A multisensor fusion approach to improve LAI time series

Aleixandre Verger, Frédéric Baret, Marie Weiss

Version définitive du manuscrit publié dans / Final version of the manuscript published in : Remote Sensing of Environment, 2011, In Press, DOI: 10.1016/j.rse.2011.05.

A multisensor fusion approach to improve LAI time series

Aleixandre Verger [a,b,*], Frédéric Baret [b], Marie Weiss [b]

[a] Departament de Física de la Terra i Termodinàmica, Universitat de València. C/ Dr. Moliner, 50. 46100 Burjassot, València, Spain

[b] INRA, Université d'Avignon. UMR114, EMMAH, Domaine Saint-Paul, Site Agroparc, F-84914 Avignon, France

[*] Corresponding author. E-mail address: aleixandre.verger@uv.es (A. Verger)

Abstract

High-quality and gap-free satellite time series are required for reliable terrestrial monitoring. Moderate resolution sensors provide continuous observations at global scale for monitoring spatial and temporal variations of land surface characteristics. However, the full potential of remote sensing systems is often hampered by poor quality or missing data caused by clouds, aerosols, snow cover, algorithms and instrumentation problems. A multisensor fusion approach is here proposed to improve the spatio-temporal continuity, consistency and accuracy of current satellite products. It is based on the use of neural networks, gap filling and temporal smoothing techniques. It is applicable to any optical sensor and satellite product. In this study, the potential of this technique was demonstrated for leaf area index (LAI) product based on MODIS and VEGETATION reflectance data. The FUSION product showed an overall good agreement with the original MODIS LAI product but exhibited a reduction of 90% of the missing LAI values with an improved monitoring of vegetation dynamics, temporal smoothness, and better agreement with ground measurements.

Keywords: LAI time series, VEGETATION, MODIS, temporal smoothing, gap filling, data fusion

Version définitive du manuscrit publié dans / Final version of the manuscript published in : Remote Sensing of Environment, 2011, In Press, DOI: 10.1016/j.rse.2011.05.

Manuscrit d'auteur / Author manuscript
Manuscrit d'auteur / Author manuscript

1 Introduction

Leaf area index (LAI), defined as half the total developed area of green leaves per unit of horizontal ground area (Chen and Black, 1992), is a key vegetation variable used for a wide range of ecological, agricultural and meteorological applications. LAI has been recognized as an Essential Climate Variable for its key role in land-atmosphere interactions (GCOS, 2010). Several international programs and initiatives have established guidelines for monitoring climate and environment, stressing the need for realistic, high quality, temporally, and spatially stable, continuous and long time series of validated LAI global products (GCOS, 2010; GEO, 2010). Moderate resolution sensors provide frequent observations of the surface and allow monitoring LAI at the global scale. Several LAI products are available to the science community, each of them being routinely produced from a unique sensor: VEGETATION/SPOT (Baret et al., 2007; Deng et al., 2006), MODIS/TERRA-AQUA (Myneni et al., 2002) and SEVIRI/MSG (García-Haro et al., 2009). However, each of these LAI products is not spatially and temporally continuous and contains occasional missing data mainly due to persistent clouds, snow cover and optically thick aerosols (Garrigues et al., 2008). Further, extreme solar geometries and long periods of darkness in winter limit the use of optical sensors at very high latitudes (Beck et al., 2006). Problems related to the sensor, data storage or transmission may also limit the continuity of observations. Finally, instability of retrieval algorithms to residual atmospheric effects, cloud contamination in surface reflectance, or directional effects degrade the LAI product accuracy and reduce the probability to find a reliable solution: in the case of MODIS collection 5, about 10–15% of the pixels with mostly herbaceous cover and about 15–30% of the pixels with woody vegetation are retrieved with the back-up algorithm, triggered when the main algorithm fails, resulting in lower quality products (Samanta et al., 2009). Compositing procedures are used to ensure a continuous and time-consistent delivery of products by reducing the impact of missing data and of unexpected day-to-day variations in retrieved LAI. The CYCLOPES LAI algorithm (Baret et al., 2007) is applied to composited reflectances estimated by fitting a surface BRDF (Bidirectional Reflectance Distribution Function) model (Roujean et al., 1992) over 30 days of observations. Conversely, the MODIS LAI algorithm is first run for each single observation date and the resulting LAI values are composited over 8 days (Myneni et al., 2002). However, this approach which selects one observation over 8 potential ones induces some significant scattering which results in lower temporal consistency as shown by Garrigues et al. (2008). To mitigate this problem, Verger et al. (2008) demonstrated that using truly composited reflectances as derived from the MODIS 16 days product (MOD43B4 (Schaaf et al., 2002)) and neural network based algorithm provides more consistent LAI products. Another approach

could be a better compositing technique of the daily LAI products. However, the lack of clear-sky observations over periods longer than the compositing window will anyway result in missing products. Although increasing the length of the compositing window would increase the probability to get actual observations, it would induce artifacts due to the intrinsic assumption of the stability of the surface during the compositing period (Roy et al., 2006).

Several temporal and/or spatial techniques have been proposed to fill gaps resulting from missing data. Borak and Jasinski (2009) evaluated several interpolation techniques for MODIS LAI time series and concluded that temporal interpolation performs more efficiently over non-forest cover types, while spatial interpolation performs better in forest areas. Gao et al. (2008) proposed to fill gaps with an ecosystem-dependent dynamics model within a limited spatial window around the pixel. When no such data is available in the vicinity of the considered pixel, an ancillary seasonal curve is used. However, the methods based on the spatial dimension may fail to represent actual landscapes where LAI could vary widely within a short distance (Garrigues et al., 2006). To overcome these difficulties Fang et al. (2008) proposed a temporal spatial filtering algorithm based on a vegetation continuous field (VCF) ecosystem curve fitting method (Hansen et al., 2002). This algorithm uses the MODIS VCF product which contains the fractional cover of trees, bare, and herbaceous within the pixel. It allows imposing regional dependant phenological behavior onto each target pixel's temporal data, while maintaining pixel-level spatial and temporal integrity. However, an intrinsic limitation of such reconstruction methods is their inability to capture underlying atypical modes of seasonality (García-Haro et al., 2009) including natural and human induced disturbances (Jönsson and Eklundh, 2002). Further, all these methods are very sensitive to possible classification errors.

