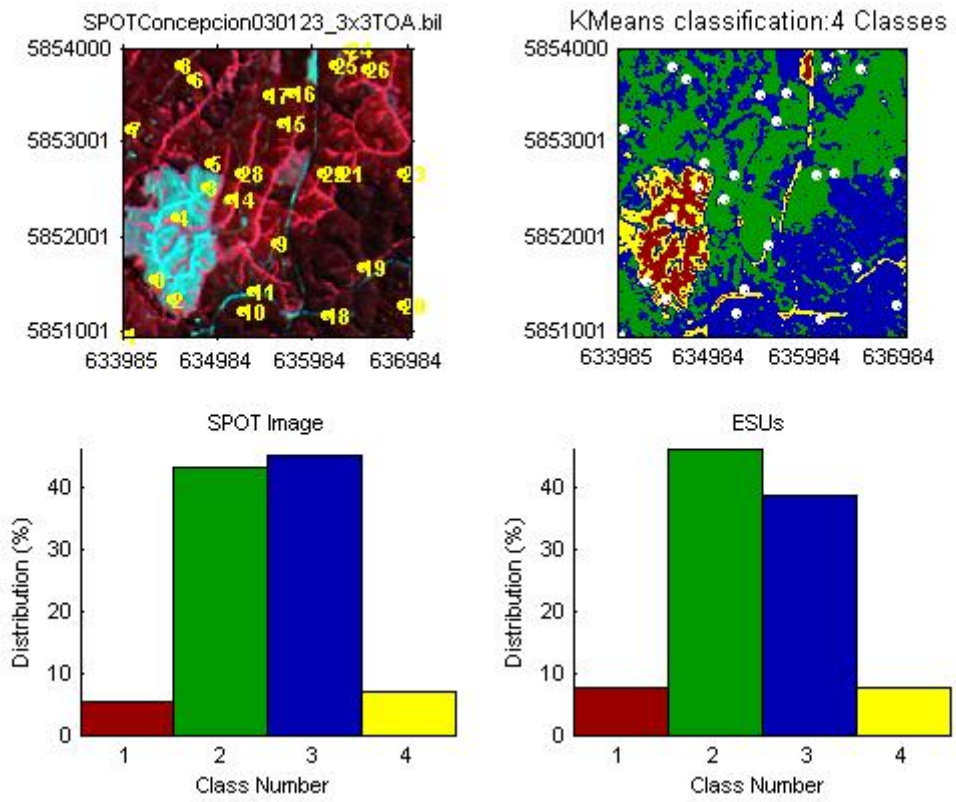

Figure 4. Classification of the SPOT image. Comparison of the class distribution between the SPOT image and sampled ESUs.

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. As there was understorey in most of the ESUs, 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, fCover, and fAPAR. The ESU biophysical variable was then computed as:

- LAI, LAI57 : LAI(above)+LAI(below).
- fCover :1-(1- fCover(above))*(1-fCover(below)). This assumes that independency of the gaps inside the understorey and the gaps inside the trees which is not true but it the only way to get the total fCover.
- fAPAR: [1-(1-fAPAR(below))*(1-fAPAR(above))], since 1-fAPAR can be considered equivalent to a gap fraction.

LAI derived from directional gap fraction and LAI derived from gap fraction at 57.5° are consistent. To build the relationships between biophysical variables and SPOT data, the reflectance of a given forest ESU (classes 2 and 4) considered as the average reflectance over the central pixel + the 8 surrounding pixels. This takes into account the fact that the height of the trees are about 30m which makes the fisheye observing an area of $\pi x[30x\tan(60°)]^2 \cong 8500m^2$, *i.e.* the area of 9 SPOT pixels (=8100m²), when using a maximum view zenith angle of 60°.

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. It appears that a single

relationship can be established for all the classes between biophysical variable and reflectances. It must be notices also that the fCover-NDVI relationship is very bad and this may be due to the fact that we assumed independency of the gap fractions between the two layers which does not seem to be the case for this site. Note that for higher zenith angle (fAPAR at 10AM), the assumption of independence appears to be verified since the fAPAR-NDVI relationship corresponds to what is expected. Therefore, in the following, we will not provide a fCover map.

