

# GEOV1: LAI and FAPAR essential climate variables and FCOVER global time series capitalizing over existing products. Part1: Principles of development and production

Frédéric Baret, M Weiss, R Lacaze, F Camacho, H Makhmara, P Pacholcyzk,

B Smets

## ► To cite this version:

Frédéric Baret, M Weiss, R Lacaze, F Camacho, H Makhmara, et al.. GEOV1: LAI and FAPAR essential climate variables and FCOVER global time series capitalizing over existing products. Part1: Principles of development and production. Remote Sensing of Environment, 2013, 137, pp.299-309. 10.1016/j.rse.2012.12.027 . hal-01321330

# HAL Id: hal-01321330 https://hal.science/hal-01321330

Submitted on 29 May 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

#### 1 GEOV1: LAI, FAPAR Essential Climate Variables and FCOVER global time series

- 2 capitalizing over existing products. Part1: Principles of development and production.
- 3
- 4 F. Baret<sup>1</sup>, M. Weiss<sup>1</sup>, R. Lacaze<sup>2</sup>, F. Camacho<sup>3</sup>, H. Makhmara<sup>4</sup>, P. Pacholcyzk<sup>4</sup> and B. Smets<sup>5</sup>
- 5 <sup>1</sup> INRA-EMMAH UMR 1114, Avignon, France
- 6 <sup>2</sup> HYGEOS, Toulouse, France
- 7 <sup>3</sup> EOLAB, Valencia, Spain
- 8 <sup>4</sup> CNES, Toulouse, France
- 9 <sup>5</sup> VITO, Mol, Belgium
- 10 Corresponding author: baret@avignon.inra.fr

#### 11 ABSTRACT

Essential Climate variables such as LAI or FAPAR are required for the monitoring, 12 understanding and modeling of land surfaces at the global scale. While several products were 13 already developed from the current medium resolution sensors, the few validation exercises 14 currently achieved highlighted significant discrepancies and inconsistencies. The objective of 15 this study is to develop improved global estimates of LAI, FAPAR and FCOVER variables by 16 capitalizing on the development and validation of already existing products. In a first step, the 17 performances of the MODIS, CYCLOPES, GLOBCARBON and JRC-FAPAR products were 18 reviewed. The MODIS and CYCLOPES products were then selected since they provide 19 higher level of consistency. These products were fused to generate the improved LAI, FAPAR 20 and FCOVER values that were later scaled to closely match their expected range of variation. 21 Finally, neural networks were trained to estimate these fused and scaled products from SPOT-22 23 VEGETATION top of canopy directionally normalized reflectance values. The resulting GEOV1 products are associated to quality control flags as well as quantitative estimates of 24 uncertainties. Performances of the GEOV1 products are finally evaluated in a companion 25 26 paper. The GEOV1 products are freely available to the community at www.geoland2.eu from 1999 up to present, globally at 1/112° spatial sampling grid at the dekadal time step. 27

28

## 29 **1 Introduction**

The importance of continuously monitoring the Earth's surface was recently recognized by 30 GCOS (Global Climate Observing System) (GCOS 2006). Essential Climate Variables (ECV) 31 related to land surfaces such as LAI (Leaf Area Index) and (Fraction of Absorbed 32 Photosynthetic Active Radiation) may be derived from observations in the reflective solar 33 domain. These vegetation biophysical variables play a key role in several surface processes, 34 including photosynthesis, respiration and transpiration. LAI is defined as half the total 35 developed area of green elements per unit horizontal ground area (Chen and Black 1992). 36 FAPAR is defined as the fraction of radiation absorbed by the green vegetation elements in the 37 400 - 700 nm spectral domain under specified illumination conditions. FAPAR is one of the 38 main inputs in light use efficiency models (McCallum et al. 2009). In addition to LAI and 39 40 FAPAR, FCOVER, the fraction of green vegetation as seen from nadir, is requested by some

Version définitive du manuscrit publiée dans / Final version of the manuscript published in : Remote Sensing of Environment (2013), Vol. 137, p. 299-309, DOI: 10.1016/j.rse.2012.12.027 Journal homepage: www.elsevier.com/locate/rse

- 41 users for vegetation monitoring (Lacaze et al. 2009) as well as for partitioning contributions
- 42 between soil an vegetation within specific models for Numerical Weather Prediction, regional
- 43 and global climate modeling, and global change monitoring (Avissar and Pielke 1989).
- 44 FCOVER is independent from the illumination conditions as opposed to FAPAR while
- 45 showing sensitivity to vegetation amount intermediate between *FAPAR* and *LAI*.

# 46 Table 1. The currently LAI, FAPAR and FCOVER products available globally at

### 47 approximately 1 km spatial sampling distance.

Products	Sensors	LAI	FAPAR	FCOVERR	Spatial sampling distance at equator	Temporal sampling (days)	Time period	Reference
MODIS C5	MODIS	~	~		1 km	8	2000-present	(Yang et al. 2006b)
CYCLOPES V3.1	VEGETATION	~	✓	~	1 km	10	1999-2007	(Baret et al. 2007)
GLOBCARBON	VEGETATION	~	✓		1 km	30	1999-2007	(Deng et al. 2006)
JRC-FAPAR	SEAWIFS <sup>(1)</sup>		~		2.17 km	1	1997-2006	(Gobron et al. 2006)

<sup>(1)</sup> JRC-FAPAR may be derived from several sensors including SEAWIFS, MERIS, MODIS
 and VEGETATION but produced globally over a long time period only from SEAWIFS.

Few global LAI, FAPAR and FCOVER products have already been generated from 50 VEGETATION, SEAWIFS, MODIS and MERIS sensors with a spatial sampling distance 51 52 close to 1km. Improved atmospheric correction, radiometric calibration and model formulation have incrementally enhanced the retrieval accuracy of the successive 53 reprocessing. Recent validation activities have shown however that significant discrepancies 54 55 were existing between these global products 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 56 products that would reconcile these differences. 57

The FP7 Geoland2 project (http://www.gmes-geoland.info) intends to implement the 58 GMES (Global Monitoring for Environment and Security) Land Monitoring Core Service that 59 corresponds to the European contribution to GEOSS (Group of Earth Observation System of 60 Systems). An operational system is developed to provide biophysical products that meet the 61 users' needs (Lacaze et al. 2009) for monitoring natural ecosystems and managed lands. The 62 main requirements correspond to 1 km spatial sampling, 10 days (dekadal) frequency with 63 products generated in near real time (less than one week), for a time period as long as possible 64 (Ganguly et al. 2008b), and associated with quantitative uncertainties. These requirements are 65 partly answering the recently updated GCOS ones for LAI and FAPAR ECVs with 250m 66 spatial resolution relaxed to 2 km for near term products, 2-weekly temporal frequency and 67 accuracy better than max(20%, 0.5) for LAI, max(10%, 0.05) for FAPAR and a stability better 68 69 than max(10%, 0.25) for LAI and max(3%, 0.02) for FAPAR (GCOS-138). The resulting products will thus eventually contribute to fulfill one of the GCOS task dedicated to the 70 operationalization of the generation of and LAI Essential Climate Variables. 71

The objective of this paper is to describe the first version of Geoland2 *LAI*, and *FCOVER*products called GEOV1. The principles used to derive the products will first be presented.
Then, the algorithm development will be described. A companion paper (Camacho et al.
2012) will finally report the validation results derived according to the CEOS/LPV guidelines
(Baret et al. 2009; Morisette et al. 2006).

## 77 2 ALGORITHM DEVELOPMENT

#### 78 2.1 General principles

Since significant efforts have already been made to develop and validate biophysical products as stated in the introduction, it is thus proposed to capitalize on the existing products and associated validation results for the development of the GEOV1 products. Previous works (Verger et al. 2008) have demonstrated the capacity of learning machines such as neural networks to estimate biophysical products including *LAI* from several sensors. This property will be exploited here, allowing to use several sensors in order to build a long time series of products. The algorithm is made of three main steps as sketched in Figure 1:

*1-Generation of the training dataset:* already existing products are first combined to provide
the 'best estimates' of the biophysical variables that will constitute the training dataset.