Concurrent use of data from several satellite systems will provide more potential observations within a given time period, increasing the availability of data in many regions with a diurnal variability of clouds. The fusion of multisensor data appears thus as an appealing strategy to fully exploit the current Earth observation missions and to fulfill user requirements in terms of spatio-temporal continuity, consistency and quality of products.

This paper aims at developing improved continuous LAI products derived from the fusion of available VEGETATION and MODIS observations. For this purpose, consistent LAI products are generated from MODIS and VEGETATION composited data following Verger et al. (2008) and are then smoothed and gap filled using a dedicated technique. The satellite products and the reference ground-based measurements are first presented. Then, the principles of the algorithm are described, validated based on simulated time series,

and applied over a subsample of situations globally representative of the range of conditions and surface types. Finally, the performances of the developed products are analyzed and compared with the original MODIS LAI product with emphasis on the continuity, consistency and agreement with ground measurements.

2 Data

To properly design, implement and validate the multisensor fusion algorithm, a data base of satellite reflectances and LAI is required over a globally representative ensemble of sites and over a time period long enough to capture their seasonality. The BELMANIP network, including 397 sites aiming to represent the variability of surface conditions over the Earth (WWW1; Baret et al., 2006), is used. The BELMANIP sites were generally chosen to be the most homogeneous (at 1 km resolution) as possible over a 3x3 km² area. This allows to minimize the point spread function effect and possible geolocation discrepancies between the two considered sensors, VEGETATION and MODIS. We considered the 2001-2003 period to well represent seasonal and interannual variability. The reflectance products are first presented. Then, the MODIS LAI product used to calibrate the algorithm for the fusion of VEGETATION and MODIS observations are depicted. Finally, the ground-based measurements used to validate the satellite LAI estimates are described.

2.1 The VEGETATION and MODIS reflectance composited products

CYCLOPES (derived from SPOT VEGETATION) and MOD43B4 (derived from Terra MODIS) reflectance products correspond to nadir viewing normalized reflectances in the red, near infrared and short wave infrared spectral bands (centered, respectively, at 645 nm, 835 nm, and 1665 nm for VEGETATION, and at 648 nm, 858 nm and 1640 nm for MODIS).

CYCLOPES reflectance products are available at WWW2. They correspond to 1/112° spatial sampling interval (about 1 km at the equator) and are projected in plate-carrée. The temporal sampling interval is 10 days with a 30 days compositing window. A Gaussian weighting function is applied to put more emphasis on observations close to the centre of the window. Top of atmosphere reflectance are first cloud screened and atmospheric correction is applied. The Roujean et al. (1992) BRDF model is then adjusted over the resulting top of canopy reflectance available within the compositing window to compute a reflectance value for nadir viewing and for the median solar zenith angle during the compositing period for each of the VEGETATION bands. Note that a minimum of two valid observations in the compositing window was required. An initial

version of the algorithm was successfully validated by Hagolle et al. (2005). More details are available in Baret et al. (2007).

MODIS reflectance products were downloaded from WWW3. They correspond to 1 km spatial sampling interval using a sinusoidal projection system. Nadir BRDF adjusted reflectance product (MOD43B4) collection 4 provides atmospherically corrected (Vermote et al., 1997), cloud-free, normalized reflectance values for each of the MODIS spectral bands at the mean solar zenith angle of each 16 day period (Schaaf et al., 2002). MOD43B4 product is computed using the Ross-Li BRDF model (Lucht et al., 2000). When the number of available observations is too small for a full BRDF inversion, a global database of archetypal BRDF model parameters is used to predict the shape of the BRDF (Schaaf et al., 2002). MOD43B4 products were validated by Jin et al. (2003). In addition to the nadir normalized reflectance values, the MOD43B4 product provides extensive quality information. Only data labeled as 'good quality', i.e. full inversion with at least seven cloud-free observations, were included in our dataset.

The consistency between VEGETATION and MODIS reflectance composited products was already demonstrated by Verger et al. (2008).

2.2 The MODIS LAI product

MODIS/TERRA LAI collection 5 product (MOD15A2) (available at WWW3) was used. The spatial sampling characteristics are similar to those of the MOD43B4 reflectance product. The main MODIS LAI retrieval algorithm relies on a radiative transfer model which ingests red and near infrared bidirectional reflectance factor values, their associated uncertainties, the view-illumination geometry, and biome type (within eight types) (Myneni et al., 2002; Shabanov et al., 2005). The MODIS algorithm accounts for clumping through three-dimensional radiative transfer formulation and assumptions on canopy architecture specific per biome class (Knyazikhin et al., 1998). If this algorithm fails, a back-up procedure is triggered to estimate LAI from biome specific NDVI based relationships. MODIS LAI product is generated by selecting the daily LAI value which corresponds to the maximum FPAR (Fraction of Photosynthetically Absorbed Radiation) value in an 8-day compositing period. In this study, only the MODIS LAI values generated by the main algorithm were considered. Back-up retrievals were not used because they have generally lower quality mostly due to residual clouds and poor atmospheric correction (Yang et al., 2006). Further only the pixels assigned to the eight biome land cover types were considered which excludes barren/sparse vegetation (rock, tundra, and desert), permanent wetlands/inundated marshlands or perennial snow/ice.

Although MODIS LAI product has been extensively validated (e.g. Ganguly et al., 2008; De Kauwe et al., 2011), high level of noise was inducing shaky temporal profiles and unrealistic seasonality (Kobayashi et al., 2010) which limits the interest of directly using this product in the proposed fusion algorithm.

2.3 Ground-based LAI measurements

Garrigues et al. (2008) proposed a set of validation sites where ground values of LAI were derived from indirect methods based on light transmittance measurements. For some of the sites, measurements were repeated at several dates. However, the ground measurements were not always achieved within the 2001-2003 period considered in our study. To increase the number of available points, the ground measurements completed outside the 2001-2003 period were included in the validation exercise (section 4) for the forest sites appearing as steady over few years based on remote sensing observations. Local ground measurements were scaled up by using empirical transfer functions and ancillary high-resolution images (Morisette et al., 2006). The accuracy of ground-based reference maps depend on errors in field measurements but also on uncertainties of fine-resolution satellite data, sampling and spatial scaling errors (Weiss et al., 2004). Note that ground measurements could be derived from several devices and interpretation techniques, and may provide estimates of effective LAI values (assuming the canopy as a turbid medium) or true LAI values when leaf clumping is accounted for (Weiss et al., 2004). To be consistent with the definition of satellite products, special attention was paid to the nature of LAI measurements and only LAI values accounting for clumping were selected yto provide a more reliable assessment of products accuracy.