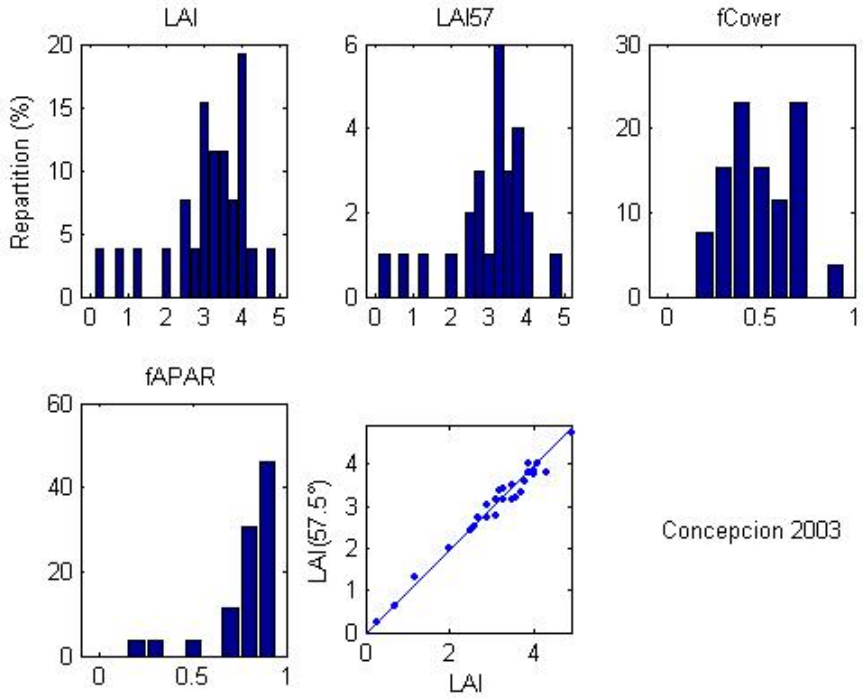

Figure 5. Distribution of the measured biophysical variables over the ESUs.

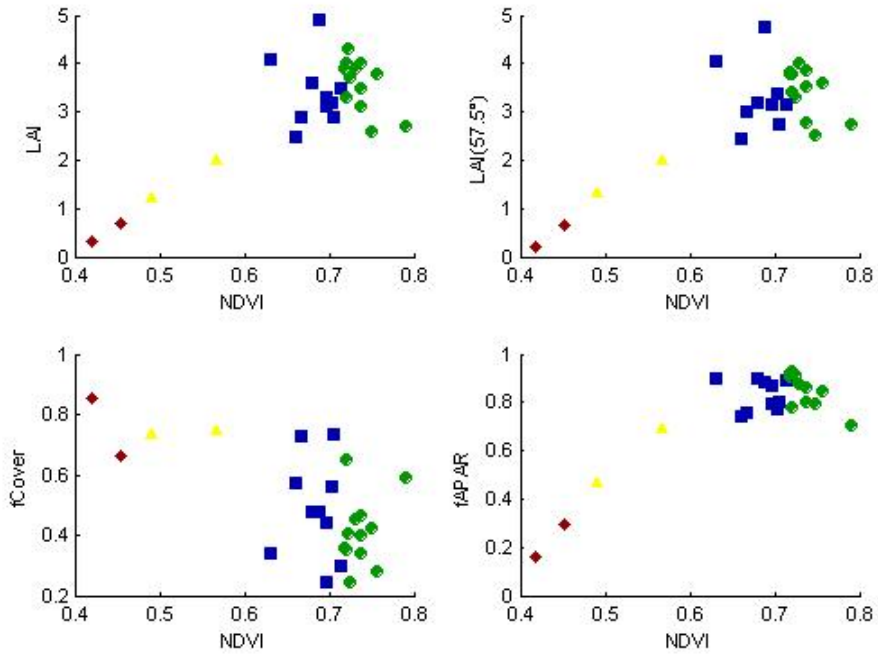

Figure 6. NDVI-Biophysical Variable relationships as a function of SPOT classes

3 Determination of the transfer function for the 3 biophysical variables: LAI, LAI57, fAPAR

3.1 Tested Transfer functions

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

- 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 Concepcion 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, NDVI and WDVI. As both methods 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 (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 Concepcion site

For all the 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

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) represents well the whole site, REG and LUT methods are both tested. The REG method provides similar results for LAI and LAI57 and much better results for fAPAR than LUT. Therefore, we will use the REG method for all the variables.

- Using either the logarithm of the reflectance or the reflectance itself provides very similar results in terms of cross-validation RMSE, for the best band combinations. 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.