2-Neural network calibration: a neural network is trained to estimate these 'best estimates'
from the input reflectance values as observed by specific sensors and the associated
geometrical configuration. Quality flags and quantitative uncertainties are also derived.

*3-Application of the network:* once the network is calibrated, it is run to provide estimates of
 the biophysical variables for each of the sensors considered, along with the quality flags and
 quantitative uncertainties.

Note that it would have been possible to follow more formally the scheme proposed by 94 95 (Verger et al. 2008) and later developed in (Verger et al. 2011). However this would need to use concurrently and in real time two (or more) sensors. This was not compatible with the 96 available processing capacity for GEOV1. Further, the use of a single product in the learning 97 98 database as proposed in (Verger et al. 2011) such as MODIS collection 5 would not allow 99 improvement of the biases sometimes observed, but would mainly decrease the frequency of missing data and smooth the temporal series. Finally, the proposed GEOV1 algorithm is 100 101 designed to work with no prior information on the vegetation type since Verger et al. (2008) showed that a single training across all biomes was performing similarly as multiple specific 102 103 training for each biome type.

Version définitive du manuscrit publiée dans / Final version of the manuscript published in : Remote Sensing of Environment (2013), Vol. 137, p. 299-309, DOI: 10.1016/j.rse.2012.12.027 Journal homepage: www.elsevier.com/locate/rse



#### 104

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

#### 106 2.2 Generation of training dataset

The way the training dataset is generated from already existing products is sketched in Figure
1, top box. Four main steps are identified: (1) selection of the most relevant products, (2)
setting the products on consistent spatial and temporal supports, (3) fusing the products and
(4) eventually scaling the fused products. Details of each of these steps are given in the
following.

#### 112 2.2.1 Selection of products

113 The available candidate global products listed in Table 1 are first evaluated to select the most 114 relevant ones. For this purpose, the results from previous validation exercises are used here.

The validation exercise achieved by Garrigues et al. (2008) for LAI products was showing few 115 missing values for the GLOBCARBON LAI product, in relation to its monthly temporal 116 resolution. However, many artifacts were observed, including unexpected outliers and shifts 117 in the phenology due to its coarse temporal resolution and low sampling frequency. Garrigues 118 et al. (2008) and Weiss et al. (2007) demonstrated that CYCLOPES was providing very 119 smooth temporal course while showing a saturation for LAI values larger than 4. For MODIS 120 collection 5 LAI products, only few local validation activities were reported (De Kauwe et al. 121 2011; Kraus 2008; Sprintsin et al. 2009) while Yuan et al.(2011) presents some results of a 122 123 global validation exercise focusing on the evaluation of a smoothed version of MODIS LAI to

Version définitive du manuscrit publiée dans / Final version of the manuscript published in : Remote Sensing of Environment (2013), Vol. 137, p. 299-309, DOI: 10.1016/j.rse.2012.12.027 Journal homepage: www.elsevier.com/locate/rse

reduce the spatial and temporal inconsistencies observed at the local spatial or temporal
scales. Ganguly et al. (2008a) validated an adaptation of MODIS *LAI* products for AVHRR
data, showing fair consistency with CYCLOPES *LAI* product over the sites considered.
However, Verger et al. (2011) demonstrated that MODIS was showing more shaky temporal

128 LAI profiles as compared to the CYCLOPES ones.

Among the global FAPAR products available, MODIS and GLOBCARBON correspond 129 roughly to the same definition, i.e. black-sky values at the time of the satellite overpass, i.e. 130 131 around 10:35 for MODIS aboard Terra and 10:30 for VEGETATION. However, if JRC-FAPAR corresponds also to black-sky FAPAR at the time of the satellite overpass, SEAWIFS 132 is crossing the equator around 12:20, i.e. for significantly smaller sun zenith angles as 133 compared to MODIS and VEGETATION. The CYCLOPES FAPAR product corresponds to 134 black-sky FAPAR at 10:00 illumination conditions which is a close approximation of the daily 135 integrated black-sky FAPAR value (Baret et al. 2005; Baret et al. 2007). Therefore, MODIS 136 aboard Terra, GLOBCARBON and CYCLOPES FAPAR products share very similar 137 138 definitions, while JRC-FAPAR derived from SEAWIFS is expected to show lower values because of the smaller sun zenith angles experienced close to solar noon. Further, it is limited 139 to sun zenith angles lower than 50° which may pose a problem for the higher latitudes and/or 140 part of the season. Comparison between MODIS collection 5 and CYCLOPES FAPAR 141 products was showing a good agreement, with however some overestimation of MODIS for 142 the low FAPAR values (McCallum et al. 2010). GLOCARBON and JRC-FAPAR were found 143 144 much lower in magnitude, while a good consistency in spatial and temporal trends were found between CYCLOPES and JRC-FAPAR (McCallum et al. 2010). GLOBCARBON products 145 146 were showing the largest discrepancies with all the other products.

The MODIS and CYCLOPES FAPAR products were therefore selected ensuring a good 147 consistency between LAI and FAPAR values since both MODIS and CYCLOPES provide 148 149 concurrently LAI and FAPAR products. However, CYCLOPES and MODIS products are 150 based on different assumptions and inversion techniques. CYCLOPES considers canopies as a turbid medium for all the biome types, while allowing pixels to be a mixture of pure bare soil 151 and pure vegetation patches, i.e. including some possible clumping at the landscape level. The 152 inversion of the radiative transfer model is achieved using a neural network. Conversely, 153 MODIS algorithm is biome dependent and assumes some clumping at the plant level for some 154 155 biomes including savanna and forests. The retrieval of LAI is achieved using a look-up-table inversion technique. 156

Apart from CYCLOPES FCOVER products, no other global FCOVER product is currently 157 available apart from the SAF-LAND products covering the METEOSAT disk (Camacho-de 158 Coca et al. 2006). However, several studies have pointed out that NDVI could be a good 159 proxy for FCOVER (Baret et al. 1995; Carlson and Ripley 1997; Gutman 1991). Camacho-de 160 161 Coca et al. (2006) compared several regional FCOVER products over Africa and showed that the CYCLOPES FCOVER product was very consistent with other products although a 162 significant and systematic bias was observed. This was also confirmed by Verger et al. 163 (2009). It is therefore proposed to select the CYCLOPES FCOVER original product while 164 rescaling it to provide values more consistent with ground measurements as detailed later. 165

#### 166 2.2.2 Spatial and temporal sampling for the training dataset

The training dataset is generated over a sample of sites representative of the global
distribution of vegetation types and conditions. For this purpose, the BELMANIP2 set of sites
was used. It corresponds to 420 sites located in relatively flat and homogeneous areas (at a

Version définitive du manuscrit publiée dans / Final version of the manuscript published in : Remote Sensing of Environment (2013), Vol. 137, p. 299-309, DOI: 10.1016/j.rse.2012.12.027 Journal homepage: www.elsevier.com/locate/rse

- 170 kilometric resolution over 10 x 10 km<sup>2</sup> domains). BELMANIP2 differs from the original
- 171 BELMANIP set of sites (Baret et al. 2006) by improving the global representativeness and
- 172 homogeneity. Figure 2 shows the distribution of sites and their coordinates are available at
- WWW1. Years 2003 and 2004 were selected to represent the whole seasonality as well as
- 174 significant inter-annual variability.