3 Methods

A fusion algorithm was developed to estimate LAI from the synergistic use of VEGETATION and MODIS observations while keeping the spatial and temporal sampling as close as possible to the original MODIS LAI product. The proposed approach is based on the use of neural networks and temporal compositing techniques. The strategy included two main steps. First, for each of the 8 MODIS biomes, two neural networks were trained to estimate MODIS LAI products, one from MOD43B4 and the other from CYCLOPES reflectance products. Then, a specific Temporal Smoothing and Gap Filling algorithm (TSGF) was applied to the neural network LAI estimates resulting in the 'FUSION' product. The same TSGF algorithm was also applied to the main MODIS LAI values ('SmoothedMOD') to better evaluate the advantage of fusing several products. Table 1 lists the products that are used in section 4 for performances evaluation showing

differences in the temporal sampling and resolution, but sharing approximately the same spatial support. The flow chart of the proposed algorithm is given in Fig. 1.

[Table 1]

3.1 Neural networks LAI estimation

The versatility and performance of neural networks to learn a particular LAI product from several input reflectance products was demonstrated by Verger et al. (2008). MODIS LAI product estimated from CYCLOPES reflectances ('CYCrhoMODlai') and MOD43B4 reflectances ('MODrhoMODlai') combinations were here investigated.

The neural networks were trained per land cover class (Verger et al., 2008) using MOD12Q1 land cover map (Friedl et al., 2002) both for CYCLOPES and MOD43B4 reflectances. Seven main biomes were considered: shrubs, savanna, grasses and cereal crops, broadleaf crops, needleleaf forest (deciduous and evergreen), deciduous broadleaf forest and evergreen broadleaf forest (Table 2).

The LAI estimation approach does not need a very accurate absolute calibration of reflectance products from the two considered sensors: the main requirement here is a very high degree of spatial and temporal consistency of reflectances in each band (Verger et al., 2008). The training data set should represent the same spatial and temporal sampling support. CYCLOPES reflectance products were first re-projected into MODIS sinusoidal projection system using a bi-cubic resampling method (Reichenbach and Geng, 2003). For all inputs and outputs, the median value over 3×3 pixels area was computed when at least 5 high quality level pixels were available. This allows minimizing geometrical uncertainties while removing most outliers inside the 3×3 pixels area. The 'dominant biome class' is considered. However, sites with less than 5 pixels belonging to the same biome class in the 3×3 pixels area during the 3 year period were rejected from the training dataset. Remaining outliers in reflectance values due to cloud contamination, residual atmospheric and directional effects or snow masking errors, were further eliminated by rejecting the 10% of samples showing the largest discrepancies between CYCLOPES and MOD43B4 reflectances over the three spectral bands. The temporal sampling frequency was set to the 16 days of MOD43B4 reflectance product corresponding to the lowest one among inputs and output variables. The 8-day original MOD15A2 LAI product was thus composited to get a 16-day temporal sampling: the central MOD15A2 value was averaged with the ones from the previous and the next compositing windows, assigning a 0.5 weight to these bracketing observations. However, when one or two of the border observations were missing, the product

was still computed from the remaining observations. Conversely, when the central observation was missing, no product was computed. The CYCLOPES reflectances were finally linearly interpolated at the central date of MOD43B4 product if the two closest CYCLOPES observations were within ±15 days from this date. Further details about the generation of the training dataset can be found in Verger et al. (2008).

Back-propagation networks made of one layer of five tangent-sigmoid and one layer with one linear transfer function were considered (Verger et al., 2011). They result in 26 synaptic and 6 bias coefficients to be adjusted. The inputs (reflectances in the red, near infrared and short wave infrared and the sun zenith angle) and LAI output were scaled by their minimum and maximum values. The training dataset was split in two subsets by randomly selecting cases. Two thirds of the cases were used to train the networks while the remaining one third of cases was used to evaluate the performances and avoid over-fitting.

Results (Table 2) show that performances of LAI estimates are very close between CYCLOPES and MOD43B4 reflectances, with a slightly smaller root mean square error (RMSE) values for CYCrhoMODlai. RMSE values are low for non-forest biomes (RMSE<0.2) and reasonable for savanna (<0.5) and forest biomes (<0.9). Note that the RMSE values increase with the LAI values across biomes, with non-forest and evergreen broadleaf forest experiencing respectively the lowest and the largest ones. These results agree well with those of Verger et al. (2008).

[Table 2]

The neural networks were then applied to the reflectance products at the pixel level (sinusoidal MODIS 1km sampling grid) and at the original temporal sampling interval (10 days for CYCLOPES and 16 days for MOD43B4). Since the neural networks were trained with MODIS LAI, the definition of CYCrhoMODlai, MODrhoMODlai and FUSION LAI products agrees with the MODIS LAI definition.

3.2 The temporal smoothing and gap filling (TSGF) method

The 1 km spatial sampling interval MODrhoMODlai and CYCrhoCYClai were fused to provide LAI values every 8 days using temporal smoothing and gap filling processes (Fig. 1).

1. Temporal smoothing

   After evaluating several widely used temporal filters, a simple but robust method based on the adaptive Savitzky-Golay (SG) filter (Savitzky & Golay, 1964; Chen et al., 2004) was selected to smooth the LAI temporal profiles. The SG filtering captures subtle and rapid changes while being little sensitive to outliers (Fang et al., 2008). Chen et al. (2004) showed the effectiveness of an

Version définitive du manuscrit publié dans / Final version of the manuscript published in : Remote Sensing of Environment, 2011, In Press, DOI: 10.1016/j.rse.2011.05.

adaptive SG filter in comparison to the Best Index Slope Extraction method (Viovy et al., 1992) and the fast Fourier transform technique for reconstructing SPOT VEGETATION high-quality NDVI time-series. SG also worked well for MODIS LAI series (Fang et al., 2008).