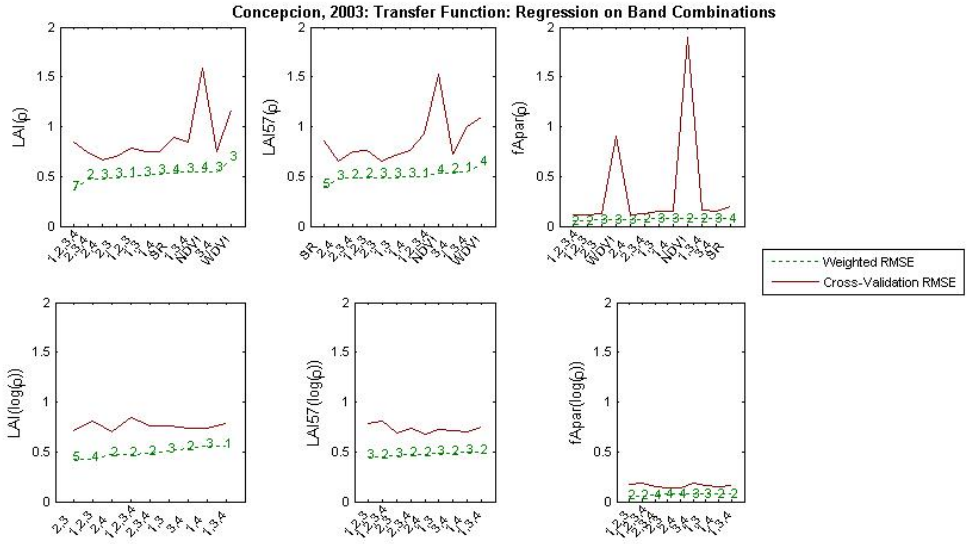

Figure 7. Transfer function: test of multiple regressions applied on different band combinations. Band combinations are given in abscissa. The estimated biophysical variable is given in ordinate. Top graphs correspond to regression made on reflectance (ρ): the weighted root mean square error (RMSE) is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

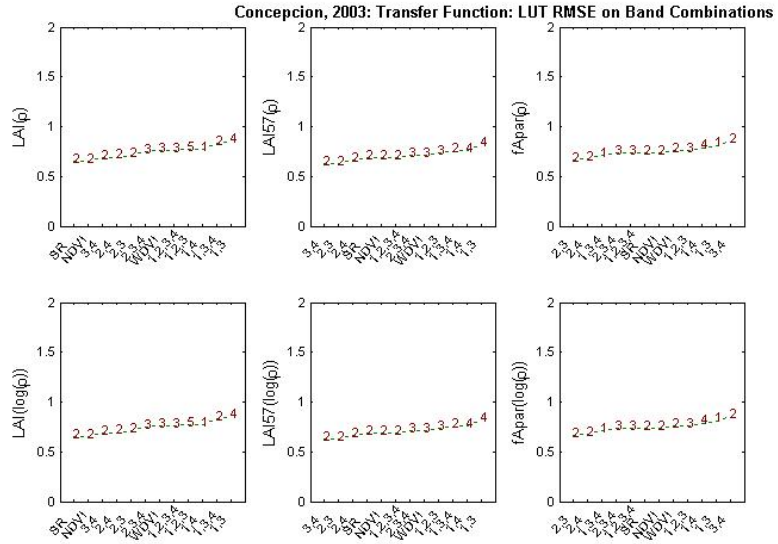

Figure 8. Transfer function: test of LUT applied on different band combinations. Band combinations are given in abscissa. The estimated biophysical variable is given in ordinate. Top graphs correspond to regression made on reflectance (ρ): the root mean square error is presented in green. The numbers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

For the effective LAI and LAI at 57.5°, Figure 9 and Figure 11 show that some ESUs have systematically a weight lower than 0.7 for almost all tested band combination: ESU 6 and 18. Although, ESU 18 was shifted of

one pixel down since it was on a border of a path, it is still rejected by the robust regression algorithm. ESU 6 corresponds to pine (pinus radiate) + understorey (quila) and the can-eye processing report, as well as the hemispherical images of the ESU do not show particular problem.
The (XS2,XS3,XS4) combination for LAI and (XS2,XS4) for LAI at 57.5° were selected since they present a good compromise between the RMSE values and the number of points with a weight lower than 0.7(Figure 10). Note that the RMSE is quite low (around 0.6 for LAI varying between 0 and 4).