#### 175

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

The same spatial and temporal supports are required to allow combining several products 177 together. The MODIS LAI and FAPAR products selected were thus re-projected onto the 178 cylindrical projection system with  $1/112^{\circ}$  sampling grid used as a reference for the 179 VEGETATION, CYCLOPES and GEOV1 products. Because of the point spread function of 180 the several products considered as well as the possible geometrical uncertainties on pixel 181 localization, a 3x3 pixels spatial support was used. Note that the algorithm is trained over 3x3 182 pixels although it will ultimately apply to single pixels. This scale discrepancy is acceptable 183 because of the homogeneity of the BELMANIP2 sites. 184

185 The temporal sampling used to fuse MODIS and CYCLOPES products will be that of the CYCLOPES original products, i.e. dekadal (10 days). It will allow using directly the 186 normalized reflectance values derived from VEGETATION based on the CYCLOPES 187 188 preprocessing algorithm (Baret et al. 2007) that will also constitute the GEOV1 temporal sampling. The values to be fused at a given GEOV1 dekad were computed as follows: for 189 CYCLOPES, the product values corresponding to the GEOV1 dekad are considered; for 190 MODIS 8 days products, all the values available within  $\pm 10$  days around the GEOV1 dekad 191 are first gathered resulting in a maximum of 27 values (3 MODIS dates times 9 pixels). Note 192 193 that the CYCLOPES 30 days temporal resolution (with Gaussian weighing) is still larger than 194 that of the 16 days (2 times 8 days) of the MODIS selected values. However, widening the temporal window for MODIS resulted in artifacts on the seasonality. Then the 3x3 aggregated 195 values for each GEOV1 dekad are computed only if at least 5 valid individual values (over the 196 197 9 for CYCLOPES and 27 for MODIS) are available. Valid pixels are defined by the main and main + saturation QC (Quality Control) for MODIS, and valid input reflectance QC for 198 CYCLOPES. For MODIS, because of the relatively large variability observed over time and 199 space, values were further filtered using the difference  $\Delta = LAI_{f=90\%} - LAI_{f=70\%}$  between the 200 f=90% and f=70% percentiles (f) computed over the valid LAI values: when  $\Delta > 0.2$ 201

#### Postprint Version définitive du manus

Version définitive du manuscrit publiée dans / Final version of the manuscript published in : Remote Sensing of Environment (2013), Vol. 137, p. 299-309, DOI: 10.1016/j.rse.2012.12.027 Journal homepage: www.elsevier.com/locate/rse

corresponding to an unexpectedly large scattering of the data over homogeneous sites and 202 203 during a short time period, the cases were rejected. Finally, for the GEOV1 dates fulfilling the 204 above criterions both for CYCLOPES and MODIS products, the LAI and FAPAR values corresponding to the 70% percentile was computed for CYCLOPES and MODIS. This allows 205 minimizing the influence of possible residual cloud contamination and atmospheric effects 206 that negatively biased the product values (Chen et al. 2006). However, because of the 207 homogeneity of the sites and the short time period considered, the values once filtered as 208 described above, should be closely distributed around the median, i.e. LAI or FAPAR values at 209 50% and 70% frequencies should be very close together for a given date and site. 210

### 211 2.2.3 Fusing the products

No fusion process was applied for FCOVER since GEOV1 will derive only from CYCLOPES *FCOVER* products. Conversely, for *LAI* and *FAPAR*, MODIS and CYCLOPES products were fused to benefit from their complementarities. An optimal fusion scheme would be a weighted average between the two products, with weights,  $w_{FAPAR}$  and  $w_{LAI}$  respectively for *FAPAR* and *LAI*, being driven by the uncertainties associated to each product:

217 
$$\begin{cases} FAPAR_{fused} = FAPAR_{MOD}w_{FAPAR} + FAPAR_{CYC}(1 - w_{FAPAR}) \\ IAI_{fused} = IAI_{MOD}w_{LAI} + IAI_{CYC}(1 - w_{LAI}) \end{cases}$$
[1]

Where the subscripts '*fused*' '*MOD*', '*CYC*' correspond respectively to the fused, MODIS and 218 219 CYCLOPES products. However, the uncertainties attached to the CYCLOPES and MODIS products only refer to the theoretical performances, and model assumptions as well as the 220 structure of the uncertainties were not accounted for. Uncertainties may be also derived from 221 the comparison with ground measurements as already achieved for MODIS and CYCLOPES. 222 However, these ground measurements are not very numerous (Camacho et al. 2012; Garrigues 223 et al. 2008). Further, it is not advisable to use the validation data to calibrate an algorithm in 224 order to preserve the required independency between the calibration and the validation 225 processes. For these reasons, the weight used in the fusion between MODIS and CYCLOPES 226 227 were based on heuristic arguments.

Garrigues et al. (2008) and Weiss et al. (2007) reported that CYCLOPES *LAI* was showing
some saturation for *LAI* values around 4. Conversely, MODIS *LAI* and *FAPAR* values were
generally higher than expected for the very low vegetation amounts (Figure 4). Further, the
MODIS algorithm assigns zero values for *LAI* and *FAPAR* over pixels classified as bare,
which may pose problems in case of misclassification. It was thus proposed to fuse the
products by reducing the contribution of MODIS products for low *LAI* and *FAPAR* values and
enhancing the MODIS contribution for the larger *LAI* and *FAPAR* values as sketched in

Figure 3 with  $w_{LAI} = w_{FAPAR} = \min(1, \frac{LAI_{CYC}}{4})$ . The weight, *w*, is driven by  $LAI_{CYC}$  (Figure 3) since  $LAI_{CYC}$  appears more stable as compared to MODIS *LAI* (Verger et al. 2011). The threshold of  $LAI_{CYC}$ =4 corresponds to the value when  $LAI_{CYC}$  starts to saturate. The parallel processing applied to both *LAI* and *FAPAR* (equation 1) is expected to keep a good

239 consistency between these two variables.

Version définitive du manuscrit publiée dans / Final version of the manuscript published in : Remote Sensing of Environment (2013), Vol. 137, p. 299-309, DOI: 10.1016/j.rse.2012.12.027 Journal homepage: www.elsevier.com/locate/rse



240



#### 242 2.2.4 Scaling the fused products

The fused *FAPAR* products showed that the maximum values (at 99% percentile) are around 0.90 (Figure 4b, dashed black line) although the maximum values are expected to be close to 0.94 (Baret and Guyot 1991) corresponding to full cover dense vegetation with albedo in the PAR domain close to 0.06. Therefore, the fused values were scaled according to:

247 
$$FAPAR_{scaled} = \frac{0.94}{0.90} FAPAR_{fused}$$
 [2]

The highest  $FCOVER_{CYC}$  value was approximated by the 99% percentile value, i.e. 248  $FCOVER_{CYC}(99\%) = 0.69$  (Figure 4c). This is in agreement with results obtained by Verger 249 (2008) showing that CYCLOPES FCOVER product was underestimating actual values, 250 although being strongly linearly correlated with other regional products including the SAF-251 Land one. It is thus proposed to correct for this systematic underestimation by applying a 252 scaling factor. This factor was computed considering the highest FCOVER<sub>CYC</sub> values 253 observed that should correspond to full coverage (FCOVER=1). The 'scaled' FCOVER 254 product, FCOVER<sub>scaled</sub> used later to train the neural network will thus be computed according 255 256 to:

257 
$$FCOVER_{scaled} = \frac{1}{0.69} FCOVER_{CYC}$$
 [3]

Version définitive du manuscrit publiée dans / Final version of the manuscript published in : Remote Sensing of Environment (2013), Vol. 137, p. 299-309, DOI: 10.1016/j.rse.2012.12.027 Journal homepage: www.elsevier.com/locate/rse



#### 258

# Figure 4. Cumulated frequency of CYCLOPES (green solid line), MODIS (red solid line) fused (dashed black line for *FAPAR*) and scaled (black solid line) *LAI* (a), *FAPAR*(b) and *FCOVER* (c) products as observed over the 420 BELMANIP2 sites during years 2003-2004.

263 Note that oppositely to FCOVER and FAPAR products, no specific theoretical upper limit exists for LAI since values larger than 10 are often reported at least at the local scale (Scurlock 264 et al. 2001). Regarding the theoretical lower bound (*LAI=FAPAR=FCOVER=0*) 265 266 corresponding to bare soil, the cumulated frequencies displayed in (Figure 4b) do not show particular problems for the fused product. However MODIS products show an offset of 0.25 267 for LAI (Figure 4a) and 0.18 (Figure 4b) for FAPAR values, confirming the previous 268 observations of McCallum et al. (2010). This bias was corrected by the fusion process used 269 for deriving GEOV1 where CYCLOPES LAI and FAPAR products contribute the most for 270 these low vegetation amounts. 271

#### 272 2.3 Training the neural networks

The inputs and outputs of the training data set and the neural network architecture andlearning process are described in this section.

