

Retrieving crops Green Area Index from high temporal and spatial resolution remote sensing data

Amanda Veloso¹
Valérie Demarez¹
Eric Ceschia¹

¹ Centre d'Etudes Spatiales de la Biosphère (CESBIO)
18 av. Edouard Belin, bpi 2801, 31401 Toulouse cedex 4, France
{amanda.veloso, valerie.demarez, eric.ceschia}@cesbio.cnes.fr

Abstract. This paper aims at firstly evaluating the correspondence between Normalized Difference Vegetation Index (NDVI) products from Formosat-2 (F2) and SPOT sensors and then to perform a comparative analysis of two methods for retrieving Green Area Index from high spatial and temporal resolution satellite data (F2 and SPOT). For this purpose, an empirical approach using NDVI plus field data and a Neural Network approach using the PROSAIL model are compared over four different crops: wheat, sunflower, maize and soybean. The performance of both methods were evaluated and compared with *in-situ* direct (destructive) and indirect (hemispherical photos) measurements. Results suggest better performances for the empirical approach (R^2 , RMSE). Still the physically-based method leads to good results (R^2 , RMSE). The latter seems to be more promising due to its portability and independence from field measurements.

Keywords: Green Area Index (GAI), Normalized Difference Vegetation Index (NDVI), Formosat-2, SPOT, Crops.

1. Introduction

Green Area Index (GAI) is an important vegetation variable used for different ecological, agronomical and meteorological applications. In agriculture and crop monitoring domains, GAI has a key role linking Earth observation imagery with crop models.

The estimation of biophysical variables such as GAI and also Fraction of Absorbed Photosynthetically Active Radiation (Fapar) and Fraction of Vegetation Coverage (Fcover) from remote sensing observations have been studied since the first space missions (e.g. Landsat in 1972). With the advances in technology, new missions such as Sentinel-2 [Martimor et al. (2007)] and Venus [Dedieu et al. (2006)] will provide new perspectives for land surfaces monitoring. The data from these satellites will combine both the high spatial resolution and high revisit frequency. These characteristics are essential for an accurate cartography and modeling of croplands, which have often small area and frequent temporal changes (due to plant development stages and management practices).

For this study, the need of continuous time series of high resolution images led to the combined use of data from Formosat-2 and Spot satellites. For this reason, the compatibility between these two datasets was investigated, aiming at having a complementary and homogeneous set of remote sensing data.

Many methods have been developed to relate GAI to optical remote sensing signal [Baret and Buis (2008)]. The main objective of this work is to perform a comparative analysis of two methods for retrieving GAI from high spatial resolution and high temporal frequency satellite data. The first approach consisted in an empirical method, linking NDVI time series with ground measurements. The second approach was based on the application of neural network inversion of a radiative transfer model (PROSAIL [Baret et al. (1992)]). Further studies will integrate the estimated GAI (from the best performing method) into a simple crop model for estimating biomass and yield over complete cultural cycles.

2. Data

2.1. Site description

The study was carried out in the southwest France, within a 24 km x 24km² area (1°10'E, 43°27'N). The climate is temperate mild, characterized by warm and dry summers, sunny autumns, soft winters and by rainy and stormy springs. Annual precipitation is about 656 mm and annual mean temperature is 13°C. The study area is mainly covered by arable lands (around 60%).

2.2. Field data

The study was carried out from 2006 to 2010, in the area around two 3km × 3km experimental sites (called Lamasquère and Auradé [Béziat et al. (2009)], and where meteorological measurements and flux studies have been performed). Destructive and non-destructive measurements of GAI were done over four types of crops: winter wheat, sunflower, maize and soybean.