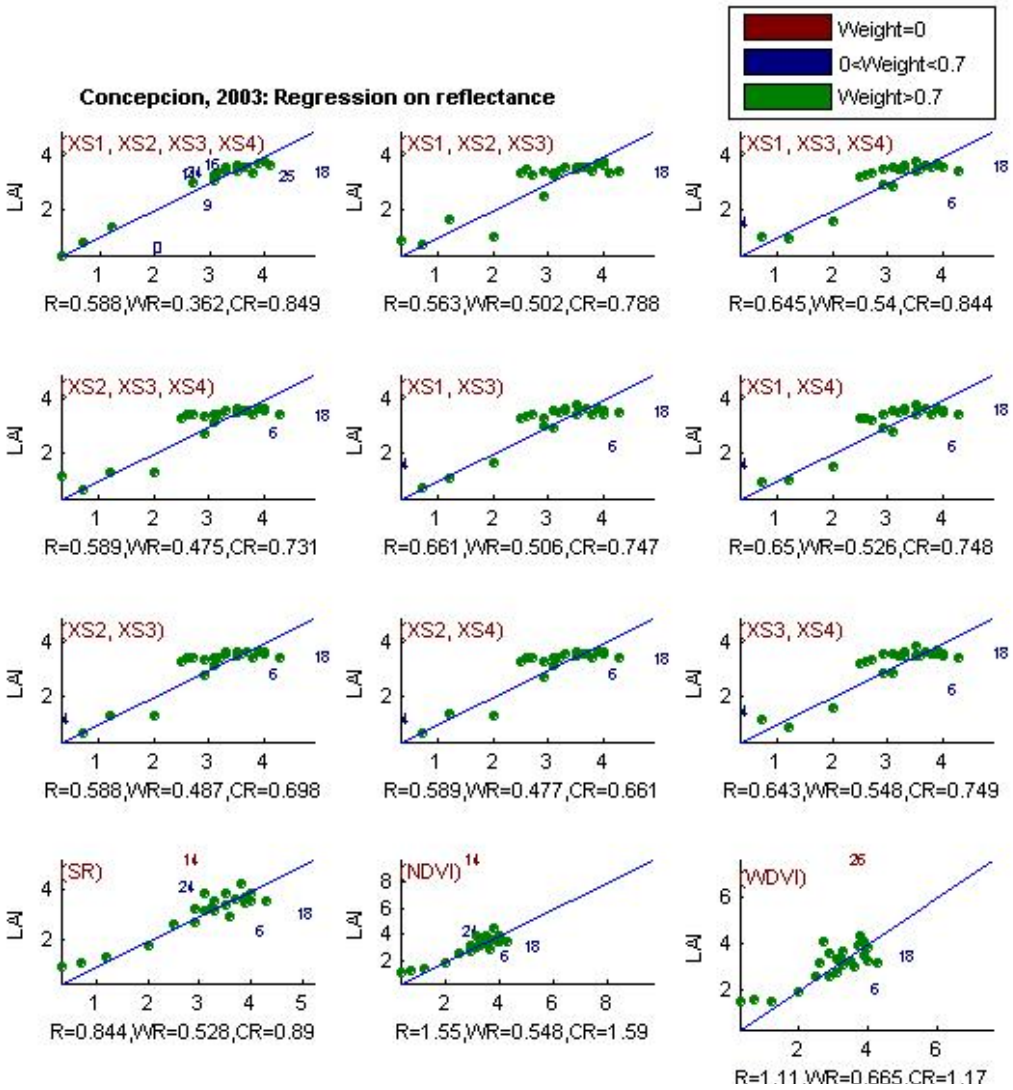

Figure 9. Effective Leaf Area Index: results for regression using different band combinations. R is the root mean square error computed between LAI and estimated LAI. WR is the weighted root mean square error and CR is the cross validation root mean square error.