The proposed approach used local polynomial functions similarly to the SG-filtering. However, the fixed and symmetric compositing window of the standard SG was replaced by an adaptative process with variable length and allowing asymmetry in time:

1.1. The first step consisted in defining the compositing window and identifying the high-quality data used to fit the polynomial over the available LAI values. The length of the compositing period is one of the main drivers of the smoothing performances: a short compositing window captures subtle and rapid changes in the time series but is more sensitive to noise and gaps (Chen et al., 2004); conversely, a long compositing window produces smoother temporal profiles at the expense of flattening sharp peaks. After some trial and error tests, the period was defined by selecting the 3 closest observations before and 3 closest after the considered date within a maximum 128-day compositing window centered on the considered date (±64-day). If no such 6 LAI values were available within the 128 days compositing period, the smoothing was not applied, resulting in a missing data in the smoothed product.

1.2. Weights W_i at date i were set to the frequency of the corresponding products, i.e. $W_i = 10$ for CYCrhoMODlai and $W_i = 16$ for MODrhoMODlai. This allows assigning the same importance to the LAI estimates derived from CYCLOPES (10-day frequency) and MOD43B4 reflectances (16-day). These weights were further normalized to ensure an even contribution of each side of the compositing period:

$$W_{i_-}^{'} = \frac{W_{i_-}}{\sum_{i_-=1}^{3} W_{i_-}} ; \; W_{i_+}^{'} = \frac{W_{i_+}}{\sum_{i_+=1}^{3} W_{i_+}}$$

where the signs refer to the observations *i* before (-) or after (+) the considered compositing date i_0. When an actual observation is available at date i_0, its weight, $W_0^{'}$ is set to $W_0^{'} = 2\frac{\sum_{i_0=1}^{n_{i_0}} W_{i_0}}{\sum_{i_-=1}^{3} W_{i_-} + \sum_{i_+=1}^{3} W_{i_+}}$ where n_{i_0} is the number of actual observations ($0 \le n_{i_0} \le 2$ when 2 reflectance products are considered as in this study) at the

Version définitive du manuscrit publié dans / Final version of the manuscript published in : Remote Sensing of Environment, 2011, In Press, DOI: 10.1016/j.rse.2011.05.

compositing date i_0. A second order polynomial function was then fitted using a weighed least square technique within the compositing window.

1.3. The local polynomial fitting should be used cautiously to reproduce adequately the seasonality of the original data, especially during rapid change of growth or senescence rate: too large moving window will prevent describing realistically these features. To capture the corresponding sudden rise in data values and limit the flattening of sharp peaks sometimes observed in the fitted data, only a relatively small window can be used. To minimize this bias artifact, a correction was applied over a smaller window (±32-day) centered on the peaks by applying a linear regression between the fitted values and the original data if a minimum of 4 observations exist within the ±32-day window.

2. Gap filling

A simple linear interpolation in a local moving window of 128- day length was applied to fill gaps in the smoothed data series. The gain of reconstructed data achieved with the first iteration was further exploited to improve the robustness and the continuity in the time series through a second iteration of the linear interpolation.

[Fig. 1]

4 Validation

Validation of multi-temporal methods applied to moderate resolution satellite data is often not straightforward, since independent ground measurements for a broad range of conditions (seasonal period of vegetation or atmospheric conditions) are required. Most of ground-based reference maps are single-date maps derived under clear-sky conditions in the maximum development state of vegetation (Baret et al., 2006). The lack of ground measurements representative of the MODIS 1 km products for a long enough time period hampers the quantitative assessment of MODIS time series. For testing temporal analysis methods in a controlled environment, simulation of time series appears complementary to the application to satellite data and comparison with ground data (Verbesselt et al., 2010). The temporal smoothing and gap filling capacity of TSGF was therefore validated by (1) simulating 8-day LAI time series, and (2) applying the method to actual LAI data. simulated range of8-day LAI time series was simulated with different noise levels and percentage of gaps in order to robustly test TSGF (section 4.1). Then the method was applied to both 8-day MODIS satellite LAI time series and 10-day and 16-day neural network LAI estimates. The developed

SmoothedMOD and FUSION products were validated by analyzing their continuity, consistency and accuracy. The continuity of LAI products was evaluated by characterizing the spatio-temporal and biome distribution of gaps (section 4.2) and the distribution of length of periods without products values (Length of Gaps, LoG) (section 4.3). Temporal consistency was evaluated by the smoothness level of the time series (section 4.4). The consistency of temporal profiles was further analyzed over a set of sites based on expert knowledge and auxiliary information (section 4.5). Finally, the several LAI products investigated were first intercompared (section 4.6) and their accuracy was evaluated over the set of available ground measurements (section 4.7).

4.1 Assessment of TSGF in simulated LAI time series

Simulated 8-day LAI time series are generated for a 3-year period by using an asymmetric Gaussian function for each season (Verbesselt et al., 2010) (dotted line in Fig. 2). Second, different levels of white noise and randomly distributed gaps were created in the simulated time series (crosses in Fig. 2). The noise component was generated using a random number generator that follows a normal distribution $\aleph(0, \sigma)$ (mean value equal to 0 and variance, σ^2), i.e. $LAI_{noisy} = LAI + \aleph(0,\sigma)$. The values for the absolute LAI uncertainty used in this study were varying between σ= 0 and σ= 0, 0.5 by 0.05 steps. The percentages of simulated gaps were ranging between 0 and 0.7 by 0.1 steps. Third, the TSGF method was applied to the degraded data (solid line in Fig. 2). Fourth, the reconstructed time series were compared with the original reference data by computing the RMSE. In order to obtain representative values, the RMSE was derived as the mean value of 500 iterations for each level of noise and gap proportion.

[Fig. 2]

Fig. 2 illustrates that the TSGF method reproduces adequately the seasonality and the time of changes. A good correspondence between the reconstructed time series and the original test LAI data is generally observed. However, the method shows some limitations to reconstruct sharp peaks when the proportion of gaps is high since a minimum of 4 LAI data values within a ±32-day window centered on the peak is required for applying the flattening correction (step 1.3 of TSGF method, section 3.2). Some artifacts are also observed when the signal to noise ratio is low as observed around the minima of the time series (Fig. 2). Results show as expected that the uncertainty in the reconstruction of temporal series increases with the level of noise and the fraction of missing data (Fig. 3). For continuous time series (i.e. fraction of missing data equal to 0), the mean RMSE increases linearly with the noise level with an offset close to zero. For

discontinuous time series, the RMSE increases slowly with the noise level when the noise level is small compared to the percentage of missing data.

