MAPPING LEAF AREA INDEX MEASUREMENTS AT DIFFERENT SCALES FOR THE VALIDATION OF LARGE SWATH SATELLITE SENSORS: FIRST RESULTS OF THE VALERI PROJECT

Marie Weiss, Louis de Beaufort, Frédéric Baret, Denis Allard, Nadine Bruguier, Olivier Marloie

INRA, Bioclimatologie, Site Agroparc, Domaine Saint-Paul, 84914 Avignon, France
INRA, Biométrie, Site Agroparc, Domaine Saint-Paul, 84914 Avignon, France

ABSTRACT - *Nowadays and in the coming years, many large swath satellite sensors such as VEGETATION, POLDER, MERIS, NOAA/AVHRR, MSG, MODIS, MISR, etc ... are launched to characterize the biosphere at the global scale. Biophysical products are thus derived from satellite data, such as the leaf area index (LAI) that plays a key role in canopy functioning and soil-vegetation-atmosphere transfers. Many algorithms, generally dedicated to one sensor, have been developed to derive LAI products from reflectance data. Data are now acquired (soil measurements as well as satellite data) and processed to validate those algorithms and provide an estimate of the associated uncertainty. The MODLAND (for the EOS platform) and VALERI (for European sensors) projects both consist in the development of a network of validation sites for large swath satellite data that are representative of the Earth surface (agricultural areas, boreal or tropical forests, sparse vegetation,...).*

Considering that the lowest spatial resolution for large swath satellite sensors is 7km (POLDER), the validation sites cover 10km² areas. Exhaustive LAI measurements at this scale is a difficult task since these ground acquisitions are very local (about few meters). In the frame of the VALERI project, a spatial sampling strategy is developed to allow the extension of local measurements to the whole area observed by the sensors, using high spatial resolution data (SPOT). The aim of this study is to provide a high spatial resolution LAI map for the whole 10km^2 area from a selection of ground measurements, as well as the associated error of estimation. We compare three geostatistical methods: (1) LAI ground measurement ordinary kriging, (2) kriging of the error between measured LAI and estimated LAI from SPOT data and (3) co-located kriging. In this study, both LAI and associated error of estimation maps are compared using these three geostatistical methods. One validation site corresponding to an agricultural area was considered here. The high spatial resolution LAI maps will then allow to derive low spatial resolution maps (from 250m for MERIS to 7km for POLDER) to evaluate the algorithms of LAI estimation from reflectance data.

1 - INTRODUCTION

In the very next years, a range of large swath satellite data such as AVHRR, VEGETATION, POLDER, SeaWifs, MERIS, AATSR, MSG, MODIS, MISR will concurrently fly over the Earth surface. These sensors have potential to estimate the key biophysical variables, such as leaf area index, used as inputs, in both Soil-Vegetation-Atmosphere Transfer models and Net Primary Production models. Up to now, large swath sensors were used to produce global maps of these biophysical products through empirical relationships derived from spectral indices. However, these relationships provide generally poor accuracy. This is the reason why new algorithms, based on physical model inversion techniques such as look-up-tables or neural networks, were developed in synergy with sensor technology improvements. Considering the complexity of the biophysical

processes involved in the satellite signal, the only way to ensure a proper validation of biophysical products is to perform an actual direct validation of these algorithms by comparing *LAI* ground measurements to estimates from actual satellite data [Justice 2000].
The objective of the VALERI (VALidation of European Remote sensing Instruments) project is to develop a network of validation sites over the Earth surface and a methodology designed to directly measure the biophysical variables of interest at a proper spatial and temporal scales. This allows the evaluation of existing algorithms as well as the intercomparison between sensors and algorithms. The main difficulty in the validation of large swath sensors products is to provide ground measurements at low spatial resolution (typically from 250m for MERIS or MODIS to 7km for POLDER), since ground measurements are generally local and represents few meters. It is therefore not feasible to exhaustively sample the whole area corresponding to one large swath pixel. The aim of this study is thus to define a ground sampling strategy based on the combination of ground measurements and high spatial resolution remote sensing data. It can be divided into four steps:

1. Segmentation of a high spatial resolution image (SPOT, TM) of the validation site, allowing to identify the main class types in the image to design a ground measurement sampling strategy for each class.
2. Ground measurements of the leaf area index. A maximum set of 50 high spatial resolution pixel is sampled compatibly with manpower resources and time.
3. Spatial extension of the ground measurements to the whole site thanks to the high spatial resolution image.
4. Evaluation of the leaf area index at the large swath sensor resolution.

The first year of the VALERI project was devoted to the refinement of the methodological aspects. For this purpose, four sites have been sampled:

- Romilly an agricultural area in the north of France,
- Järvselja, a mixed forest in Estonia,
- Nezer Forest in the south west of France made of pine trees, and
- Gourma, a semi-desertic area in Mali, with grass and very sparse tree cover.

However, this preliminary study will concentrate on the Romilly site only. In the same way, lthough VALERI is dedicated to *LAI*, fAPAR, fCover, albedo and Chlorophyll content products, the present study is restricted to *LAI* for the demonstration.

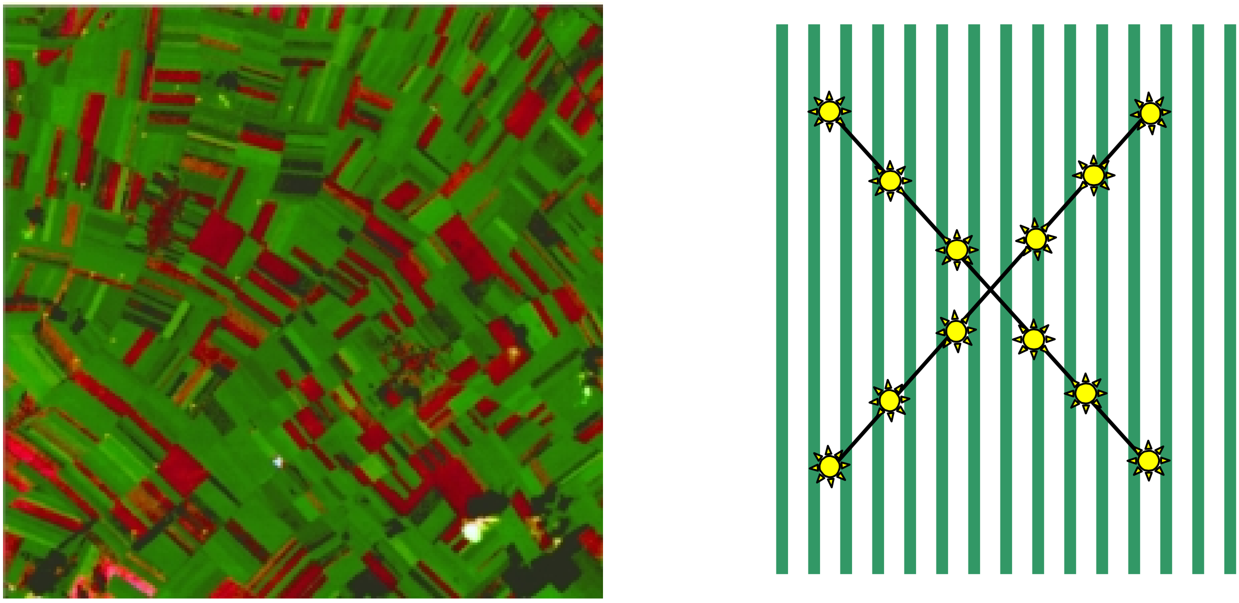

Figure 1. left: SPOT image of the Romilly VALERI site (Red : XS3, Green: XS1, Blue: XS2). Right: the local LAI2000 measurements (the stars) organized along a 20m long cross. The vertical lines represent the row direction.

2 - MATERIEL AND METHODS

The 10km by 10km Romilly site is an agricultural area composed of winter crops (wheat, barley, pea, alfalfa, rape-seed, hemp, poppy) and summer crops (maize, sugar beet, sunflower, potato) and few urban and woody zones (figure 1).
A SPOT-4 image acquired in May 2000 allowed to define a simple sampling strategy by dividing the area into 49 1.4km-size cells. In each sub-area, at least one field was selected for ground measurements. The field was chosen to characterize the main crop in the cell, and also to approach the proportion of the considered crop within the whole area. Except for rape-seed, 12 local cross-shaped measurements of LAI2000 were made in the 49 fields corresponding to the 49 cells (figure 1). Each measurement is composed of one above and four below canopy measurements. The measurements are distant from 4m to each other, leading to a 20m long cross. In case or row crops, the 20m cross is always placed diagonally to the rows. The *GPS* positioning is achieved at the center of the cross, with 1m accuracy in 95% of the points. The measurements were performed at more than 40m from the field edges, except for rape-seed crops that were too dense to penetrate inside. For rape-seed, measurements have been acquired on the field edge (no cross shape).

3 - STATISTICAL AND GEOSTATISTICAL DATA ANALYSIS

A second SPOT-1 image has been acquired in June 2000. Due to some problem during the acquisition, we could only get half the whole area. In this study, we directly work on the digital counts of the images, and thus, the image is not corrected from atmospheric effects. Woody areas (that have not been sampled) and urban zones are masked. The statistical analysis of the image shows classical results. Red and green bands are strongly correlated (correlation coefficient of 0.96) since those bands are quite close. Correlation is lower between visible and near infrared domains (-0.41 for red and –0.19 for green band). This is confirmed by principal component analysis performed on the whole image (Table 1). For the first axis, weights corresponding to the red and green bands are quite similar and negative while positive and smaller for the near infrared. The second axis shows the reverse situation. The two first axes explain more than 99% of the variability of the digital counts, which shows high redundancy of the information in the visible.

	Inertia (%)	NIR	Red	Green
Axe 1	70	0.37	-0.68	-0.64
Axe 2	29	0.91	0.12	0.39
Axe 3	1	0.19	0.72	-0.66

Table 1. Principal component analysis between SPOT bands on Romilly site.

To better describe the spatial structure of the image and the measurements, experimental variograms were also computed.
For the Romilly SPOT image, the isotropic variogram showed no trend differences as compared to the directional ones (along raw or along column). To reduce computation time, we thus decided to consider only directional, direct and crossed variograms (equation 3.1)

$$\gamma_{ij}(h) = \frac{1}{2N(h)} \sum_{N(h)} \left[Z_i(x+h) - Z_i(x)\right]\left[Z_j(x+h) - Z_j(x)\right] \qquad (3.1)$$

where Z_i is the digital count in the band *i*, *h* is the distance between two pixels of the image, and *N(h)* the number of couples (x,$x+h$) in the image.

Results show that the direct variograms are not very different between bands, except the sill value, which is higher for bands with high-level digital counts. The within field variability that can be approximated by the variogram value for short distances is much lower than the between field

variability (see figure 2 for an example with the near infrared). The crossed variograms between visible and near infrared bands, did not show particular patterns. Conversely, the crossed variogram between red and green bands (figure 2) show correlations for short distances (<20 pixels).
The *LAI* experimental variogram was first computed on the elementary 49x12 points measured within the crosses. This allowed the estimation of a spherical covariance model (range: 80m, sill:0.72m, and nugget:0.16), which was used to estimate the mean *LAI* for each cross by ordinary kriging. The correlation coefficient between this estimation and the actual mean *LAI* value is 0.98. As there is no significant difference, we decided to use only the mean *LAI* values computed over the 12 measurements of each cross.

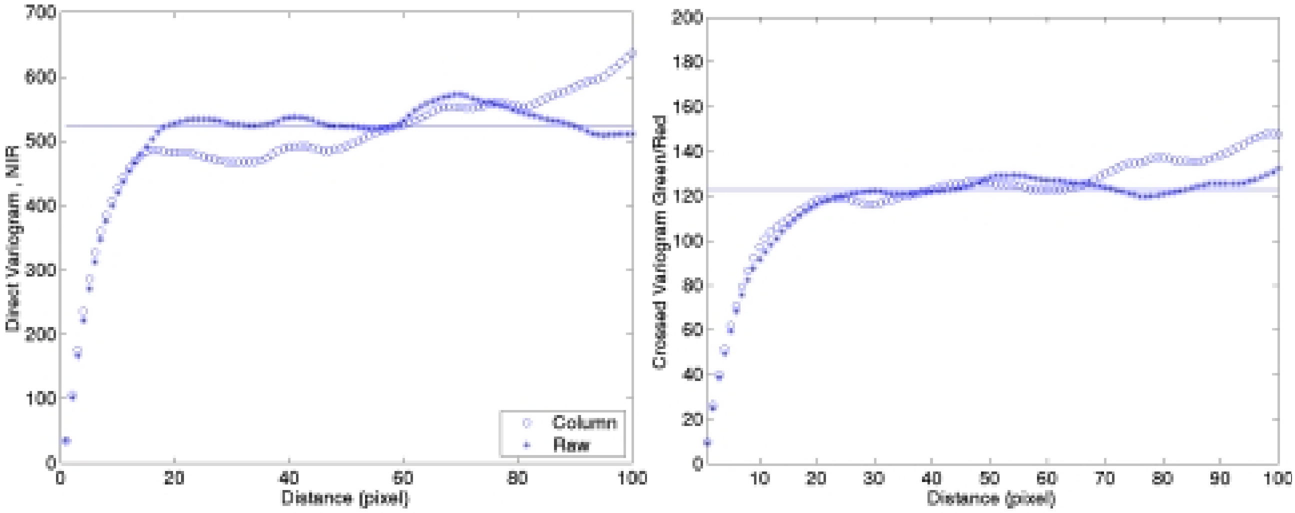

Figure 2. Experimental directional direct (left) and crossed (right) variograms computed for Romilly The solid line corresponds to the experimental variance (or covariance) value.

4 - GEOSTATISTICAL METHODS TO DERIVE *LAI* MAP

To derive *LAI* maps from few ground measurements, three geostatistical methods [Wackernagel 95] were tested: ordinary kriging, kriging the residual and co-kriging. This was investigated over a 3km by 3km sub-area within the 10 by 10 original site because of the shift of SPOT image from the initial site.

4.1 - Ordinary kriging

The first step was to model the spatial covariance, using *LAI* measurements. As the 49 *LAI* values measured was insufficient to directly compute the variogram, we approximated the *LAI* variogram from a combination of the SPOT image and the *LAI* measurements. For this purpose, a supervised classification is primarily applied to the whole SPOT image (7 classes). Each class was associated to the average *LAI* value measured within this class. Variograms computed on this synthetic image show a range at 300m. Results show that the variogram can be modeled as the sum of two spherical models corresponding respectively to short distances (<300m, representing the discontinuities between fields) and longer distances (300m<d<2000m). This variogram is finally used to estimates the *LAI* map by ordinary kriging of the *LAI* measurements.

4.2 - Kriging the residual.

Different relationships between SPOT data and *LAI* have been considered (exponential, linear) and the best result is obtained using a simple multi-linear regression (Eq 4.21). This relationship is valid for canopies, i.e. for *LAI* > 0.0. *LAI* is set to 0.0 elsewhere (bare soil, buildings, ….).

$$LAI^{*}=0.59+0.009Red+0.028NIR-0.026Green,\mathrm{RMSE}=0.86 \tag{4.21}$$

The residual error ε^* is the difference between *LAI* and LAI^* estimated using eq. (4.21). This error is known at the 49 cross measurement points:

$$\varepsilon^*(v_i) = LAI^*(v_i) - LAI(v_i) \quad (4.22)$$

where LAI^* is given by equation 4.21. The spatial covariance model of the error $\varepsilon^*(v_i)$, was determined assuming that the residual error is the sum of 3 independent errors:

(1) a spatial component (linked to soil influence for example),

(2) a component linked to the measurement itself (LAI2000 instrument, raw effect) which remains the same inside a vegetation class

(3) a component linked to the error of the *LAI* estimation by the average *LAI* measured on the 12 points of the cross.

We finally applied ordinary kriging to $\varepsilon^*(v_i)$.

4.3 - Colocated kriging.

Because only few ground measurements are available, prior information, such as SPOT data, would be welcome to improve the estimation. The co-located kriging has then been applied by using *LAI* as the principal variable and estimated LAI^* as the secondary variable. The problem was to model the crossed covariance between *LAI* and LAI^* as the secondary variable. This was achieved using a co-regionalisation linear model [Goovaerts 97].

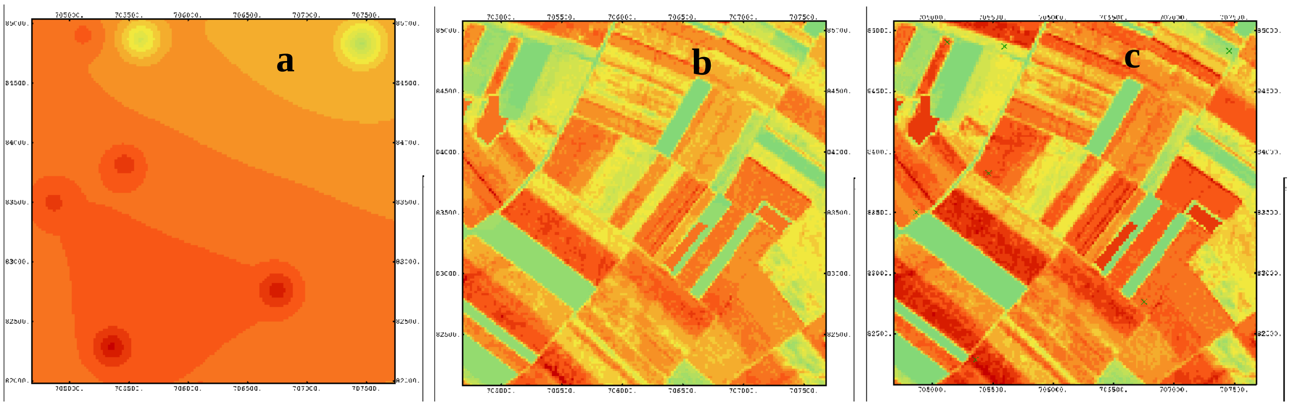

Figure 3. *LAI* maps over the 3*3km² obtained using ordinary kriging (a), kriging the residual (b) and co-located kriging (c). *LAI* values ranging from green (low) to red (high).

4.4 - Results

Results for the 3 methods are presented in figure 3. Ordinary kriging is not recommended since it relies only on few ground measurements (7 for this 3km by 3km area) that are certainly not representative of the whole Romilly area. It was used here just as a base line method. The two other methods introduce *a priori* information, using high spatial resolution satellite data and require the knowledge of the crossed covariance model. In the case of co-located kriging, kriging weights for the principal variable (*LAI*) is quite low (10%), showing that the spatial information provided by the SPOT image is prevalent. This is not surprising according to the low density of ground *LAI* measurements. The main interest in residual kriging is that the relationship f^* between *LAI* and reflectance can be non linear. In the Romilly case, the interest of residual kriging is not outlined since it has been shown that linear regression provides the best fitting.

5 - CONCLUSION

The final step of this study was to provide a *LAI* map at 1km resolution characteristic of the VEGETATION sensor, using the 3 previously described geostatistical methods. The low spatial resolution estimates of *LAI* was computed by aggregation of the high spatial resolution estimates using the arithmetic average. For the first method, block kriging was used to provide the 1km *LAI* map as well as the kriging variance. The computation of kriging variance for the two other methods is much more complex and details can be found in [de Beaufort 2000]. Results show little difference between co-located and residual kriging in terms of *LAI* and an overestimation for ordinary kriging as compared to the other methods (figure 4). If we consider the kriging variances, the most interesting method seems to be the co-located kriging.

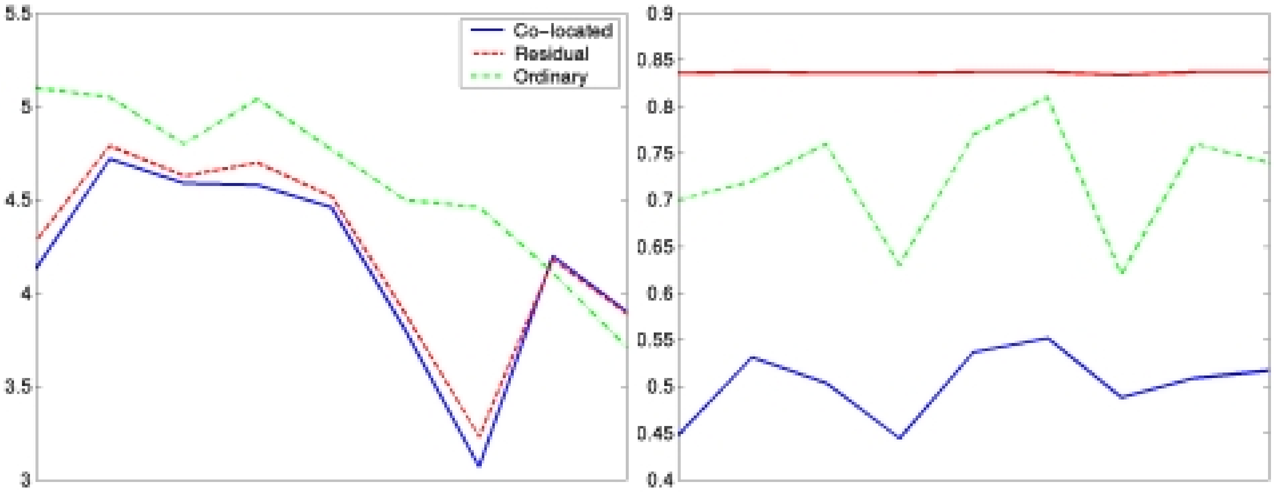

Figure 4. On the left, *LAI* aggregation for the 9 VEGETATION pixels constituting the 3km by 3km sub-area. On the right, kriging variance for the three methods.

We thus propose the co-located kriging as the most efficient method. However, in the frame of the VALERI project, different types of canopies have to be considered. It is therefore necessary to test the 3 methods on the other sites. We also propose to reduce the 10km by 10km areas to 3km by 3km to allow better sampling and be more confident in the kriging methods.

H. Wackernagel : “Multivariate geostatistics, an introduction with applications”, Springer Verlag, 1995, 257pp.

P. Goovaerts : “Geostatistics for natural resources evaluation.Applied geostatistics series”, Oxford university press, 1997.

C. Justice, A. Belward, J. Morisette, P. Lewis, J. Privette, and F. Baret, "Developments in the validation of satellite sensor products for the study of the land surface". International Journal of Remote Sensing v. 21(17):. 2000.

L. de Beaufort : "Définition d'une méthode de cartographie d'indice foliaire destinée à la validation des produits de capteurs à large champ",. Internal report, INRA Bioclimatologie, Avignon. 2000.