#### 275 2.3.1 Inputs

The inputs of the neural network correspond to the bidirectional reflectance factor (BRF) as 276 measured by VEGETATION aboard SPOT. They correspond to directionally normalized top 277 of canopy reflectance in the red, NIR and SWIR bands as derived from the CYCLOPES L3a 278 products. The preprocessing steps include cloud screening, atmospheric correction based on a 279 climatology of aerosols, and BRDF normalization using a robust fit of Roujean's model 280 (Hagolle et al. 2004; Roujean et al. 1992). Details about the processing from L0 (raw signal) 281 to L3a can be found in (Baret et al. 2007). However, the original CYCLOPES L3a products 282 were reprocessed to benefit from updated radiometric calibration coefficients for the 283 VEGETATION sensors. The blue band was not considered here since it brings only little 284 extra information on the surface as compared to the red, NIR and SWIR bands, while being 285 very sensitive to errors in the atmospheric correction (Bacour et al. 2006). To match the 286 spatial support used for the 'best estimates' described earlier, the median was computed over 287 the valid pixels within the 3x3=9 potential ones. The median for each band was preferred 288 here, since the use of the 70% percentile for each band would result in possible spectral 289 inconsistencies because the sensitivity of reflectance to canopy variables is different between 290 bands. Note however that the CYCLOPES L3a products show a good stability over the 291 292 BELMANIP2 sites which are relatively homogeneous as earlier demonstrated by Verger et al. 293 (2008).

In addition to the BRF in the red, NIR and SWIR bands, the algorithm used the cosine of the median value of the sun zenith angle ( $\theta_s$ ) corresponding to the valid observations during the 30 days window over which the directionally normalized top of canopy reflectance values were composited.

#### 298 2.3.2 Outputs

The outputs correspond to the targeted LAI, FAPAR and FCOVER variables. To further verify 299 that the resulting 'best estimates' were consistent, they were plotted against NDVI=(BRF<sub>NIR</sub>-300 BRF<sub>Red</sub>)/(BRF<sub>NIR</sub>+BRF<sub>Red</sub>) values computed from the CYCLOPES BRF products in the red 301 (BRF<sub>Red</sub>) and in the near infrared (BRF<sub>NIR</sub>). Cases with NDVI<0.05 were rejected since these 302 values are not expected over bare soil or vegetation pixels (results derived from NDVI values 303 computed for VEGETATION sensor bands based on a large data base of soil reflectance 304 305 available at WWW2 and described in Liu et al. (2002). Then, for each class of NDVI values 306 (20 classes over the [0, 1] domain of variation), the cases with FAPAR or LAI values lower (respectively higher) than the 5% percentile (respectively 95%) were rejected. This allowed to 307 further improve the consistency of outputs with input reflectance values through NDVI as 308 attested by Figure 5. Relatively few cases were observed for LAI in between 3 and 4. This is 309 explained by the fact that forests that are the more likely to show such median to high LAI 310 values (Weiss et al. 2007) represent only 25% of the global land surfaces. Further, many 311 forest sites are frequently covered by clouds in addition to snow cover and poor illumination 312 conditions that are frequently observed in winter for the high northern latitudes. Although this 313 situation is not ideal, the neural networks should be able to interpolate efficiently between the 314 cases available on both sides of this area. 315



316

# Figure 5. Relationships between *NDVI* and both *FAPAR* (a) and *LAI* (b) 'best estimates'. The red points correspond to the cases rejected that are outside the [5% 95%] percentile range.

The consistency of the output variables was further evaluated. As expected, the relationship
between *LAI* and *FAPAR* was keeping very consistent as compared to the original
CYCLOPES and MODIS products (Figure 6). The same consistency is observed between *LAI*and *FCOVER* as well as between *FAPAR* and *FCOVER* (results not presented for the sake of
brevity).

Version définitive du manuscrit publiée dans / Final version of the manuscript published in : Remote Sensing of Environment (2013), Vol. 137, p. 299-309, DOI: 10.1016/j.rse.2012.12.027 Journal homepage: www.elsevier.com/locate/rse



325

Manuscrit d'auteur / Author manuscript

#### Figure 6. Relationship between LAI and FAPAR for CYCLOPES (CYCV31 a), MODIS 326 (MODC5 b) and 'best estimates' (c) as observed over the 420 BELMANIP2 sites during 327 328 2003-2004 period (14200 points).

#### Neural network architecture and learning process 329 2.3.3

The previously described dataset is finally made of 14200 cases where consistent top of 330 331 canopy directionally normalized reflectance values are paired with 'best estimates' of LAI, FAPAR and FCOVER values. This represents roughly 47% of the total 30200 data potentially 332 existing over the 420 sites during the 2003-2004 period (72 dekads). The available data was 333 randomly split into a training data set made of 90% of the data available, and a test data set 334 (10% of the data) used for testing the hyper-specialization of the training process and 335 evaluating the theoretical performances. The large fraction of data used for the training 336 process allows getting a better representativeness of surface types and conditions considering 337 the limited time period (2 years) and sites considered (420 sites) and the large fraction of 338 missing data. The inputs and outputs are normalized to prevent possible numerical problems 339 during the training process. Normalization is achieved by scaling between -1 and +1 the range 340 341 of variation of input and output values according to:

342 
$$x_{norm} = 2 \frac{x - x_{\min}}{x_{\max} - x_{\min}} - 1$$
 [4]

343 where x represents either the inputs or outputs,  $x_{min}$  and  $x_{max}$  are respectively the minimum and maximum values of x and  $x_{norm}$  is the corresponding normalized value. 344

A back-propagation network architecture has been used since it proves very efficient in 345 similar problems (Baret et al. 2007; Verger et al. 2010a). It is made of one hidden layer of 5 346 tangent-sigmoidal neurones, and one output layer made of a single linear neurone. This 347 348 architecture includes 31 coefficients to be adjusted (25 synaptic weights and 6 biases) providing more than 400 cases per coefficient to be adjusted. The Levenberg-Marquardt 349 optimization algorithm (Ngia and Sjoberg 2000) is used for adjusting the synaptic weights and 350 neuron bias to get the best agreement between the output simulated by the network and the 351 corresponding value of canopy biophysical variable in the training dataset. The initial values 352 of the weights and biases were set to a random value between -1.0 and +1.0. Three networks 353 were trained in parallel for each targeted variable with variation in the initial values. The one 354 providing the best performances over the test dataset was selected. Results obtained during 355 the training process showed that the three parallel networks were performing very similarly, 356 indicating a robust training process. A dedicated neural network for each variable was 357 preferred here to using a unique network for estimating concurrently the three output variables 358

Version définitive du manuscrit publiée dans / Final version of the manuscript published in : Remote Sensing of Environment (2013), Vol. 137, p. 299-309, DOI: 10.1016/j.rse.2012.12.027 Journal homepage: www.elsevier.com/locate/rse

- because the associated architecture is simpler, leading to an easier training process while still
- 360 providing very good consistency between LAI, FAPAR and FCOVER estimated variables,
- similarly to what is observed on the training dataset (Figure 6).



362

#### Figure 7. Structure of the neural network used to derive *LAI*, *FAPAR* and *FCOVER* (the 'Var' box) from VEGETATION input reflectance in red (B2), NIR (B3) and SWIR as well as illumination geometry $(\cos(\theta_s))$ .

The theoretical performances were evaluated over the test dataset. It shows that the training was very efficient for the three variables. The dispersion around the 1:1 line is very small and no bias is observed over the whole range of variation of the three variables (Figure 8). Note that *LAI* shows less accurate estimates for 3 < LAI < 4 because of the slightly lower number of data available for this *LAI* range as seen in Figure 5.

371



Figure 8. Comparison between the 'best estimates' of *LAI* (a), *FAPAR* (b) and *FCOVER* (c) products and the values estimated from the trained neural networks (called NNT estimates). The solid thick line corresponds to the median value. Dark, medium and light gray areas correspond respectively to [25% 75%] [10% 90%] and [1% 99%] of the estimated cases for each class of actual values (20 classes are considered, from 0 to the maximum of actual values). The data correspond to the test data set (n=1420).

Further evaluation was achieved to check the specific performances of the network for the main great vegetation types. Results show (Figure 9) that no biases are observed in the estimated *LAI*, except a slight underestimation (LAI<sub>est</sub>-LAI<sub>train</sub> $\approx$ 0.1). The positive and negative residuals are always well balanced. The residuals are increasing in absolute value from the low *LAIs* (Non Forest) up to the largest one (EBF). In case of EBF, most of the variability is imputed to remaining cloud contamination.

385

Version définitive du manuscrit publiée dans / Final version of the manuscript published in : Remote Sensing of Environment (2013), Vol. 137, p. 299-309, DOI: 10.1016/j.rse.2012.12.027 Journal homepage: www.elsevier.com/locate/rse



Figure 9. Difference between the LAI estimated from the NNT (LAI<sub>est</sub>) and the 'best estimates'
 LAI value used for the training (LAI<sub>train</sub>) for the four main vegetation types : Evergreen
 Broadleaf Forest (EBF), Needle leaf Forest (ENF), Deciduous Broadleaf Forest (DBF) and Non Forest. The red line, blue box and black whiskers represent respectively 50%, [25%,75%] and
 [5%-95%] percentiles. Outliers are indicated by stars.

#### 391 2.4 Associated uncertainties and quality assessment

All the quality control flags associated to the top of canopy reflectance values are available along with the products. They describe the nature of the surface (land/sea), the presence of snow, the possible contamination by clouds or cloud shadow, the aerosol characteristics used for the atmospheric correction, and the possible saturation of the radiometric signal. Two additional qualitative assessment criterions more directly dedicated to the biophysical products are provided along with a quantitative estimation of the associated uncertainties. The way they are computed is described here after.

#### 399 2.4.1 Input out of range

Since the algorithm is based on a learning machine approach, it is important to verify whether 400 the inputs of a given observation keeps within the range of variation of the training dataset 401 called here the definition domain. If this condition is not fulfilled, the network will run in 402 extrapolation mode, with no warranty about the realism of the outputs. The definition domain 403 is limited by the convex hull formed in the BRF feature space by the cases used in the training 404 405 process (Figure 10). For the sake of simplicity and ease of implementation, the 3D feature space formed by B2, B3 and SWIR bands was gridded by dividing the range of variation of 406 each band (Figure 10 and Table 2) into 10 equally spaced classes. The ensemble of cells 407 containing data used for the training form the definition domain. When the input BRF values 408 are outside the definition domain, i.e. outside cells containing data used for the training, an 409 'input out of range' flag is raised. Note that the sun zenith angle was not included in the 410 description of the definition domain. This would have induced increased complexity for 411 marginal gain since the definition domains corresponding to several sun zenith angles are 412 largely overlapping. 413

Version définitive du manuscrit publiée dans / Final version of the manuscript published in : Remote Sensing of Environment (2013), Vol. 137, p. 299-309, DOI: 10.1016/j.rse.2012.12.027 Journal homepage: www.elsevier.com/locate/rse



414

Figure 10. Definition domain for input VEGETATION Top of canopy directionally normalized
BRF values in B2, B3 and SWIR bands. Axes are scaled between 1 to 10, corresponding to the 10
classes of reflectance values ranging between the minimum value (class 1) and the maximum
value (class 10). The cells containing data from the training dataset are represented in black.
They form the definition domain.

Table 2. Range of variation of BRF values (Minimum and maximum) observed in the training
dataset and used to compute the definition domain.

	Min	Max
B2	0.000	0.429
B3	0.036	0.547
SWIR	0.000	0.648

#### 422 2.4.2 Output out of range

The physical limits for the three variables are described in Table 3. However, for LAI, the 423 424 upper limit is not a physical limit, but a value just slightly higher than the maximum value 425 that can be reached by the MODIS and CYCLOPES original products. Because of the several sources of uncertainties associated to the inputs, the algorithm calibration process including 426 427 uncertainties attached to the original MODIS and CYCLOPES products, a tolerance is set for the extreme values: when the neural network provides biophysical variable estimates outside 428 their definition range the value will be always set to the closest bound of the range, i.e. either 429 the minimum or the maximum values (Table 3). The product uncertainty value will be also set 430 to its maximum value. However, The output status flag is thus raised only when the output is 431 outside the output range enlarged by the tolerance values  $[P^{tol}_{min}, P^{tol}_{max}]$  as defined in Table 432 3. 433

# Table 3. Minimum, Maximum, Resolution and Tolerance values used to raise the output out of range flag.

	Min	Max	Resol.	Tol <sub>min</sub>	Tol <sub>max</sub>
LAI	0	7.0	0.01	-0.2	7.20
FAPAR	0	1.0	0.01	-0.05	1.05
FCOVER	0	1.0	0.01	-0.05	1.05

#### 436 2.4.3 Product uncertainties

437 The uncertainties associated to each biophysical variable are computed over the training 438 dataset: for each case in the training dataset, a RMSE value is computed over the biophysical 439 variables that have their corresponding input BRF values within the uncertainty domain 440 (Figure 11). Further, the sun zenith angles selected have also to be within  $\pm 5^{\circ}$  around the 441 direction of the case considered. Uncertainties on reflectance are derived from Baret et al.

Manuscrit d'auteur / Author manuscript

- 442 (2007): the standard deviation,  $\sigma$ , is computed for each waveband according to the uncertainty
- 443 model reported in Table 4:
- 444  $\sigma = a + b \cdot BRF$  [5]
- 445 Where *a* and *b* are coefficients.

Observation

Confidence interval (Sigma)



446

# Figure 11. Scheme showing how the uncertainties attached to the products were computed

Table 4. Values of the uncertainty model used for the input BRFs. From Baret et al.
(2007).

	B2	B3	SWIR
a	0.005	0.003	0.005
b	0.05	0.03	0.03

A neural network is then trained to relate these computed RMSE values to the corresponding 451 input variables. This learning process is similar to what was previously achieved for the 452 453 product estimation itself and the network has the same architecture as the one used for the derivation of the products (Figure 7). The same normalization as used for the variables is also 454 applied for the inputs and outputs of the neural network dedicated to uncertainties. Results 455 (Figure 12) show that the model of uncertainties is relatively robust for LAI, FAPAR and 456 FCOVER, with however some degradation of performances for the larger values of the 457 biophysical variables. 458

Version définitive du manuscrit publiée dans / Final version of the manuscript published in : Remote Sensing of Environment (2013), Vol. 137, p. 299-309, DOI: 10.1016/j.rse.2012.12.027 Journal homepage: www.elsevier.com/locate/rse





Figure 12. Theoretical performances of the neural network model used to describe RMSE values
for *LAI* (a) *FAPAR* (b) and *FCOVER* (c) from input reflectance values and geometry of
observation.

### 463 **3** Operational production and dissemination

The GEOV1 processing line was derived from the CYCLOPES V3.1 processing line which 464 465 has been used to generate the 1999-2007 time series of CYCLOPES products (WWW3). It has been adapted to consider the specificities of the GEOV1 products in terms of algorithm, 466 and output format requested by the users. It has also been consolidated to fit the software 467 conventions (general coding rules, computer platform issues, filename and directory 468 conventions) defined by the operational processing center. At different stages of the 469 development process, reviews have been carried out to check that the several input 470 specifications were answered in the design of the processing line. Further, a rigorous 471 472 methodology of validation based upon unit and integration tests, as well as scientific analysis of the output using tools including visual control and statistical metrics has been applied. 473

474 Before its integration in the operational processing center, the processing line has been run to generate 2 years (2003–2004) of GEOV1 demonstration products used to perform the 475 476 validation exercise presented in Camacho et al. (2012). They have also been supplied to 477 Geoland2 users who checked the consistency with their requirements before starting the operational production of real time and the whole 12 years time series processing of 478 observations. The processing line has been optimized to ensure operational management with 479 480 permanent quality control while benefiting from the variable computation resources available for parallel processing. 481

The GEOV1 products are generated in multi-band hdf5 format (the variable, its uncertainty, the quality flag, the number of input observations, the land-sea mask) and in tiles of 10°x10° covering the land surfaces of the whole globe. They are available in open access through the Geoland2 web platform (WWW4) where users can browse the catalogue, order the products after registration, and subscribe to receive the products. The GEOV1 products are also disseminated via the Eumetcast system to African and South American users.