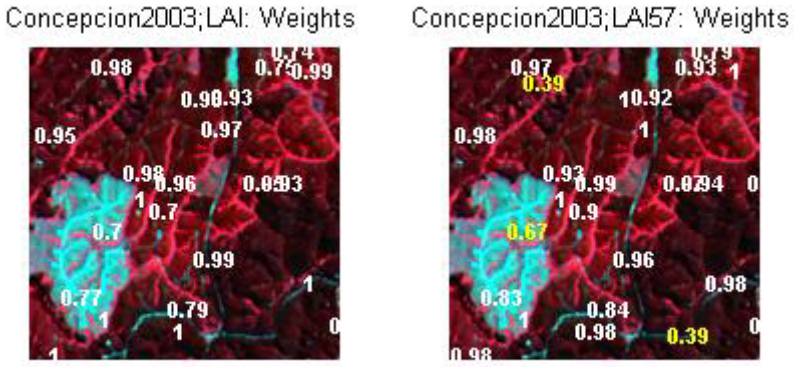

Figure 10. Weights associated to each ESU for the determination of LAI (left) and LAI57 (right) transfer function.

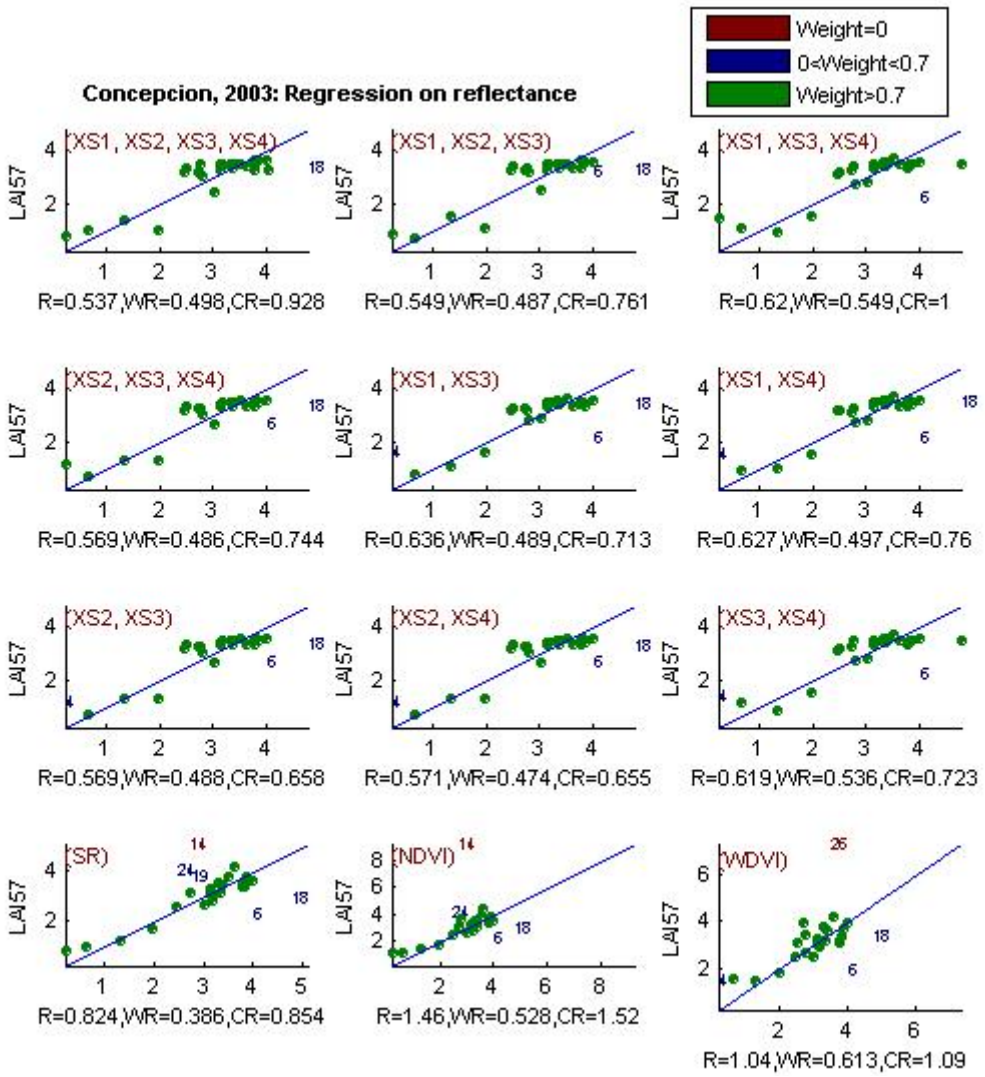

Figure 11. Effective Leaf Area Index at 57.5°: results for regression using different band combinations

For fAPAR (XS1,XS2,XS3,XS4), ESU 0 seems to show a strange behaviour in band XS2 since it is rejected when using any regression including band XS2. As for LAI, fAPAR for ESU6 is generally underestimated when using the REG. The same is observed for ESU9 but this underestimation is still reasonable.