Each non-destructive GAI value was estimated from digital hemispherical photographs (DHP) taken using the VALERI protocol [<http://w3.avignon.inra.fr/valeri/>]. An amount of 13 DHPs were taken from above the canopy over a 20m × 20m sampled area called Elementary Sampling Unit (ESU). The photographs were taken with a Nikon Coolpix 8400 camera having FC-8 fisheye lens. The DHPs were processed using the imaging software CAN-EYE V.5.1 [<http://www4.paca.inra.fr/can-eye>], providing estimates of *effective* and *true* GAI. The latter takes into account the clumping effect [Demarez et al. (2008)]. GAI was also measured using a destructive method by cutting the entire plant and measuring the surface of the green organs (leaves and stem) with a planimeter (LI3100, LiCor, Lincoln, NE, USA).

2.3. Remote sensing data

2.3.1. Formosat-2 & Spot data

Formosat-2 (F2) satellite provides images with spatial resolution of 8 m in four reflective bands centered at 488, 555, 650 and 830 nm. The sensor has a footprint of 24 km × 24 km and an orbital cycle of one day. Images are taken at near constant viewing angles (around 29°). A set of 112 images is available from 2006 to 2010 over the study area.

The SPOT images were taken by SPOT 2, 4 and 5 satellites. Data have spatial resolution of 20m (SPOT 2 and 4) or 10m (SPOT 5) in the green, red and near infrared spectral bands for SPOT 2 plus short-wave infrared band for SPOT 4 and 5. The field of view is 60 km. The viewing angles varied between +/- 27°. A total of 52 SPOT images were acquired over the five years period.

2.3.2. Remote sensing data pre-processing

SPOT and FORMOSAT images were processed with the KALIDEOS processing chain [<http://kalideos.cnes.fr>]. It provides valuable atmospheric, radiometric and geometric corrections required for accurate time series studies (detailed information is available in Lafrance et al. (2012)).

2.4. Land use map

The cultivated plots were identified by means of the RPG (*Registre Parcellaire Graphique*). The RPG is a detailed land use database describing the cultivated fields (surface, location, crop species and management).

3. Methods

3.1. Comparison of Formosat-2 and Spot data

To allow crop monitoring for winter and summer crops, continuous time series of high resolution satellite data are required. The use of multi-sensor data is then strongly recommended. However, there are several differences between the sensor systems in addition to their spectral bands responses differences. These additional differences are primarily due to differing orbital characteristics and spatial resolution. In this study, Formosat-2 data are compared with Spot data.

The NDVI products of F2 and SPOT were firstly compared. For that purpose, the SPOT pixels were re-sampled to the F2 pixels size of $8m \times 8m$, permitting a pixel to pixel comparison (not shown). A field scale comparison was performed as well, by calculating the average NDVI of each field over the study area. The mean NDVI was calculated based on the fields contours determined by the RPG.

Because of the difficulty of having SPOT and F2 acquisitions exactly on the same day, images acquired for time intervals of ± 3 days between the sensors were accepted (Table 1). View and solar angles of the pair image differed. Crop characteristics were assumed to be unchanged during the period between the paired acquisitions (3 days). No rain event was observed between the acquisitions.

The NDVI products were compared for the different crops encountered in our area of study. Only the dates within the period of vegetative development were used, from 2006 until 2010. Considering the Formosat-2 NDVI as the standard NDVI, a linear correction of the type $y = ax + b$ was applied to the SPOT NDVI to match the standard F2 [Steven et al. (2003)]. This intercalibration aims at obtaining a homogeneous NDVI dataset, regarding future application in GAI estimates. The same could be done to establish a SPOT standard product. This step allowed having an estimate of the correlation between the two sets of NDVI values. We evaluated afterward the influence of the different view angles on the residual errors between NDVI products.

Table 1 : Specifications of Spot and Formosat-2 paired dates of acquisitions: view and solar zenith (θ_v, θ_s) and azimuth (ϕ_v, ϕ_s) angles.

Date	Sensor	$\theta_v(^{\circ})$	$\theta_s(^{\circ})$	$\phi_v(^{\circ})$	$\phi_s(^{\circ})$
04/04/06	SPOT	-28	98	41	149
05/06/06	F2	22	349	27	133
23/06/06	SPOT	20	-75	21	152
23/06/06	F2	22	348	27	130
17/07/06	SPOT	-29	98	27	135
17/07/06	F2	22	348	28	131
23/11/06	SPOT	-1	101	64	168
23/11/06	F2	23	346	65	161
15/02/07	SPOT	-24	99	59	155
15/02/07	F2	26	226	60	153
06/07/07	SPOT	27	-72	22	150
07/07/07	F2	28	233	26	133
11/02/08	SPOT	21	-75	59	163
11/02/08	F2	34	245	61	155
20/06/08	SPOT	-16	103	25	134
19/06/08	F2	31	240	26	133
28/06/08	SPOT	-16	100	26	133
26/06/08	F2	31	240	26	132
19/03/09	SPOT	-22	99	48	150
17/03/09	F2	29	236	51	145
24/06/09	SPOT	26	-74	22	149
23/06/09	F2	29	236	29	124

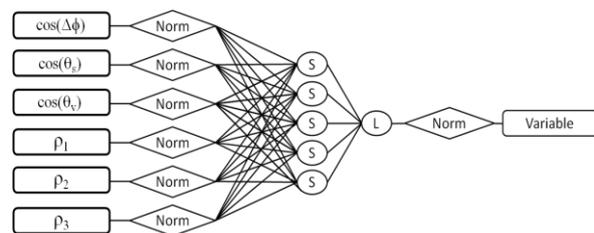


Figure 1. Neural Networks architecture [Weiss and Baret (2010)].

3.2. Retrieving GAI from remote sensing data

A wide range of methods have been developed to estimate biophysical variables from remote sensing data [Baret and Buis (2008)]. Two methods were chosen for this study: an empirical approach, linking biophysical variables and NDVI and a physical approach based on the application of neural network inversion of a radiative transfer model.

3.2.1. Empirical method

NDVI is often used to characterize vegetation. The use of this index reduces the anisotropic effects from the surface because the directional signatures are similar in its wavebands (red and infrared). NDVI, however, remains sensitive to changes in the observation geometry [Bacour et al. (2006)]. The empirical is one of the first and simpler approaches for retrieving GAI and other biophysical variables such as Fapar and Fcover. Nevertheless a large and representative amount of field data is needed, which is a drawback for up-scaling studies.

The relationship between NDVI and GAI can be easily established. For this study, the following equation was used (Eq.1).

$$GAI = \alpha \times (e^{\beta \cdot NDVI} - e^{\beta \cdot NDVI_s}) \quad (1)$$

where $NDVI_s$ is the bare soil NDVI. It was fixed to 0.13. The parameters α and β were calibrated using ground truth GAI measurements, by minimizing the Root Mean Square Error (RMSE) between estimated and observed GAI.

3.2.2. Physical method

This method is based on the inversion of the radiative transfer model PROSAIL using neural networks (NNT). This approach relies on three main steps: a) creation of a learning database, b) training the neural networks and finally c) applying the created networks for estimating the requested biophysical variables. This approach is also denominated as *BV-NET* tool (*Biophysical variable neural network* [Baret et al. (2007)]).

a) Creation of a learning database

The learning database consists of a table linking the input variables, which are the canopy reflectances simulated by PROSAIL model, with the output variables. This paper focuses on the GAI as main output variable, but estimates of Fapar and Fcover were accomplished too.

PROSAIL Model [Baret et al. (1992)]

PROSAIL is a combination of PROSPECT [Jacquemoud et al. (2009)] and SAIL models. The first one simulates leaf reflectances and transmittances, which are used as input variables for SAIL model. PROSPECT requires input parameters describing leaf properties: the mesophyll structure parameter (N), chlorophyll content, both a and b (Cab), water content (Cw), brown pigment content (Cbp) and dry matter content (Cdm). SAIL model provides directional reflectances, Fapar and Fcover variables. It is based on canopy extent. Input parameters describing structure of canopy and background soil reflectance are: green area index (GAI), average leaf angle (ALA), hotspot and brightness coefficient (Bs). The distribution of input parameters and variables used in this step is shown in Table 2.

b) Neural networks training

A different network is generated for each one of the sensors. SPOT satellites have two kinds of sensor instruments on board (HRG1 and HRG2). The analysis of their spectral responses showed that differences between instruments of a same satellite are not significant.