[Fig. 3]

4.2 Distribution of gaps

The spatio-temporal distribution of missing data was computed per 10° latitudinal bands and for each month over the BELMANIP sites. Results (Fig. 4) show that the gaps are mostly observed around the equator due to persistence of clouds and at high latitudes in wintertime mainly due to snow and cloud cover as well as large sun zenith angles. A high frequency of gaps is also observed at mid latitudes during summertime which may be related to cyclonic activity (Fig. 4). The original MODIS product presents the highest rate of invalid data (34%). This may be partially explained due toby the short composition period of the MOD15A2 LAI. Note that the MOD15A2 back-up LAI product, which represents 22% of cases, was considered as invalid data. Barren/sparse vegetation pixels were not included in this study since MOD15A2 was not available for these cases.

The MODrhoMODlai product shows less than half the fraction of invalid data (16%) as that of MOD15A2(34%) because of differences in the length of the compositing period (16 and 8 days, respectively). Similar spatio-temporal patterns were also observed between these two LAI products. However, MODrhoMODlai presents a significant lower frequency of gaps for the northern latitudes. This may result from the greater tolerance to the quality of input reflectance of the neural network algorithm as compared to the main MOD15A2 algorithm.

Even though the compositing period of the reflectance data used as input into CYCrhoMODlai (30 days) is twice as long as that of MODrhoMODlai (16 days) and the minimum number of valid observations required in the compositing period is lower for the first (only two observations were required in the CYCLOPES algorithm while MOD43B4 reflectances were generated only if a minimum of seven observations were available, section 2.1), the gap frequency is slightly higher for CYCrhoMODlai (20%) than for MODrhoMODlai (16%). Both products show clear and similar spatio-temporal distributions of gaps, with large fraction of gaps located at the equator and the higher latitudes in winter (Fig. 4). However, the main differences are observed at these higher latitudes where MODrhoMODlai shows lower frequency of gaps. This may be explained by differences in snow and cloud screening performances as well as to the realism of the BRDF models used in the compositing algorithms (Lucht & Lewis, 2000).

The SmoothedMOD product shows a significantly reduced rate of gaps (13%) as compared to MOD15A2 (34%), demonstrating the efficiency of the proposed TSGF method. The remaining gaps in SmoothedMOD (Fig. 4) are still located over the equatorial and high latitudes where the original MODIS product typically presents long periods (at least 128-day) with almost no LAI values.

The FUSION product shows a very low fraction of gaps (4%). Gaps in the FUSION product are mostly located at the equator while almost no gaps are observed for the higher latitudes demonstrating the effectiveness of the synergy between CYCrhoMODlai and MODrhoMODlai products.

[Fig. 4]

The distribution per biomes (Table 3) shows that most of the gaps in the MOD15A2 product correspond to forest areas mainly located at the equator and the higher latitudes, and also to shrubs largely located also at the higher latitudes (see MODIS land cover map in Myneni et al., 2002). Further, forests may correspond to high LAI values with possible saturation conditions in the MOD15A2 algorithm. The derived SmoothedMOD product shows a significant reduction of gaps for all biomes, particularly for the deciduous broadleaf forest and all the non-forest biomes except shrubs. The CYCrhoMODlai shows around 10% frequency of gaps for non-forest biomes while forests show particularly high fraction of gaps that culminates with the evergreen broadleaf forests (41%). Conversely, around 15% of data are missing for MODrhoMODlai product except for evergreen broadleaf forest that reaches 34% gaps. The FUSION product presents less than 3% gaps except for evergreen broadleaf forest that however provides still the lower frequency of gaps with 17% although similar to the SmoothedMOD product (18%). This clearly indicates the difficulties to monitor this particular biome located near the equator with VEGETATION and MODIS satellites. However, the comparison between the FUSION and the SmoothedMOD products for the other biomes and particularly for shrubs and needleleaf forest, demonstrates the better performances of the multisensor fusion approach as compared to the monosensor gap filling approach.

[Table 3]

4.3 Length of gaps (LoG)

The time interval between two consecutive valid observations of LAI products on each side of a gap, i.e. the length of gaps (LoG), is a very pertinent metrics to describe the continuity of products required for monitoring land surfaces. The histogram of LoG (Fig. 5) shows that the FUSION and the SmoothedMOD products have very low frequencies of gaps with lengths about evenly distributed. The three other LAI products show very

similar patterns after a LoG value around 50 days. However, for short LoG values, MOD15A2 and MODrhoMODlai products showed similar patterns with the highest frequency of LoG. Conversely, CYCrhoMODlai presents a higher frequency of long gaps as compared to MODrhoMODlai.

[Fig. 5]

4.4 *Smoothness*

The smoothness level of LAI temporal series was evaluated as proposed by Weiss et al. (2007) using the absolute difference, δ, between $LAI(t)$ product value at date t and the mean value between the two bracketing dates: $\delta=|1/2(LAI(t+\Delta t)+LAI(t-\Delta t)) - LAI(t)|$, where Δt is the temporal sampling interval. Difference δ is computed only if the two bracketing *LAI* values at $(t-\Delta t)$ and $(t+\Delta t)$ exist. The smoother the temporal evolution, the smaller the δ difference should be.

The original MODIS product is the most 'shaky' in relation with its short compositing period (Fig. 6). MODrhoMODlai shows an improvement in the smoothness due both to the increase of the temporal compositing window, the nature of the compositing algorithm (BRDF model fit) and the use of neural network based biophysical algorithm that provides more stable results than the MOD15A2 algorithm as already noticed by Verger et al. (2008). The CYCrhoMODlai shows intermediate smoothness level because of the increase of the compositing period as compared to MODrhoMODlai. Finally, SmoothedMOD and FUSION products exhibit the smoothest temporal profiles because of the TSGF process and its associated long ($\leq$ 128-day) compositing period.

[Fig. 6]

4.5 *Temporal profiles*

The LAI temporal profiles of the central pixel of five BELMANIP sites (Table 4) selected to represent typical features were investigated (Fig. 7). Visual inspection of the temporal profiles confirms that the original MODIS product is extremely shaky, while the FUSION product is extremely smooth, MODrhoMODlai, CYCrhoMODlai and SmoothedMOD showing intermediate situations. The seasonality is generally well depicted by the different products for Mongu site with a good agreement with the few available ground measurements along its phenological cycle. However some discrepancies are observed over the Dahra site, that is partly explained because of the small spatial support of the ground measurements that may not

represent accurately coarse resolution pixels. This may also reveal a limitation of the temporal smoothing algorithm to capture faithfully the rapid changes when the percentage of missing data is very high as commented in section 4.1. And a limitation of the fusion approach to deal with differences in the original products when the LAI changes abruptly as observed in the MODrhoMODlai and CYCrhoMODlai products for the Dahra site. The situation is more complex for the forest sites: it is very difficult to describe any seasonality at the pixel level with the original MODIS product because of its extremely shaky character while the SmoothedMOD product exhibits significant artifacts due to the small number of available original MODIS data and the large associated uncertainties. Conversely, the FUSION product benefits from the complementarities in temporal consistency of MODrhoMODlai and CYCrhoMODlai products and shows the most reasonable temporal profiles, with a marked seasonal pattern for Hyytiälä and Soignes which is determined by the presence of understory and deciduous species and with very limited seasonality as expected for Counami (Table 4).

[Table 4]

[Fig. 7]

4.6 Comparison between LAI products

The intercomparison of LAI products was achieved at the pixel scale and the closest date. For the sake of simplicity, the RMSE values were computed for the 'Forest' and the 'Non-Forest' biome classes since similarities are observed within these broad classes. RMSE values are clearly lower for the non-forest classes as compared to the forest classes (Table 5), in relation with the LAI levels experienced by these classes in agreement with earlier observations. As expected the SmoothedMOD product agrees the best with the original MODIS LAI product. Neural networks estimates of MODIS LAI from the MODIS reflectances (MODrhoMODlai) are very consistent with estimates from the CYCLOPES reflectances (CYCrhoMODlai). This supports our hypothesis that the spatial sampling interval of VEGETATION and MODIS sensors is comparable and that training over the same biophysical product with input reflectance derived from different sensors provides very consistent output products in agreement with Verger et al. (2008). Larger differences are found when comparing MODrhoMODlai with MODIS which justifies a posteriori why it was preferred to include MODrhoMODlai instead of either the original MODIS product or the SmoothedMOD product in the fusion process. The RMSE associated to the comparison between MODIS and the FUSION product provides

an approximate quantification of the MODIS temporal noise. The FUSION product shows high consistency with both CYCrhoMODlai and MODrhoMODlai as expected.

[Table 5]

4.7 Direct validation

For the direct validation of the MODIS-based LAI products (closer to true LAI than to effective LAI definition),only ground LAI measurements proposed by Garrigues et al. (2008) (c.f. section 2.3.), which take into account the clumping effect were selected.. Further, to provide a consistent comparison of product performances, validation sites were selected if the MOD15A2, SmoothedMOD and FUSION LAI products were concurrently available within less than 8 days to the date of ground measurements acquisition. To minimize registration errors, the mean LAI product values over 3×3 pixels areas for the date the closest to the ground measurement campaign was compared to the ground LAI over the typical 3x3 km² size of ground validation sites. Our study, in agreement with the CEOS recommendations (Morisette et al., 2006), benefits from the high number (N=71) and representativeness of ground measurements which cover most of the physical range of LAI variation. However, there are not enough measurements for broadleaf forests where the largest differences between LAI products are expected (Table 2). Note that ground-based measurements are associated to errors and biases that are difficult to estimate (e.g. uncertainties in the estimation of clumping and woody area ratio, lack of understory quantification or spatial scaling errors).

The overall performances of LAI products were quantified by the RMSE decomposed into accuracy (B) and precision (S) components. The accuracy (B) is measured as the mean value of the differences between products and ground measurements while the precision (S) is computed as the standard deviation of estimates around the best linear fit also reflected by the correlation coefficient (R^2).

Results (Fig. 8) show that the FUSION product agrees the best with LAI ground measurements. The improvement in the overall performances (RMSE) of the FUSION product is due to the observed increment of the precision (scoring the highest R^2 and the lowest S values). In the fusion, the precision is increased through a double step process. First, the precision of intermediate neural network LAI estimates (MODrhoMODlai, CYCrhoMODlai) is improved as compared to that of MOD15 LAI product as proven in Verger et al. (2008) because neural networks smooth out noise and they are based on 16-day and 30-day composited reflectances as compared to 8-day maximum compositing of MOD15. Second, the fusion of MODrhoMODlai and CYCrhoMODlai results in more data and consequently in less noise. However, the

FUSION product presents a slight negative bias (B). These systematic differences are also observed in the comparison between the MOD15 product and the ground measurements. Since the fusion is based on the MOD15 LAI, the FUSION product cannot correct from possible biases in the original MOD15 product. Note also that most noisy points should be below the 1:1 line, since noise caused by clouds and poor atmospheric conditions is negatively biased. No significant difference is observed between the original and the smoothed MODIS LAI, with a slight improvement of performances through the smoothing process which reduces some artifacts (SmoothedMOD shows higher correlation (R^2) and lower scattering (S) than MOD15).

[Fig. 8]

5 Conclusions

This study introduced an innovative fusion approach based on the use of neural networks and temporal filtering (TSGF) techniques. The algorithm fills gaps and smoothes out consistent products derived from two sensors. It is a generic approach applicable to any combination of satellite products. It was here applied to improve MODIS LAI product using reflectance data from VEGETATION and MODIS sensors.

The performances of the approach for restoring the spatial and temporal distribution of MODIS LAI products is demonstrated based on the evaluation of the fraction of valid data, length of periods with missing data and occurrence of missing data in the original and proposed LAI products. The original MODIS product presents the highest rate of invalid data (34%) due to the shortest composition period but also to the failure of the main retrieval algorithm to provide a solution in sub-optimal conditions with high uncertainties in input data. The use of composited reflectances provides a more stable and continuous input dataset which allows reducing by more than a half the fraction of invalid data from the neural network estimates (MODrhoMODlai (16% of invalid data) and CYCrhoMODlai (20%)). The TSGF method appears to be very efficient to improve the continuity of LAI products (13% of invalid data in the SmoothedMOD product) but the instabilities of MODIS LAI are transferred in part to the SmoothedMOD product in periods and locations with frequent missing MODIS LAI product. The proposed fusion approach smoothes out efficiently most instabilities in the original MODIS product and allows filling most of the gaps (only 4% of gaps are remaining in the FUSION product) through the use of composited reflectances, neural network training and temporal smoothing and gap filling techniques. Further, the FUSION product agrees the best with ground measurements. However, residual gaps persist in evergreen broadleaf forests in the equatorial latitudes revealing the difficulties to monitor these particular areas due to persistent cloudiness. The investigated products show limitations to

describe rapid LAI changes as observed over vegetation surfaces with a short growing season. A refinement of the temporal filtering algorithm and the use of daily observations would probably mitigate this problem. Geostationary satellite sensors (e.g. SEVIRI/Meteosat) which deliver continuously-updated information could also contribute to improve the continuity and reliability of LAI products. Further studies should therefore concentrate on designing strategies that take advantage of the complementary characteristics of geostationary satellite products which provides temporal continuity and stability but coarser spatial resolution and polar orbit based observations with a finer spatial resolution but more affected by the multi-temporal noise and longer periods of missing data. Note that the proposed approach is entirely driven by remotely sensed observations and uses only information from the time series to be filled. For operational purposes this approach could be complemented by the use of a climatology resulting from the compilation of data over an extended temporal period or based on dedicated functioning models with the capacity to simulate the dynamics of LAI.

The proposed fusion approach assumes some degree of spatial and temporal consistency between the products to be fused. Although the temporal aspect is probably minor and could be solved using the daily observations from different sensors rather than the composited reflectance products, the spatial aspect might be more limiting. The results reported here demonstrate that the fusion algorithm accommodates the differences in the spatial support of VEGETATION and MODIS over the BELMANIP sites. However, co-registration errors, projection systems and point spread function of both sensors may compromise the applicability of the proposed approach over heterogeneous areas where more attention to the scaling effects should be paid.

Our results show the high potential of the proposed multisensor fusion approach as compared to the standard temporal filtering methods for improving the performances of satellite products in terms of continuity, consistency and agreement with ground measurements. This innovative sensor-independent approach may contribute to generate continuous long-term Earth System Data Records from remote sensing data collected with several sensors over the past three decades and to extend the time series with sensors to be launched in the near future.

Acknowledgments

The CYCLOPES products are generated under EU/FP5 CYCLOPES project, using algorithms developed by CNES, CNRM, INRA, and Noveltis. They are produced and provided by the POSTEL Service Centre at MEDIAS-France. MODIS products are generated under NASA support, using algorithms developed by

Boston University. They are provided by LPDAAC. Ground measurements result from VALERI/CNES, BigFoot/NASA, SAFARI2000/NASA and CCRS initiatives. Aleixandre Verger is funded by the VALi+d postdoctoral program (FUSAT, GV-20100270). The authors would like to thank the three anonymous reviewers for their professional and insightful comments. Many thanks also to Jan Verbesselt for his advice on the validation based on simulated time series. Finally we thank also the two anonymous reveiwers for their constructive comments.

References

Baret, F., Morissette, J., Fernandes, R., Champeaux, J., Myneni, R., Chen, J., Plummer, S., Weiss, M., Bacour, C., Garrigues, S., & Nickeson, J. E. (2006). Evaluation of the representativeness of networks of sites for the global validation and intercomparison of land biophysical products: Proposition of the CEOS-BELMANIP. *IEEE special issue: Validation and accuracy assessment of global products*, *44(7)*, 1794–1803.

Baret, F., Hagolle, O., Geiger, B., Bicheron, P., Miras, B., Huc, M., Berthelot, B., Niño, F., 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. *Remote Sensing of Environment*, *110*, 275−286.

Beck, P. S. A., Atzberger, C., Høgda, K. A., Johansen, B. & Skidmore, A. K. (2006). Improved monitoring of vegetation dynamics at very high latitudes: A new method using MODIS NDVI. *Remote Sensing of Environment*, *100*, 321–334.

Borak, J.S. & Jasinski, M.F. (2009). Effective interpolation of incomplete satellite-derived leaf-area index time series for the continental United States. *Agricultural and Forest Meteorology*, *149*, 320-332.

Chen, J. M., & Black, T. A. (1992). Defining leaf area index for non-flat leaves. *Plant Cell Environment*, 15, 421−429.

Chen, J., Jönsson, P., Tamura, M., Gu, Z.,Matsushita, B., & Eklundh, L. (2004). A simple method for reconstructing a high quality NDVI time-series data set based on the Savitzky–Golay filter. *Remote Sensing of Environment*, *91*, 332−344.

De Kauwe, M.G., Disney, M.I., Quaife, T., Lewis, P., & Williams, M. (2011). An assessment of the MODIS collection 5 leaf area index product for a region of mixed coniferous forest. *Remote Sensing of Environment*, *115*, 767−780.

Deng, F., Chen, J. M., Plummer, S., Chen, M., & Pisek, J. (2006). Algorithm for global leaf area index retrieval using satellite imagery. *IEEE Transactions Geoscience and Remote Sensing*, *44*, 2025–2047.

Fang, H., Liang, S., Townshend, J. R. & Dickinson, R. E. (2008). Spatially and temporally continuous LAI data sets based on an integrated filtering method: Examples from North America. *Remote Sensing of Environment*, *112*, 75–93.

Fensholt, R., Sandholt, I., & Rasmussen, M. S. (2004). Evaluation of MODIS LAI, fAPAR and the relation between fAPAR and NDVI in a semi-arid environment using in situ measurements. *Remote Sensing of Environment*, *91*, 490−507.

Friedl, M. A., McIver, D. K., Hodges, J. C. F., Zhang, X. Y., Muchoney, D., Strahler, A. H., Woodcock, C. E., Gopal, S., Schneider, A., Cooper, A., Baccini, A., Gao, F., & Schaaf, C. (2002). Global land cover mapping from MODIS: Algorithms and early results. *Remote Sensing of Environment*, *183*, 287−302.

Ganguly, S., Samanta, A., Schull, M. A., Shabanov, N. V., Milesi, C., Nemani, R. R., Knyazikhin, Y., & Myneni, R. B. (2008), Generating vegetation leaf area index Earth system data record from multiple sensors. Part 2: Implementation, analysis and validation, *Remote Sensing of Environment*, *112*, 4318-4332.

Gao, F, Morisette, J., Wolfe, R., Ederer, G., Pedelty, J., Masuoka, E., Myneni, R., Tan, B., & Nightingale, J. (2008). An Algorithm to Produce Temporally and Spatially continuous MODIS-LAI Time Series. *IEEE Geoscience and Remote Sensing Letters*, *5(1)*, 60−64.

García-Haro, F. J., Camacho, F., Verger, A., & Meliá, J. (2009). Current status and potential applications of the LSA SAF suite of vegetation products. *29th EARSeL Symposium, Chania, Greece*. Available at http://www.earsel.org/symposia/2009-symposium-Chania/09EARSEL_garciaharoetal_LSASAF.pdf.

Garrigues, S., Allard, D., Baret, F., & Weiss, M. (2006). Quantifying spatial heterogeneity at the landscape scale using variogram models. *Remote Sensing of Environment*, *103*, 81-96.

Garrigues, S., Lacaze, R., Baret, F., Morisette, J. T, Weiss, M., Nickeson, J. E., Fernandes, R., Plummer, S., Shabanov, N.V., Myneni, R., & Yang, W. (2008). Validation and intercomparison of global leaf area

index products derived from remote sensing data. *Journal of Geophysical Research*, *113*, G02028, doi:10.1029/2007JG000635.

GCOS. (2010). Implementation Plan for the Global Observing System for Climate in Support of the UNFCCC, November 2009, 153 pp. (Document online at http://www.wmo.int/pages/prog/gcos/documents/GCOSIP-10_DRAFTv1.0_131109.pdf).

GEO. (2010). Group on Earth Observations. Report on Progress, November 2010, 147 pp. (Document online at http://www.earthobservations.org/documents/ministerial/beijing/MS2_The GEO Report on Progress.pdf)

Hagolle, O., Lobo, A., Maisongrande, P., Cabot, F., Duchemin, B., & De Pereyra, A. (2005). Quality assessment and improvement of temporally composited products of remotely sensed imagery by combination of VEGETATION 1 and 2 images. *Remote Sensing of Environment, 94*, 172-186.

Hansen, M.C., DeFries, R.S., Townshend, J.R.G., Sohlberg, R., Dimiceli, C., & Carroll, M. (2002). Towards an operational MODIS continuous field of percent tree cover algorithm: examples using AVHRR and MODIS data. *Remote Sensing of Environment, 83*, 303-319.

Jin, Y., Schaaf, C. B., Woodcock, C. E., Gao, F., Li, X., Strahler, A. H., Lucht, W., & Liang, S. (2003). Consistency of MODIS surface BRDF/Albedo retrievals: 2. Validation. *Journal of Geophysical Research*, *108(D5)*, 4159. doi: 10.1029/2002JD002804

Jönsson, P., & Eklundh, L. (2002). Seasonality extraction by function fitting to time–series of satellite sensor data. *IEEE transactions on Geoscience and Remote Sensing*, *40(8)*, 1824−1832.

Knyazikhin, Y., Martonchik, J. V., Myneni, R. B., Diner, D. J., & Running, S. W. (1998). Synergetic algorithm for estimating vegetation canopy leaf area index and fraction of absorbed photosynthetically active radiation from MODIS and MISR data. *Journal of Geophysical Research*, *103*, 32257−32275.

Kobayashi, H., Delbart, N., Suzuki, R., & Kushida, K. (2010). A satellite-based method for monitoring seasonality in the overstory leaf area index of Siberian larch forest, *Journal of Geophysical Research*, *115*, G01002, doi:10.1029/2009JG000939.

Lucht, W., & Lewis, P. (2000). Theoretical noise sensitivity of BRDF and albedo retrieval from the EOS-MODIS and MISR sensors with respect to angular sampling. *International Journal of Remote Sensing*, *21*, 81–98.

Lucht, W., Schaaf, C. B., & Strahler, A. H. (2000). An algorithm for the retrieval of albedo from space using semiempirical BRDF models. *IEEE Transactions on Geoscience and Remote Sensing*, *38*, 977–998.

Morisette, J., Baret, F., Privette, J. L., Myneni, R. B., Nickeson, J., Garrigues, S., Shabanov, N.V., Weiss, M., Fernandes, R.A., Leblanc, S.G., Kalacska, M., Sanchez-Azofeifa, G.A., Chubey, M., Rivard, B., Stenberg, P., Rautiainen, M., Voipio, P., Manninen, T., Pilant, A.N., Lewis, T.E., Iiames, J.S., Colombo, R., Meroni, M., Busetto, L., Cohen, W.B., Turner, D.P., Warner, E.D., Petersen, G.W., Seufert, G., & Cook, R. (2006). Validation of global moderate resolution LAI Products: A framework proposed within the CEOS Land Product Validation subgroup. *IEEE Transactions on Geoscience and Remote Sensing*, *44*, 1804−1817.

Myneni, R., Hoffman, S., Knyazikhin, Y., Privette, J., Glassy, J., Tian, Y., Wang, Y., Song, X., Zhang, Y., Smith, G. R., Lotsch, A., Friedl, M., Morisette, J. T., Votava, P., Nemani, R. R., & Running, S. W. (2002). Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data. *Remote Sensing of Environment*, *83*, 214–231.

Privette, J. L., Tian, Y., Roberts, G., Scholes, R. J., Wang, Y., Caylor, K.K., Frost, P., & Mukelabai, M. (2004). Vegetation structure characteristics and relationships of Kalahari woodlands and savannas. *Global Change Biology*,10(3), 281–291.

Privette, J. L., Mukelabai, M. M., & Huemmrich, K. F. (2005). SAFARI 2000 Leaf Area Measurements at the Mongu Tower Site, Zambia, 2000-2002. Data set. Available on-line [http://daac.ornl.gov/] from Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, U.S.A. doi:10.3334/ORNLDAAC/781.

Reichenbach, S. E., & Geng, F. (2003). Two dimensional cubic convolution. *IEEE Transactions on Geoscience and Remote Sensing*, *12*, 857−865.

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. *Journal of Geophysical Research*, *97(D18)*, 20455–20468.

Roy, D. P., Lewis, P., Schaaf, C., Devadiga, S., & Boschetti, L. (2006). The global impact of cloud on the production of MODIS bi-directional reflectancemodel based composites for terrestrial monitoring. *IEEE Geoscience and Remote Sensing Letters*, *3*, 452−456.

Samanta, A., Ganguly, S., Schull, M. A., Shabanov, N., