Consistent and accurate LAI, FAPAR and FCOVER global products: principles and evaluation of GEOV1 products

Frédéric Baret, Marie Weiss, Roselyne Lacaze, Fernando Camacho, Philippe Pacholcyzk, Hassan Makhmara, Bruno Smets

Consistent and accurate LAI, FAPAR and FCOVER global products: principles and evaluation of GEOV1 products

F. Baret[1], M. Weiss[1], R. Lacaze[2], F. Camacho[3], P. Pacholcyzk[4] H. Makhmara[4] and B. Smets[5]
[1] *INRA-EMMAH UMR 1114, Avignon, France*
[2] *HYGEOS/CNES/GEOLAND2, Toulouse, France*
[3] *EOLAB, Valencia, Spain*
[4] *CNES, Toulouse, France*
[5] *VITO, Mol, Belgium*

baret@avignon.inra.fr

ABSTRACT *LAI, FAPAR and FCOVER variables are required for the monitoring, understanding and modelling of land surfaces at the global scale. While several products were already developed from the current medium resolution sensors, the few validation exercises achieved demonstrated that significant discrepancies and inconsistencies were observed. The objective of this study was to develop new global estimates of LAI, FAPAR and FCOVER that will build on the pros and minimize cons of already existing products. In a first step, the performances of the MODIS, CYCLOPES, GLOBACRBON and JRC-FAPAR products were reviewed. The MODIS and CYCLOPES products were selected since they provide higher level of consistency. These products were then fused to generate the 'best estimate' of LAI, FAPAR and FCOVER that were later scaled to better match their expected range of variation. Finally, neural networks were trained to estimate these best estimates products from SPOT-VEGETATION top of canopy directionally normalized reflectance values. Performances of the derived products called GEOV1 were evaluated, showing significant improvements as compared to previous products. These products will be extended back to 1981 using the AVHRR series of observation, and continued after the VEGETATION era thanks to AVHRR-METOP, PROBA-V and Sentinel3 future missions.*

1 INTRODUCTION

The importance of continuous monitoring the Earth's surface was recently recognized by GCOS (GCOS, 2006): a set of Essential Climate Variables was identified as being both accessible from remote sensing observations and intervening within key processes. Among those related to land surfaces, LAI (Leaf Area Index) and $FAPAR$ (Fraction of Absorbed Photosynthetic Active Radiation) may be derived from observations in the reflective solar domain. These vegetation biophysical variables play a key role in several processes, including photosynthesis, respiration and transpiration. LAI is defined as half the total developed area of leaf elements per unit horizontal ground area (Chen and Black, 1992). $FAPAR$ is defined as the fraction of radiation absorbed by the canopy in the 400 - 700 nm spectral domain under specified illumination conditions. It is one of the main inputs in light use efficiency models (McCallum et al., 2009). The cover fraction ($FCOVER$) defined as the fraction of background covered by green vegetation as seen from nadir appears also a very pertinent variable that can be used when separating the contribution of the soil from that of the canopy.

Products	LAI	FAPAR	FCOVER	Temporal frequency (days)	Time period	Projection	Reference
MODIS C5	✓	✓		8	2000-2008	Sinusoidal	(Myneni et al., 2002)
CYCLOPES V3	✓	✓	✓	10	1999-2007	Lat-lon	(Baret et al., 2007)
GLOBCARBON	✓	✓		30	1999-2007	Lat-lon	(Deng et al., 2006)
JRC-FAPAR		✓		1	1997-2006	Lat-lon	(Gobron et al., 2006)

Table 1. The currently available GLOBAL products.

Few global *LAI*, *FAPAR* and *FCOVER* products have already been generated from medium spatial resolution sensors such as VEGETATION, SEAWIFS, MODIS and MERIS (Table 1). Recent validation activities have shown however that significant discrepancies were existing between them as well as with ground measurements (Garrigues et al., 2008; McCallum et al., 2010; Weiss et al., 2007), calling thus for the development of new products that would reconcile these differences.

The Geoland2 project (http://www.geoland2.eu) intends to implement a Land Monitoring Core Service that corresponds to a European contribution to GEOSS (Group of Earth Observation System of Systems). The Bio-geophysical Parameters (BioPar) service within Geoland2 aims at developing pre-operational infrastructures for providing global land products both in near real time and off-line mode with long time series.

The objective of this paper is to describe the first version of Geoland2 *LAI*, *FAPAR* and *FCOVER* products that will be called GEOV1. The principles used to derive the products will first be presented. Then, the algorithm development will be described as long as some validation results.

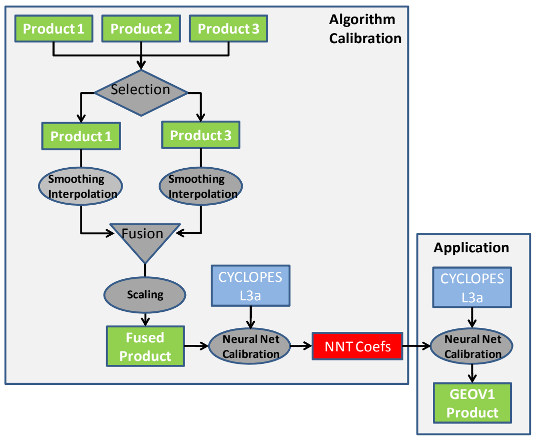

Figure 1. Schematic description of the principle used to develop the GEOV1 product.

2 ALGORITHM DEVELOPMENT

2.1 Principles

The biophysical algorithm is based on already existing products to capitalize on the efforts accomplished and get a larger consensus from the user community. Figure 1**Erreur ! Source du renvoi introuvable.**

shows the several steps used for each product. Following the published literature on products validation (Garrigues et al., 2008; Weiss et al., 2007), the best performing products were selected and combined to take advantage of their specific performances while limiting the situations where products show deficiencies. The selected products are re-projected onto the VEGETATION plate-carrée 1/112° grid, smoothed through time and interpolated at the 10 days frequency. Then the products are combined and eventually scaled to compute the fused product that is expected to provide globally the 'best' performances. The fused products are generated for 2003-2004 over the 420 BELMANIP2 set of sites (Figure 2) that is supposed to represent the possible range of surface types and conditions over the Earth (Baret et al., 2006). Neural networks are then calibrated over this set of sites to relate the fused products to the corresponding VEGETATION L3a top of canopy directionally normalized reflectances using the CYCLOPES pre-processing algorithms (Baret et al., 2007).

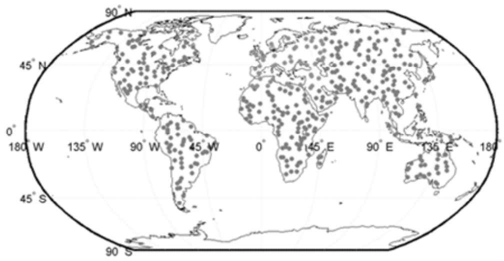

Figure 2. The 420 BELMANIP2 sites used to sample vegetation types and conditions.

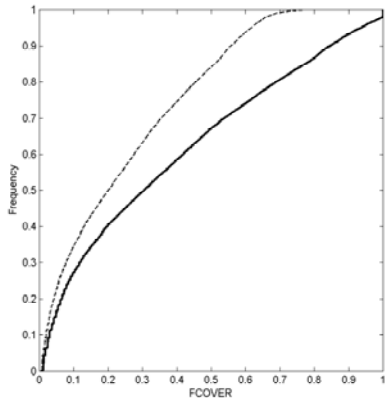

Figure 3. Cumulated frequency of $FCOVER$ CYCLOPES (dashed line) and GEOV1 products as observed over the 420 BELMANIP2 sites during years 2003-2004.

2.2 Generation of the training data base

$FCOVER$ product on one side and LAI and $FAPAR$ products on the other will be described separately because of the differences in available products.

a) $FCOVER$

Since only the CYCLOPES products were available globally, no fusion with other products was possible. However, several evaluations have shown that CYCLOPES $FCOVER$ products were suffering from a significant systematic underestimation (Verger, 2008). This was corrected for by applying a scaling factor to the CYCLOPES V3.1 products ($FCOVER_{CYCV31}$). This factor ($\frac{1}{0.6872}$) was corresponding to the inverse of the $FCOVER$ value for the 99% cumulated frequency (Figure 3) that should be expected to be very close to 1.0 since it should correspond to very dense canopies:

$$FCOVER_{best} = \frac{1}{0.68} \cdot FCOVER_{CYCV31} \quad (1)$$

Where $FCOVER_{best}$ is the value that will be used for training the neural networks.

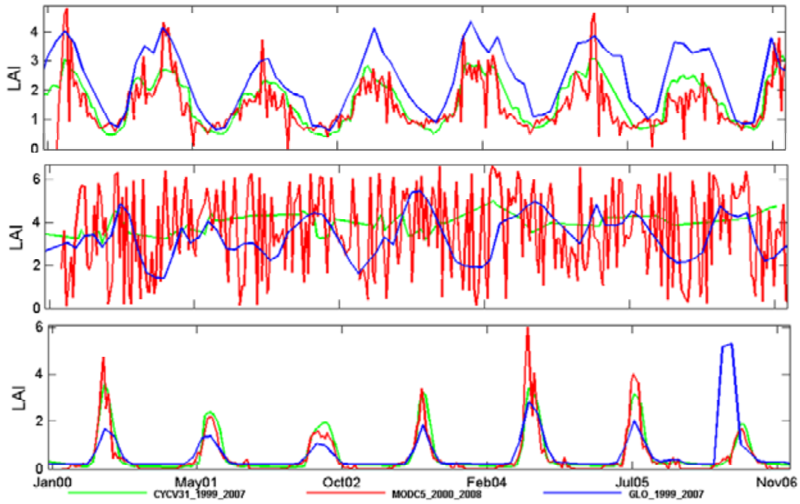

Figure 4. Example of LAI dynamics of CYCLOPES, MODIS and GLOBCARBON products for 3 typical sites.

b) LAI and $FAPAR$

The 30 days temporal sampling used for GLOBCARBON appears not very well suited to describe the seasonality of vegetation as shown in Figure 4. This is the reason why GLOBCARBON products were not selected. Further, GLOBCARBON LAI products were showing a significant number of outliers. MODIS and CYCLOPES LAI products will therefore be selected.

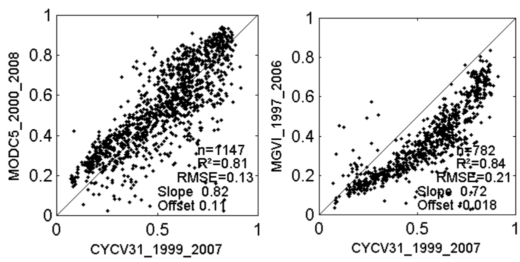

Figure 5. Scatterplot between FAPAR products.

JRC-FAPAR products derived from the SEAWIFS sensor show very similar seasonality to that of the MODIS and CYCLOPES $FAPAR$ products with however generally lower values as shown in Figure 5.

MODIS and CYCLOPES $FAPAR$ products were showing a closer agreement, particularly for the medium to high $FAPAR$ values. MODIS and CYCLOPES were thus selected to generate the $FAPAR$ products. This will further provide better consistency between LAI and $FAPAR$ products. Note that the definition of $FAPAR$ products is not very different between the several products: MODIS and JRC-FAPAR are instantaneous black-sky at the time of satellite overpass (around 10:30), while CYCLOPES corresponds to instantaneous black-sky at 10:00 which is a good approximation of the daily integrated black-sky $FAPAR$ value (Baret et al., 2007).

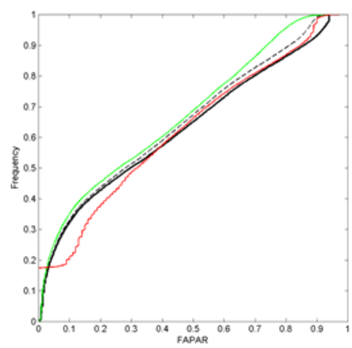

Figure 6. Cumulated frequency of $FAPAR$ CYCLOPES (green line), MODIS (red line) products as observed over the 420 BELMANIP2 sites during years 2003-2004. The dashed black line corresponds to the fusion product (according to equation 3) and the solid black line to GEOV1 products.

Investigation of the relationships between MODIS and CYCLOPES LAI and $fAPAR$ products show that:

- MODIS $FAPAR$ overestimates CYCLOPES values for the lower $FAPAR$ values (**Erreur ! Source du renvoi introuvable.**).
- Fair agreement is observed for medium to high $FAPAR$ values with however slightly lower values for CYCLOPES products (**Erreur ! Source du renvoi introuvable.**).
- Fair agreement is observed between MODIS and CYCLOPES LAI values up to values around 3.

To benefit from the better performances observed for CYCLOPES $FAPAR$ products for the lower $FAPAR$ values, and for MODIS LAI products for the larger LAI values, it was proposed to average MODIS and CYCLOPES products using the following weighing factor $w = \min(1, \frac{1}{4} LAI_{CYCV31})$:

$$\begin{cases} fAPAR_{fused} = fAPAR_{MODC5} \cdot w + fAPAR_{CYCV31} \cdot (1-w) \\ LAI_{fused} = LAI_{MODC5} \cdot w + LAI_{CYCV31} \cdot (1-w) \end{cases} \quad (2)$$

This parallel processing of LAI and $FAPAR$ is expected to keep a good consistency between LAI and $FAPAR$ products. The fused $FAPAR$ products showed that the maximum values are around 0.898 (**Erreur ! Source du renvoi introuvable.**) although maximum values are expected to be close to 0.94 (Baret and Guyot, 1991). Therefore, the fused values were scaled according to:

$$fAPAR_{best} = \frac{0.94}{0.898} \cdot fAPAR_{fused} \quad (3)$$

No particular scaling was applied to LAI since there is no obvious maximum values to set up.

Results show that, as expected, the relationship between LAI and $FAPAR$ was keeping very consistent as compared to the original CYCLOPES and MODIS products (**Erreur ! Source du renvoi introuvable.**)

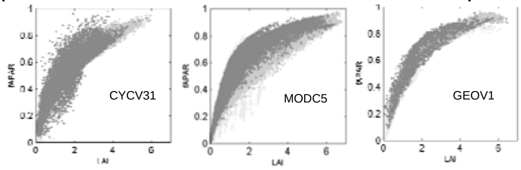

Figure 7. Relationship between LAI and $FAPAR$ for CYCLOPES, MODIS and GEOV1 products as observed over the 420 BELMANIP2 sites during 2003-2004 period.

2.3 Training the neural networks

The training was achieved over the 420 BELMANIP2 sites for the 2003-2004 period. The spatial support was 3 x 3 pixels over each site to minimize possible geometrical problems between the several products and dates used to compute the inputs and outputs of the neural network.

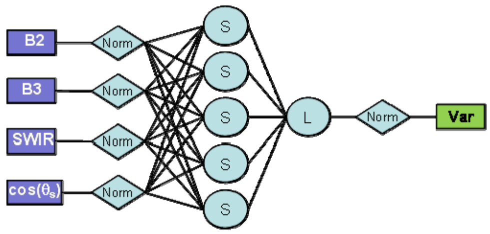

Figure 8. Structure of the neural network used to derive LAI, $FAPAR$ and $FCOVER$ from VEGETATION input reflectance.

a) Neural network architecture

A back-propagation neural network architecture was selected, with one hidden layer of 5 tangent-sigmoidal neurons and one layer with a single linear neuron (Figure 8). Inputs and outputs were normalized using the minimum and maximum values. Five networks were trained in parallel. The one providing the best performances over an independent test data set was selected.

b) Inputs

The top of canopy reflectance acquired by the VEGETATION sensor in the red (B2), near infrared (B3) and short wave infrared (SWIR) were used as inputs. The preprocessing steps are described in (Baret et al., 2007) and include cloud screening, atmospheric correction based on a climatology of aerosols, and BRDF normalization using a robust fit of the Roujean model (Hagolle et al., 2004; Roujean et al., 1992). In addition, the cosine of the sun zenith angle at the observation time is also used as input.

c) Outputs

Special attention was carried out when fusing MODIS and CYCLOPES products for the computation of the 'best' LAI and $FAPAR$ estimates. MODIS products were re-projected to the CYCLOPES lat-lon grid. All the valid MODIS data including main, main plus saturation and back-up algorithms available within ±10 days around the CYCLOPES date were considered. This may correspond to a maximum of three MODIS products. Then, if the difference between values at the 70% and 90% cumulated frequency was lower than 0.2 and 1.0 respectively for $FAPAR$ and LAI, the 70% cumulated frequency value is retained for the fusion. Note that because BELMANIP2 sites are relatively homogeneous at the 10x10 pixels scale, the 70% cumulated frequency value is very close to the mean value. However, taking the 70% cumulated frequency value instead of the mean or the median value prevents from being too sensitive to possible unscreened clouds or cloud shadow that lead to lower LAI and $FAPAR$ values.

2.4 Associated uncertainties and quality assessment

Three quality assessment criterions were provided along with the products:

- Input out of range. This represents the consistency of the measured VEGETATION input reflectances with those used in the training data base. A flag is raised when observations are outside the training definition domain. The definition domain was approximated by the convex hull formed in the reflectance feature space by the cases used in the learning process (Figure 9). When the input reflectances are outside the definition domain, a flag is raised.

Table 2. Minimum, Maximum, Resolution and Tolerance values used to raise the output of range flag.

	Unit	Min	Max	Resol.	Tol_{min}	Tol_{max}
LAI	$m^{-2}.m^{-2}$	0	8.0	0.01	-0.2	7.00
FAPAR	-	0	1.0	0.01	-0.05	1.05
FCOVER	-	0	1.0	0.01	-0.05	1.05

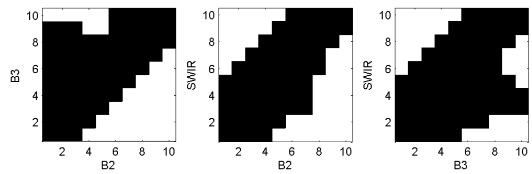

Figure 9. Definition domain of the input reflectance. The cells in black correspond to those where input reflectance were actually observed. Cells in white are outside the definition domain.

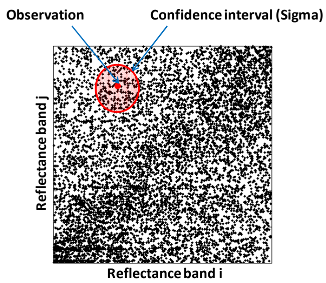

Figure 10. Scheme showing how the uncertainties attached to the products were computed.

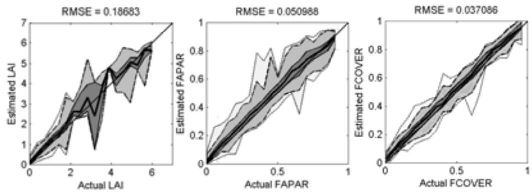

Figure 11. Relationship for LAI, $FAPAR$ and $FCOVER$ products before (actual) and after (estimated) the training process over a test dataset not used in the training process.

- Output out of range. This flag is raised only when the output is outside the output range enlarged by the tolerance values [Tol_{min}, Tol_{max}] as defined in Table 2. If the outputs of the neural network falls within the [Tol_{min}, Min] (respect. [Max, Tol_{max}]), the values are simply reset to the Minimum (respect. maximum) values.
- Estimated uncertainties. This represents the expected error expressed in RMSE between the estimated and the actual biophysical values as derived from the theoretical performances of the algorithm evaluated over an independent data set. The reflectance uncertainties are used to define a confidence interval (Figure 10). The LAI, $FAPAR$ and $FCOVER$ with corresponding reflectance inside the confidence interval are then used to compute the RMSE. A specific network is finally trained to relate the estimated uncertainties to the input reflectance and observation geometry values.

3 PERFORMANCES

3.1 Theoretical performances

Comparison between estimates by the neural network and the actual 'best' $\boldsymbol{LAI}$, $\boldsymbol{FAPAR}$ or $\boldsymbol{FCOVER}$ values as evaluated over an independent test data set show very good performances without any biases (Figure 11). The scattering as measured by the RMSE is also very small.

3.2 Case studies

Three sites were selected to evaluate qualitatively the performances. An extended validation was concurrently achieved, based on quantitative metrics as proposed by CEOS (Morissette et al. 2006; Garrigues et al. 2009).

Figure 12 shows that GEOV1 products are very smooth as expected. This is due both to the quality of the pre-processing steps as well as to the properties of the neural network. The seasonality is very consistent with that of the other products. The range of variation appears quite realistic, both for the low vegetation amounts (Figure 12, left) and the larger ones (Figure 12, left). The consistency between $\boldsymbol{LAI}$ and $\boldsymbol{FAPAR}$ is also very strong (Figure 7, right) as expected.

4 CONCLUSION

The GEOV1 $\boldsymbol{LAI}$, $\boldsymbol{FAPAR}$ and $\boldsymbol{FCOVER}$ products capitalize on the efforts undertaken this last decade in the development and validation of biophysical products from medium resolution observations. It results in robust, consistent and accurate estimates of these key biophysical variables that may be used for a range of applications including those targeted for the Essential Climate Variables.

These products are currently generated by VITO for open access to the user community. The VEGETATION derived time series starting in 1999 will be completed backward using the AVHRR series as processed by Vermote et al. (2010) to get a long time series of almost 30 years. Further, the sustainability of services foreseen within the Land Monitoring Core Service will be ensured by adapting the algorithm to AVHRR-METOP, PROBA-V and Sentinel 3 missions.

Acknowledgments.

This study was funded partly from the GEOLAND2 FP7 European project.

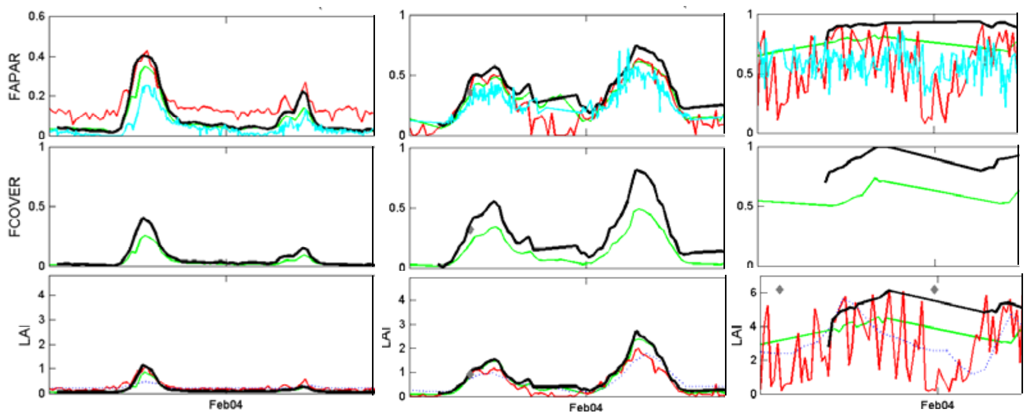

Figure 12. Temporal profiles of LAI (bottom), FAPAR (top) and FCOVER (middle) products: MODIS (red), CYCLOPES (green), GLOBACARBON (dashed blue), JRC-FAPAR (cyan) and GEOV1 (black). Gourma grassland (left 15.32°; -1.55°), Fundulea crop land (centre 44.41°; 26.58°) and Tapa evergreen broadleaf forest (right 2.87°; -54.95°) sites are presented for years 2003-2004.

References

Baret, F. and Guyot, G., 1991. Potentials and limits of vegetation indices for LAI and APAR assessment. Remote Sensing of the Environment, 35(2-3): 161-173.

Baret, F., Hagolle, O., Geiger, B., Bicheron, P., Miras, B., Huc, M., Berthelot, B., Weiss, M., Samain, O., Roujean, J.L. et al., 2007. LAI, fAPAR and fCover CYCLOPES global products derived from VEGETATION. Part 1: Principles of the algorithm. Remote Sensing of Enviroment, 110: 275-286.

Baret, F., Morissette, J., Fernandes, R., Champeaux, J.L., Myneni, R., Chen, J., Plummer, S., Weiss, M., Bacour, C., Garrigue, S. et al., 2006. Evaluation of the representativeness of networks of sites for the global validation and inter-comparison of land biophysical products. Proposition of the CEOS-BELMANIP. IEEE Transactions on Geoscience and Remote Sensing, 44(7: special issue on global land product validation): 1794-1803.

Chen, J.M. and Black, T.A., 1992. Defining leaf area index for non-flat leaves. Plant, Cell and Environment, 15: 421-429.

Deng, F., Chen, J.M., Chen, M. and Pisek, J., 2006. Algorithm for global leaf area index retrieval using satellite imagery. IEEE Transactions on Geoscience and Remote Sensing, 44(8): 2219-2229.

Garrigues, S., Lacaze, R., Baret, F., Morisette, J., Weiss, M., Nickeson, J., Fernandes, R., Plummer, S., Shabanov, N.V., Myneni, R. et al., 2008. Validation and Intercomparison of Global Leaf Area Index Products Derived From Remote Sensing Data. Journal of Geophysical Research, 113(G02028).

GCOS, 2006. GCOS-107. Supplemental details to the satellite based component of the "implementation plan for the global observing system for climate in support of the UNFCCC". WMO/TN N° 1338, GCOS/WMO, Geneve (Switzerland).

Gobron, N., Pinty, B., Aussedat, O., Chen, J.M., Cohen, W.B., Fensholt, R., Gond, V., Huemmrich, K.F., Lavergne, T., Melin, F. et al., 2006. Evaluation of fraction of absorbed photosynthetically active radiation products for different canopy radiation transfer regimes: Methodology and results using Joint Research Center products derived from SeaWiFS against ground-based estimations. Journal Of Geophysical Research-Atmospheres, 111(D13).

Hagolle, O., Nicolas, J.-M., Fougnie, B., Cabot, F. and Henry, P., 2004. Absolute calibration of VEGETATION derived from an interband method based on the Sun glint over ocean. IEEE Transactions on Geoscience and Remote Sensing, 42(7): 1472 -1481.

McCallum, I., Wagner, W., Schmullius, C., Shvidenko, A., Obersteiner, M., Fritz, S. and Nilsson, S., 2010. Comparison of four global FAPAR datasets over Northern Eurasia for the year 2000. Remote Sensing of Environment, 114(5): 941.

Myneni, R.B., Hoffman, S., Knyazikhin, Y., Privette, J.L., Glassy, J., Tian, Y., Wang, Y., Song, X., Zhang, Y., Smith, G.R. et al., 2002. Global products of vegetation leaf area and absorbed PAR from year one of MODIS data. Remote Sensing of Environment, 83: 214-231.

Roujean, J.L., Leroy, M. and Deschamps, P.Y., 1992. A bidirectional reflectance model of the Earth's surface for the correction

of remote sensing data. Journal of geophysical research, 97(D18): 20455-20468.

Verger, A., 2008. Analisi comparativa d'algorismes operacionals d'estimacio de parametres biofisics de la coberta vegetal amb teledeteccio, Universitat de Valencia, Valencia (Spain), 277 pp.

Vermote, E., Justice, C., Csiszar, I., Eidenshink, J., Mynemi, R., Baret, F., Masuoka, E. and Wolfe, R., 2010. A terrestrial surface climate data record for global change studies. In: J. Sobrino (Editor), Third International Symposium on Recent Advances in Quantitative Remote Sensing, Torrent (Spain).

Weiss, M., Baret, F., Garrigues, S., Lacaze, R. and Bicheron, P., 2007. LAI, fAPAR and fCover CYCLOPES global products derived from VEGETATION. part 2: Validation and comparison with MODIS Collection 4 products. Remote sensing of Environment, 110: 317-331.