For computing-time gain, paired-sensors were considered as one, counting a total of four networks.

The neural network is characterized by the type of neurons (the transfer function) and their architecture. In this paper a back-propagation network with two hidden layers is applied. The first layer has five neurons with sigmoid transfer function and the second layer has one neuron, with linear transfer function (Figure 1). The combination of sigmoid and linear functions is capable of fitting any type of function [Weiss and Baret (2010)]. The neural network input layer is composed of a) the angles characterizing the observation geometry (view and solar zenithal angles, $\cos(\theta_v)$ and $\cos(\theta_s)$, and the relative azimuth angle $\cos(\Delta\phi)$) and b) the top of canopy reflectances in the different wavebands of the satellite sensor. The Levenberg-Marquadt optimization algorithm is used in the training process to obtain the best agreement between the output simulated by the network and the correspondent value of the training database.

Table 2 : Input variables distribution of PROSAIL model, used for the learning database creation.

	Variable	Minimum	Maximum	Mode	Std	Class	Law
Canopy structure	GAI	0	8	2	2	8	Gauss
	ALA (°)	5	80	40	20	5	Gauss
	Hotspot	0.1	0.5	0.2	0.5	1	Gauss
Leaf optical properties	N	1.2	2.2	1.5	0.3	4	Gauss
	Cab	20	90	45	30	5	Gauss
	Cdm	0.003	0.011	0.005	0.005	4	Gauss
	Cw	0.6	0.85	0.75	0.08	4	Uniform
Soil background property	Cbp	0	2	0	0.3	4	Gauss
Soil background property	Bs	0.16	1.3	0.586	0.14	4	Log-Normal

c) Applying the neural networks

After the training step, where a relationship is established between inputs and outputs, the neural networks are applied to the Formosat-2 and SPOT images. Output products are the images of the estimated biophysical variables (GAI, Fapar and Fcover), having the same resolution as the input satellite images.

4. Results and Discussion

4.1. Comparison of Formosat-2 and Spot data

Figure 2(a) shows the scatter-plot of the Formosat-2 and SPOT NDVI products over 3957 crop fields (wheat, rapeseed, sunflower, maize and soybean), for 11 different dates. A consistent linear relationship is observed. Some points appear more dispersed, but as it can be seen on the related density plot in Figure 2(b), they represent a quite insignificant percentage of the dataset. The comparison of F2 and SPOT vegetation indices presented a strong correlation and low errors ($R^2=0.98$, $RMSE = 0.034$ and $RRMSE = 6.76\%$). The results suggest there is no particular crop effect on the relationship. Slope a and interception b of the linear relationship found are 1.001 and -0.018, respectively. These coefficients can be applied to adjust the NDVI of SPOT to match the standard NDVI of the F2 satellite (or vice versa), aiming to obtain a relatively uniform NDVI dataset.

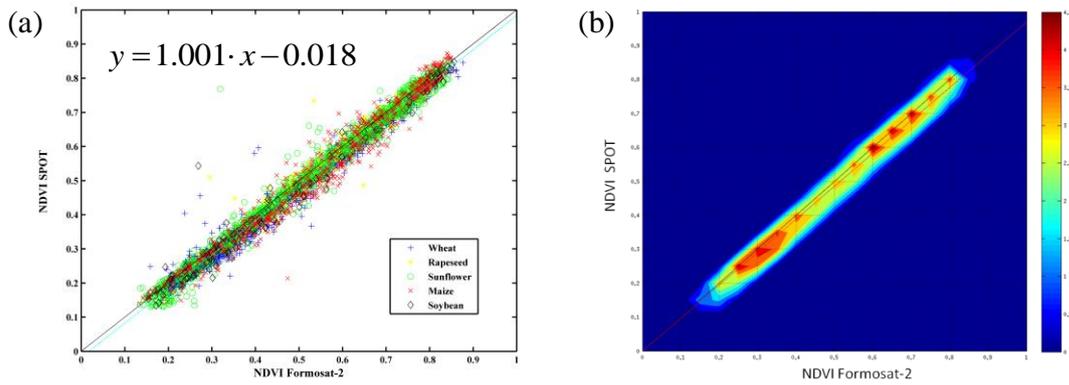


Figure 2. (a) Comparison of F₂ and SPOT NDVI products over 11 different paired dates and 3957 plots cultivated with wheat, rapeseed, maize, sunflower and soybean crops; (b) Density plot of NDVI Formosat-2 vs NDVI SPOT.

These results revealed that despite the differences in geometrical (solar and view angles) configuration of the acquisitions, NDVI from F₂ and SPOT sensors are quite similar. Figure 3 shows the relationship between NDVI residual errors and view angles differences. No significant correlation ($R^2=0.0136$) was found, suggesting a poor effect of observation geometries on these NDVI products.

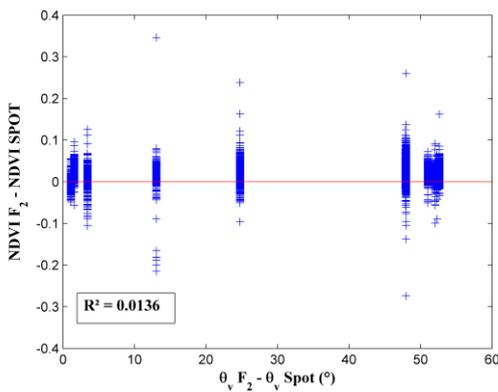


Figure 3. Residues plotted against difference between F₂ and SPOT view angles (θ_v).

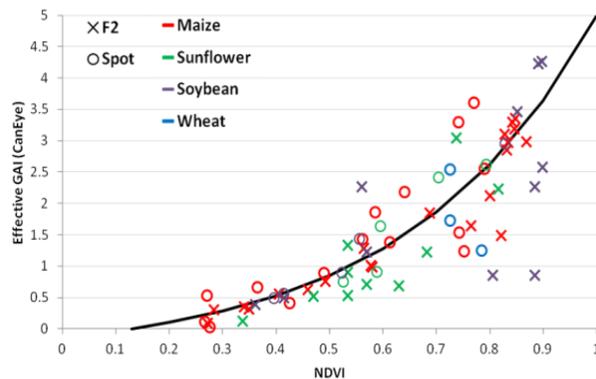


Figure 4. NDVI-GAI exponential relationship for SPOT and F₂ data.

4.2. Empirical method

The effective GAI data, obtained from the *in-situ* measurements performed over the two study area, allowed establishing a NDVI-GAI relationship for F₂ and calibrated SPOT data (Figure 4).

$$GAI = 0.31 \cdot e^{2.85 \cdot NDVI} - 0.45 \quad (2)$$

A single relationship was used for all crop types.

Only the remote sensing images acquired concurrently to ground measurements (hemispherical photographs) were used for establishing the empirical relationship (dates different from those cited on section.4.1). The empirical GAI was then calculated by applying this equation to the NDVI values of validation plots.

Figure 5(a) shows the scatter-plot of the empirical estimated GAI and the effective GAI obtained from hemispherical photographs. A strong correlation is observed ($R^2=0.86$). As expected, this approach yields good results. More accurate performance could be retrieved by establishing a different law for each crop type. However, a much larger set of field data would be necessary for this purpose.

4.3. Physical method

Figure 5(b) shows the GAI retrieved by Neural Networks approach compared with the effective GAI estimated by CAN-EYE. The performances in terms of correlation, bias, absolute and relative root mean square errors are indicated. We observe that larger values of GAI (superior than $2 \text{ m}^2 \cdot \text{m}^{-2}$) are overestimated by this method. However, different trends for the investigated crops can be distinguished. Note that for wheat and maize (*blue* and *red symbols*) the simulations agree very well with observations. For sunflower and soybean (*green* and *black symbols*), results suggest the simulations are biased, especially for greater values of GAI. The hemispherical photographs are taken at the same height during all the season. Consequently, when the vegetation is well developed the distance between the camera and the top of the canopy is shorter (than in the beginning of the season). It probably generates a saturation effect, resulting in lower estimations of effective GAI. As a result, there would be an underestimation of the GAI estimated with CAN-EYE, and not an overestimation of NNT estimates.

Furthermore, the GAI simulated by *BV-NET* tool was also compared to ground destructive measurements (Figure 5(c)). We observe that NNT retrievals are underestimated regarding the ground data. This was expected since remotely sensed products do not take into account the aggregation of the leaves (known as clumping effect), and the ground destructive samplings consider all the green organs of collected plants.

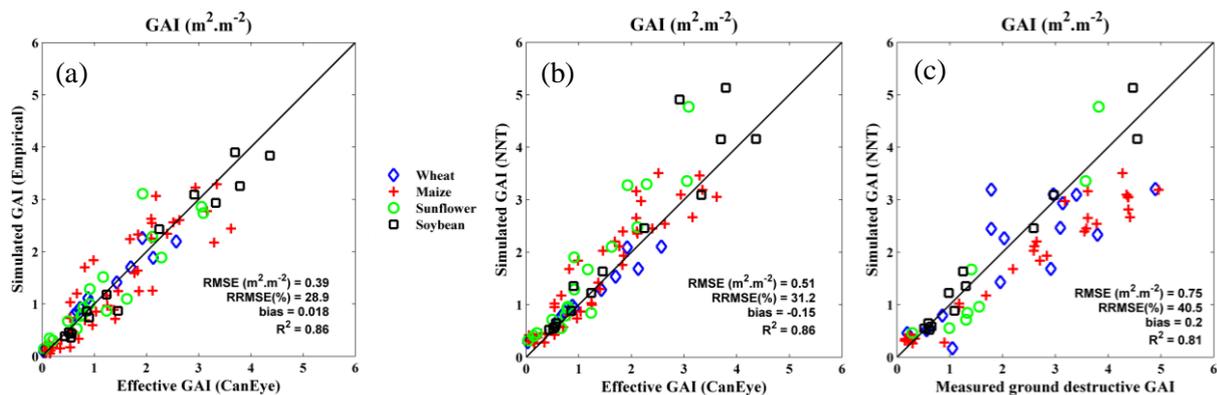


Figure 5. (a) Estimated effective GAI (with CAN-EYE) against GAI estimated through the empirical approach using NDVI from F2 and Spot satellites; (b) Estimated effective GAI (with CAN-EYE) against GAI estimated by the physically-based approach (NNT); (c) Comparison of GAI simulated through the NNT method against ground destructive GAI measurements.

5. Conclusion

In this study, NDVI derived from F2 and Spot satellites were compared. The results revealed a strong correlation between them and low influence of observation geometries. Thus with appropriate atmospheric and geometrical corrections, linear intercalibration is valuable and vegetation indices issued from the two sensors could be combined, increasing the opportunity of having cloud-free acquisitions for continuous crop monitoring.

When comparing the investigated methods for GAI retrieval, it is noticeable that the empirical approach yields better results (R^2 , RMSE). It requires however a large amount of field data, which is a common constraint. Besides, as the relationship NDVI-GAI is site-dependent, it has to be calibrated when applied to different landscapes. The physical-based method leads to good performances as well despite slight stronger errors (RMSE and RRMSE). Therefore this method is quite promising as it does not depend on any field

measurements, and offers great perspectives to regional scale applications.

The availability of accurate, homogeneous and complete series of GAI estimates over the study area will allow the assimilation of these products into crop models. It will potentially lead to better performances in crops biomass and yield estimates.

6. References

- Bacour, C., F. M. Bréon and F. Maignan. Normalization of the directional effects in NOAA-AVHRR reflectance measurements for an improved monitoring of vegetation cycles. **Remote sensing of environment**, v.102, n.3, p.402-413, 2006.
- Baret, F. and S. Buis. Estimating canopy characteristics from remote sensing observations: review of methods and associated problems. **Advances in Land Remote Sensing. New York, USA. Springer**, p.173-201, 2008.
- Baret, F., O. Hagolle, B. Geiger, P. Bicheron, B. Miras, M. Huc, B. Berthelot, F. Niño, M. Weiss and O. Samain. LAI, fAPAR and fCover CYCLOPES global products derived from VEGETATION: Part 1: Principles of the algorithm. **Remote Sensing of Environment**, v.110, n.3, p.275-286, 2007.
- Baret, F., S. Jacquemoud, G. Guyot and C. Leprieur. Modeled analysis of the biophysical nature of spectral shifts and comparison with information content of broad bands. **Remote Sensing of Environment**, v.41, n.2, p.133-142, 1992.
- Béziat, P., E. Ceschia and G. Dedieu. Carbon balance of a three crop succession over two cropland sites in South West France. **Agricultural and Forest Meteorology**, v.149, n.10, p.1628-1645, 2009.
- Claverie, M., V. Demarez, B. Duchemin, O. Hagolle, D. Ducrot, C. Marais-Sicre, J. F. Dejoux, M. Huc, P. Keravec and P. Béziat. Maize and sunflower biomass estimation in southwest France using high spatial and temporal resolution remote sensing data. **Remote Sensing of Environment**, 2012.
- Dedieu, G., A. Karnieli, O. Hagolle, H. Jeanjean, F. Cabot, P. Ferrier and Y. Yaniv. **VENUS: A joint Israel-French Earth Observation scientific mission with high spatial and temporal resolution capabilities**. Recent Advances in Quantitative Remote Sensing Valencia, Spain, 2006. 517-521 p.
- Demarez, V., S. Duthoit, F. Baret, M. Weiss and G. Dedieu. Estimation of leaf area and clumping indexes of crops with hemispherical photographs. **Agricultural and forest meteorology**, v.148, n.4, p.644-655, 2008.
- Jacquemoud, S., W. Verhoef, F. Baret, C. Bacour, P. J. Zarco-Tejada, G. P. Asner, C. François and S. L. Ustin. PROSPECT + SAIL models: A review of use for vegetation characterization. **Remote Sensing of Environment**, v.113, Supplement 1, n.0, p.S56-S66, 2009.
- Lafrance, B., X. Lenot, C. Ruffel, P. Cao and T. Rabaute. Outils de prétraitements des images optiques KALIDEOS. **Revue française de photogrammétrie et de télédétection**, n.197, p.10-16, 2012.
- Martimor, P., O. Arino, M. Berger, R. Biasutti, B. Carnicero, U. Del Bello, V. Fernandez, F. Gascon, P. Silvestrin and F. Spoto. **Sentinel-2 optical high resolution mission for gmes operational services**. Geoscience and Remote Sensing Symposium, 2007. IGARSS 2007. IEEE International: IEEE, 2007. 2677-2680 p.
- Steven, M. D., T. J. Malthus, F. Baret, H. Xu and M. J. Chopping. Intercalibration of vegetation indices from different sensor systems. **Remote Sensing of Environment**, v.88, n.4, p.412-422, 2003.
- Weiss, M. and F. Baret. **Venus biophysical variable products algorithm theoretical basis document**. EMMAH, INRA. 2010