### 488 **4** CONCLUSION

The GEOV1 *LAI*, *FAPAR* and *FCOVER* products capitalize on the efforts undertaken this last
decade in the development and validation of biophysical products from medium resolution
observations. The pragmatic approach used here is based on the fusion of CYCLOPES and

MODIS products that were demonstrated to perform the best. However, their deficiencies 492 493 observed respectively for low LAI values for MODIS and high LAI values for CYCLOPES 494 have been corrected in the fusion process. Further, the resulting FAPAR and FCOVER products have been scaled to reach the theoretical upper limit expected for high vegetation 495 amounts. Finally, one of the main advantage of the GEOV1 algorithm is that it did not 496 explicitly use a biome classification that introduces some spatial inconsistencies and 497 sometimes temporal inconsistencies (when class assignment change unexpectedly) as outlined 498 by Yang et al. (2006a) for MODIS. 499

The approach undertaken here was not calibrated using ground validation measurements, 500 preserving the required independency between calibration and validation processes. However, 501 this led to use a more pragmatic approach based on heuristic arguments that explains the 502 sometimes subjective selection of criterions used in the development of the GEOV1 algorithm 503 and several alternative solutions could have been proposed. Nevertheless the GEOV1 504 products resulting from the largely inductive approach undertaken yield robust, consistent and 505 506 accurate estimates of these key biophysical variables as demonstrated in Camacho et al. (2012). The validation shown in the companion paper demonstrates that GEOV1 significantly 507 improved the performances of currently existing products, both regarding accuracy and spatial 508 as well as temporal consistency. It provides quantified estimates of uncertainties, although 509 these are simply derived from the training dataset that reflects mostly the sensitivity of the 510 product to input reflectance values. Further investigations should be directed towards a better 511 512 quantification of the uncertainties, including 'model' assumptions since both MODIS and CYCLOPES are based on radiative transfer model inversion. 513

514 A clear understanding of the actual definition of the retrieved variables and their consistency with application requirements is one of the key aspects to consider when deriving biophysical 515 variables from remote sensing observations. For FCOVER, no particular question is raised 516 517 since the FCOVER definition is simple and clear if restricted to the green vegetation elements 518 including overstory and understory. For FAPAR, because of the good consistency between MODIS and CYCLOPES definitions, the output product is also well defined: black sky 519 520 FAPAR (green element including over and understory) at 10:15 (actually between 10:00 and 10:30). For LAI, the question is more complex since the original MODIS and CYCLOPES 521 LAI products are defined differently, at least with regards to the assumptions embedded in the 522 radiative transfer models used. For MODIS, clumping at the tree scale is accounted for, but 523 not at the shoot, nor at the landscape scales. For CYCLOPES, clumping is accounted for at 524 the landscape scale only. However, its effect will be significant mainly for the larger LAI 525 values as discussed by Garrigues et al. (2006). Therefore, GEOV1 LAI product will 526 marginally account for landscape clumping as demonstrated by the good consistency between 527 MODIS and CYCLOPES LAI products for low to medium LAI values (Camacho et al. 2012). 528 However the special case of savannas with relatively low LAI values and significant clumping 529 530 (Ryu et al. 2010). In this case, the radiation interaction between plants that are separated by significant distances is minimal, and clumping described at the plant level in MODIS may be 531 approximated at the landscape scale as in CYCLOPES. Clumping at the shoot scale will not 532 533 be accounted for in GEOV1 LAI product since neither MODIS, nor CYCLOPES account for it. Conversely, clumping at the plant scale was accounted mainly for forests and savannas in 534 the original MODIS product but not in the CYCLOPES ones. However, clumping is mainly 535 536 observed for the higher LAI values (overlap between leaves within the tree volume when the 537 number of leaves increases) (Rochdi 2003). Therefore, clumping at the plant scale as described in MODIS products should be partly preserved in the GEOV1 LAI product. The 538 539 validation achieved in the companion paper (Camacho et al. 2012) over a limited number of

sites demonstrates that the GEOV1 *LAI* products are close to the values estimated fromindirect techniques and that include the clumping effect.

These products may thus be used in a range of applications including those targeted for the 542 543 Essential Climate Variables and fulfill the needs for GMES and the GCOS task dedicated to 544 the operationalization of the generation of FAPAR and LAI products. The GEOV1 products initiate a service whose sustainability is planned within the GMES Land Monitoring Core 545 Service by adapting the algorithm, and the processing chain, to AVHRR/Metop, PROBA-V, 546 547 and Sentinel-3 missions. Further, it will be completed backward using the AVHRR data as processed by Vermote et al. (2010) to get a long and consistent time series of more than 30 548 549 years.

Although the development of this series of biophysical products should represent an important 550 step towards a more effective use of remote sensing observations, improvements are expected 551 mainly through several aspects. Clouds constitute obviously the major limitation of optical 552 systems that suffer from large areas and periods without data. Recent studies demonstrate that 553 the fusion between several sensors improves data continuity (Hagolle et al. 2005; Verger et al. 554 2010b; Yang et al. 2006b). Alternative approaches based on enhanced time series processing 555 may help removing outliers, filling gaps due to missing observations and smoothing the 556 temporal profiles (Kandasamy et al. 2013; Verger et al. 2012). However, the main limitation 557 in such global products comes mainly from the little a priori information available and 558 required to regularize the inversion process (Combal et al. 2002). The use of global 559 classification is likely to be insufficient because of its limited accuracy (Defourny et al. 2009; 560 Herold et al. 2008; Yang et al. 2006a) and because the variability in canopy architecture and 561 562 optical properties within a given class is probably as large as between classes when the seasonal variability is considered. Further, the often mixed nature of kilometric pixels poses 563 both a scaling issue and the difficulty to identify the several co-existing surface patches. 564 565 These problems call for an improved spatial resolution that will allow resolving most of the 566 vegetation patches and will authorize identifying the corresponding vegetation type from the past observations and use it as prior information. Such systems are currently being available 567 with hectometric resolutions, such as the PROBA-V (300m daily), Sentinel-3 (300 m every 2 568 days), VIIRS (370m daily). However, decametric systems such as Sentinel 2 or LDCM in 569 combination with the previous hectometric ones would probably provide the most efficient 570 571 observation system.

#### 573 Acknowledgments

572

The research leading to these results has received funding from the European Community's
Seventh Framework Program (FP7/2007-2013) under grant agreement n°218795.

- 576 WWW references
- 577 WWW1: http://calvalportal.ceos.org/cvp/web/guest/olive.
- 578 WWW2: https://www4.paca.inra.fr/emmah\_eng
- 579 WWW3: http://postel.mediasfrance.org
- 580 WWW4: http://www.geoland2.eu
- 581 **References**