Variable	Band Combination	RMSE	Weighted RMSE	Cross-valid RMSE
Effective LAI	5.9157-54.5899XS2+0.2344XS3+0.2423XS4	0.589	0.475	0.531
Effective LAI (57.5°)	5.6910-49.5057XS2-0.4535XS4	0.571	0.474	0.655
fAPAR	0.3605+27.4977XS1-28.8625XS2-0.3483XS3-0.7532 XS4	0.098	0.055	0.113

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

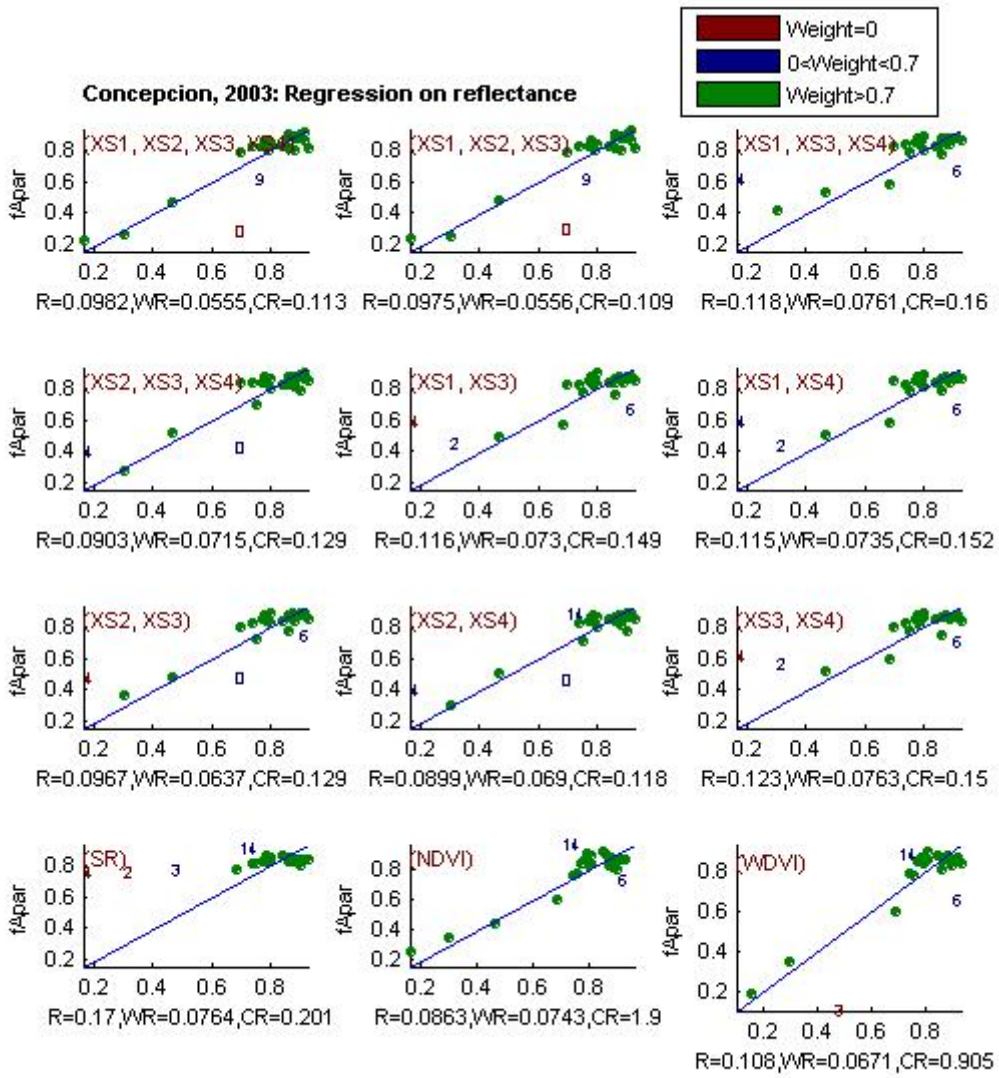

Figure 12. fAPAR : Results for regression using different band combinations

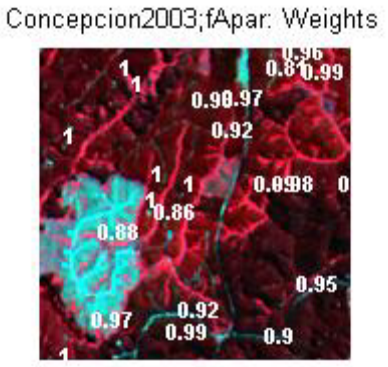

Figure 13. Weights associated to each ESU for the determination fAPAR transfer function

3.3 Applying the transfer function to the Concepcion SPOT image extraction

Figure 14 presents the biophysical variable maps obtained with the transfer function (REG) described in Table 1 for all the classes. The fCover map is not presented here since we found some problem in the ground data processing (assumption of independency between the gap fraction of trees and understorey may be not verified (see §2.3). The maps obtained for the different variables are consistent, showing similar patterns, low LAI values where low fAPAR are observed and conversely. Note that the average value for LAI and LAI57 are consistent (around 3 for both).

The flag maps show also that there is extrapolation of the transfer function all over the site, mainly in the lower left part of the image, although the test on the sampling strategy was positive (31% for LAI, 16.18% for LAI57 and 37% for fAPAR). This is mainly due to the fact that this test is performed over NDVI and that other bands than XS2 and XS3 are used for the regression. Moreover, the more bands are used for the regression, the larger is the area where extrapolation is performed. However, the estimated value of the biophysical variable in this lower left part of the image are generally low, which is consistent with the land cover map given in Figure 1.

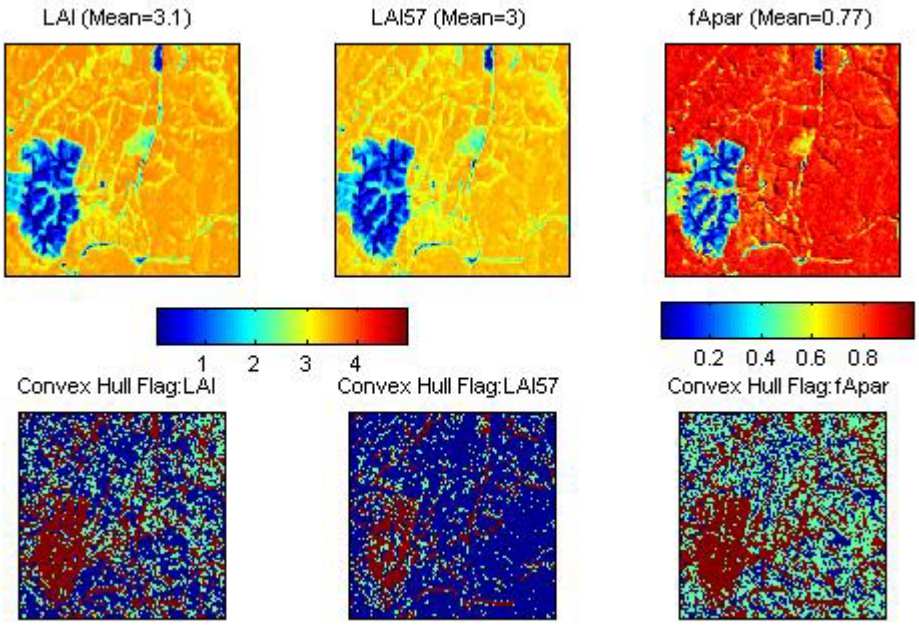

Figure 14. High resolution biophysical variable maps applied on the Concepcion site (top). Associated Flags are shown at the bottom: dark blue and green corresponds to the pixels belonging to the 'strict' and 'large' convex hulls, and red to the pixels for which the transfer function is extrapolating.

4 Conclusion

The transfer functions are finally obtained by using Reg and 26 ESUs together for all the classes. 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 available land cover map although the flag associated to each map show that the transfer function is used as an extrapolator in 30% of the cases for LAI and fAPAR.

The biophysical variable maps are available in two projections:

- Plate carrée: latitude/longitude in WGS84 at 1.7857142857e-004° resolution
- UTM, 18S, WGS84 at 20m resolution.

The transfer function was applied on the SPOT image which was geo-located in UTM18S, PSAD26. 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 all the people who participated to the campaign: Sébastien Garrigues, Gonzalo Carrasco, Bioforest headquater Eduardo Rodriguez Treskow, Christian Montes, Rodrigo Burgos Valencia, and Fernando.