582 583 584	Avissar, R., & Pielke, R.A. (1989). A parameterization of heterogeneous land surfaces for atmospheric numerical models and its impact on regional meteorology. <i>Monthly</i> <i>Weather Review</i> , 117, 2113-2136
585 586 587	Bacour, C., Baret, F., Béal, D., Weiss, M., & Pavageau, K. (2006). Neural network estimation of LAI, fAPAR, fCover and LAIxCab, from top of canopy MERIS reflectance data: principles and validation. <i>Remote Sensing of Environment</i> , 105, 313-325
588 589 590	<ul> <li>Baret, F., Bacour, C., Béal, D., Weiss, M., Bruniquel, V., &amp; Regner, P. (2005). Algorithm Theoretical Basis Document for MERIS Top of Atmosphere Land Products (TOA_VEG). In (p. 32). Avignon: INRA-CSE</li> </ul>
591 592 593	Baret, F., Clevers, J.G.P.W., & Steven, M.D. (1995). The robustness of canopy gap fraction estimates from red and near infrared reflectances: a comparison of approaches. <i>Remote sensing of environment</i> , <i>54</i> , 141-151
594 595	Baret, F., & Guyot, G. (1991). Potentials and limits of vegetation indices for LAI and APAR assessment. <i>Remote Sensing of Environment, 35</i> , 161-173
596 597 598 599	Baret, F., Hagolle, O., Geiger, B., Bicheron, P., Miras, B., Huc, M., Berthelot, B., Weiss, M., Samain, O., Roujean, J.L., & Leroy, M. (2007). LAI, fAPAR and fCover CYCLOPES global products derived from VEGETATION. Part 1: Principles of the algorithm. <i>Remote Sensing of Environment</i> , 110, 275-286
600 601 602 603 604	Baret, F., Morissette, J., Fernandes, R., Champeaux, J.L., Myneni, R., Chen, J., Plummer, S., Weiss, M., Bacour, C., Garrigue, S., & Nickeson, J. (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. <i>IEEE Transactions on Geoscience and Remote Sensing</i> , 44, 1794-1803
605 606 607	Baret , F., Nightingale, J., Garrigues, S., & Nickeson, J. (2009). Report on the CEOS Land Product Validation sub-group meeting Missoula, Montana, June 15th 2009. <i>The Earth</i> <i>Observer</i> , 21, 26-30
608 609 610 611 612	<ul> <li>Camacho-de Coca, F., Jiménez-Muñoz, J.C., Martínez, B., Bicheron, P., Lacaze, R., &amp; Leroy, M. (2006). Prototyping of the fCover product over Africa based on existing CYCLOPES and JRC products for VGT4Africa,. In J. Sobrino (Ed.), <i>Proceedings of the RAQRS'II 2nd International Symposium on Recent Advances in Remote Sensing</i>, (pp. 722-726.). Valencia (Spain): Univ. Valencia</li> </ul>
613 614 615 616	Camacho, F., Cernicharo, J., Lacaze, R., Baret, F., & Weiss, M. (2012). GEOV1: LAI, FAPAR Essential Climate Variables and FCover global time series capitalizing over existing products. Part 2: Validation and inter-comparison with reference products. <i>Remote</i> <i>Sensing of Environment, submitted</i>
617 618	Carlson, T.N., & Ripley, D.A. (1997). On the relationship between NDVI, fractional cover and leaf area index. <i>Remote Sensing of Environment</i> , 62, 241-252
619 620	Chen, J.M., & Black, T.A. (1992). Defining leaf area index for non-flat leaves. <i>Plant, Cell and Environment, 15</i> , 421-429

621	Chen, J.M., Deng, F., & Chen, M. (2006). Locally adjusted cubic spline capping for
622	reconstructing seasonal trajectories of a satellite-derived surface parameter. <i>IEEE</i>
623	Transactions on Geoscience and Remote Sensing, 44, 2230-2238
624	Combal, B., Baret, F., Weiss, M., Trubuil, A., Macé, D., Pragnère, A., Myneni, R.,
625	Knyazikhin, Y., & Wang, L. (2002). Retrieval of canopy biophysical variables from bi-
626	directional reflectance data. Using prior information to solve the ill-posed inverse
627	problem. Remote Sensing of Environment, 84, 1-15
628	De Kauwe M.G. Disney M.I. Quaife T. Lewis P. & Williams M. (2011) An assessment
620	of the MODIS collection 5 leaf area index product for a region of mixed coniferous
620	forest Remote Sensing of Environment 115, 767
030	Torest. Remote Sensing of Environment, 115, 767
631	Defourny, P., Schouten, L., Bartalev, S., Bontemps, S., Cacetta, P., de Wit, A.J.W., di Bella,
632	C., Gerard, B., Giri, C., Gond, V., Hazeu, G.W., Heinimann, A., Herold, M., Knoops, J.,
633	Jaffrain, G., Latifovic, R., Lin, H., Mayaux, P., Mucher, S.A., Nonguierma, A., Stibig,
634	H.J., Van Bogaert, E., Vancutsem, C., Bicheron, P., Leroy, M., & Arino, O., (2009).
635	Accuracy Assessment of a 300 m Global Land Cover Map: the GlobCover Experience.
636	In, Proceedings of the 33rd International Symposium on Remote Sensing of
637	Environment. Stresa (Italy)
629	Dang E. Chan, I.M. Chan, M. & Disale, I. (2006). Algorithm for global loof area index
638	Deng, F., Chen, J.M., Chen, M., & Pisek, J. (2000). Algorithm for global leaf area midex
639	<i>A</i> 4 2210 2220
640	44, 2219-2229
641	Ganguly, S., Samanta, A., Schull, M.A., Shabanov, N.V., Milesi, C., Nemani, R.R.,
642	Knyazikhin, Y., & Myneni, R.B. (2008a). Generating vegetation leaf area index earth
643	system data record from multiple sensors. Part 1: Theory. Remote Sensing of
644	Environment, 112, 4318–4332
645	Ganguly S. Schull M.A. Samanta A. Shahanov N.V. Milesi C. Nemani R.R.
646	Knyazikhin Y & Myneni R B (2008b) Generating vegetation leaf area index earth
647	system data record from multiple sensors. Part 2: Evaluation. Ramota Sansing of
649	Environment 112 A333
040	Environment, 112, 4555
649	Garrigues, S., Allard, D., Baret, F., & Weiss, M. (2006). Influence landscape spatial
650	heterogeneity on the non-linear estimation of leaf area index from moderate spatial
651	resolution remote sensing data. Remote Sensing of Enviroment, 105, 286-298
652	Garrigues, S., Lacaze, R., Baret, F., Morisette, J., Weiss, M., Nickeson, J., Fernandes, R.,
653	Plummer, S., Shabanov, N.V., Myneni, R., & Yang, W. (2008). Validation and
654	Intercomparison of Global Leaf Area Index Products Derived From Remote Sensing
655	Data. Journal of Geophysical Research, 113
656	CCOS (2006) CCOS 107 Supplemental datails to the satellite based component of the
000 657	"implementation plan for the global observing system for alimate in support of the
057	LINECCC" In (p. 102). Conque (Switzerland): CCOS/WMO
658	UNFULC . In (p. 105). Geneve (Switzerland): GUOS/WMO
659	Gobron, N., Pinty, B., Aussedat, O., Chen, J.M., Cohen, W.B., Fensholt, R., Gond, V.,
660	Huemmrich, K.F., Lavergne, T., Melin, F., Privette, J.L., Sandholt, I., Taberner, M.,
661	Turner, D.P., Verstraete, M.M., & Widlowski, J.L. (2006). Evaluation of fraction of

662	absorbed photosynthetically active radiation products for different canopy radiation
663	transfer regimes: Methodology and results using Joint Research Center products derived
664	from SeaWiFS against ground-based estimations. Journal of Geophysical Research-
665	Atmospheres, 111
666	Gutman, G.G. (1991). Vegetation indices from AV HRR data: an update and future prospects. ,.
667	Remote Sensing of Environment, 55, 121-150
668	Hagolle, O., Lobo, A., Maisongrande, P., Cabot, F., Duchemin, B., & De Pereyra, A. (2005).
669	Quality assessment and improvement of temporally composited products of remotely
670	sensed imagery by combination of VEGETATION 1 and 2 images. <i>Remote Sensing of</i>
671	Environment, 94, 172-186
672	Hagella O. Nicolas I. M. Equania P. Cabot F. & Hanry P. (2004) Absolute calibration
672	nagone, O., Nicolas, JM., Foughie, B., Cabol, F., & Henry, P. (2004). Absolute canoration
673	of vEGETATION derived from an interband method based on the Sun gint over ocean. EEET $(1.172)$ $(1.17$
674	<i>TEEE Transactions on Geoscience and Remote Sensing</i> , 42, 14/2 -1481
675	Herold, M., Mayaux, P., Woodcock, C.E., Baccini, A., & Schmullius, C. (2008). Some
676	challenges in global land cover mapping: An assessment of agreement and accuracy in
677	existing 1 km datasets. Remote Sensing of Environment, 112, 2538
670	Kandacamy S. Barat F. Vargar A. Navaux P. & Waiss M (2013) A comparison of
670	smoothing and gap filling methods for time series of remote sensing observations
690	Application to MODIS I AI products <i>Biogeospiences</i> 0, 17052, 17007
080	Application to MODIS LAI products. <i>Biogeosciences</i> , 9, 17055-17097
681	Kraus, T. (2008). Ground-based Validation of the MODIS Leaf Area Index Product for East
682	African Rain Forest Ecosystems. In. Erlangen: Universitätsbibliothek der Universität
683	Erlangen-Nürnberg
681	Lacaze R. Atzberger C. Bartholomá F. Combal B. Calvet I.C. Lefàvre V. & Olsson B.
004 60E	(2000) BioDar Usar Daquiraments. In Toulousa
065	(2009). Biol al Osel Requirements. In. Toulouse
686	Liu, G.R., Lin, T.H., & Kuo, T.H. (2002). Estimation of aerosol optical depth by applying the
687	optimal distance number to NOAA AVHRR data. Remote Sensing of Environment, 81,
688	247-252
680	McCallum I Wagner W Schmullius C Shvidenko A Obersteiner M Fritz S &
600	Nilsson S (2000) Satellite based terrestrial production afficiency modeling. Carbon
601	Ralance and Management 4:8
091	Dalance una Management, 4.8
692	McCallum, I., Wagner, W., Schmullius, C., Shvidenko, A., Obersteiner, M., Fritz, S., &
693	Nilsson, S. (2010). Comparison of four global FAPAR datasets over Northern Eurasia
694	for the year 2000. Remote Sensing of Environment, 114, 941
695	Morisette I Baret F Privette II. Myneni R B Nickeson I Garriques S Shahanov N
695	Weiss M Fernandes R Leblanc S Kalacska M Sanchez-Azofeifa GA Chubey
697	M Rivard B Stenherg P Rautiainen M Voinio P Manninen T Pilant D Lewis
692	T liames I Colombo R Meroni M Rusetto I Cohen W Turner D Warner D
600	Petersen GW Seufert G & Cook R (2006) Validation of global moderate resolution
700	LAI Products: a framework proposed within the CEOS I and Product Validation
701	subgroup IEEE Transactions on Geoscience and Remote Sensing 44 1804-1817
,01	subgroup. IEEE Transactions on Geoscience and Remote Sensing, 77, 1007-1017
	24

7 7 7	702 703 704	Ngia, L.S.H., & Sjoberg, J. (2000). Efficient training of neural nets for nonlinear adaptive filtering using a recursive Levenberg-Marquardt algorithm. <i>IEEE Transactions on Signal Processing</i> , 48, 1915-1927
7777	205 206 207	Rochdi, N. (2003). Un modèle générique d'agrégation des feuilles dans un couvert végétal: application à la simulation du transfert radiatif. In (p. 169). Paris (France): Institut National Agronomique
7 7 7	708 709 710	Roujean, J.L., Leroy, M., & Deschamps, P.Y. (1992). A bidirectional reflectance model of the Earth's surface for the correction of remote sensing data. <i>Journal of Geophysical</i> <i>Research</i> , 97, 20455-20468
7	'11 '12 '13	<ul> <li>Ryu, Y., Sonnentag, O., Nilson, T., Vargas, R., Kobayashi, H., Wenk, R., &amp; Baldocchi, D.D. (2010). How to quantify tree leaf area index in an open savanna ecosystem: A multi-instrument and multi-model approach. <i>Agricultural and Forest Meteorology</i>, 150, 63-76</li> </ul>
7	'14 '15 '16	Scurlock, J.M.O., Asner, G.P., & Gower, S.T. (2001). Worldwide Historical Estimates and Bibliography of Leaf Area Index, 1932-2000. In. Oak Ridge, Tennessee, U.S.A.: Oak Ridge National Laboratory
7777	717 718 719 720	Sprintsin, M., Karnieli, A., Berliner, P., Rotenberg, E., Yakir, D., & Cohen, S. (2009). Evaluating the performance of the MODIS Leaf Area Index (LAI) product over a Mediterranean dryland planted forest. <i>International Journal of Remote Sensing</i> , 30, 5061 - 5069
7	21 22 23	Verger, A. (2008). Analisi comparativa d'algorismes operacionals d'estimacio de parametres biofisics de la coberta vegetal amb teledeteccio. In, <i>Departament de fisica de la terra i termodinamica</i> (p. 277). Valencia (Spain): Universitat de Valencia
7777	24 25 26 27	<ul> <li>Verger, A., Baret, F., García-Haro, F.J., Camacho, F., &amp; Meliá, J. (2010a). Consistency of vegetation estimates from SEVIRI/Meteosat observations and operational algorithms. In J. Sobrino (Ed.), <i>Third International Symposium on Recent Advances in Quantitative Remote Sensing</i>. Torrent (Spain)</li> </ul>
7	28 29 30	Verger, A., Baret, F., & Weiss, M. (2008). Performances of neural networks for deriving LAI estimates from existing CYCLOPES and MODIS products. <i>Remote Sensing of Environment</i> , 112, 2789-2803
7777	'31 '32 '33 '34	Verger, A., Baret, F., & Weiss, M. (2010b). Fusion of MODIS and VEGETATION observations for improved consistency and continuity of LAI product time series. In J. Sobrino (Ed.), <i>Third International Symposium on Recent Advances in Quantitative Remote Sensing</i> . Torrent (Spain)
7	'35 '36	Verger, A., Baret, F., & Weiss, M. (2011). A multisensor fusion approach to improve LAI time series <i>Remote Sensing of Environment</i> , 115, 2460-2470
7777	737 738 739	Verger, A., Baret, F., Weiss, M., Kandasamy, S., & Vermote, E., F. (2012). The CACAO method for smoothing, gap filling and characterizing anomalies in satellite time series. <i>IEEE Transactions on Geoscience and Remote Sensing</i>

Version définitive du manuscrit publiée dans / Final version of the manuscript published in : Remote Sensing of Environment (2013), Vol. 137, p. 299-309, DOI: 10.1016/j.rse.2012.12.027 Journal homepage: www.elsevier.com/locate/rse

Verger, A., Martínez, B., Camacho-de Coca, F., & García-Haro, F.J. (2009). Accuracy 740 assessment of fraction of vegetation cover and leaf area index estimates from pragmatic 741 methods in a cropland area. International Journal of Remote Sensing, 30, 2685 - 2704 742 Vermote, E., Justice, C., Csiszar, I., Eidenshink, J., Mynemi, R., Baret, F., Masuoka, E., & 743 Wolfe, R. (2010). A terrestrial surface climate data record for global change studies. In 744 J. Sobrino (Ed.), Third International Symposium on Recent Advances in Quantitative 745 746 Remote Sensing. Torrent (Spain) Weiss, M., Baret, F., Garrigues, S., Lacaze, R., & Bicheron, P. (2007). LAI, fAPAR and 747 fCover CYCLOPES global products derived from VEGETATION. part 2: Validation 748 and comparison with MODIS Collection 4 products. Remote Sensing of Environment, 749 750 110, 317-331 Yang, W., Huang, D., Tan, B., Stroeve, J.C., Shabanov, N., Knyazikhin, Y., Nemani, R., & 751 Mynemi, R. (2006a). Analysis of leaf area index and fraction of PAR absorbed by 752 vegetation products from the Terra MODIS sensor: 2000-2005. IEEE Transactions on 753 754 Geoscience and Remote Sensing, 44, 1829-1842 Yang, W., Shabanov, N.V., Huang, D., Wang, W., Dickinson, R.E., Nemani, R.R., Knyazikhin, 755 Y., & Myneni, R.B. (2006b). Analysis of leaf area index products from combination of 756 757 MODIS Terra and Aqua data. Remote Sensing of Environment, 104, 297-312 758 Yuan, H., Dai, Y., Xiao, Z., Ji, D., & Shangguan, W. (2011). Reprocessing the MODIS Leaf Area Index products for land surface and climate modelling. Remote Sensing of 759 Environment, 115, 1171 760 761

762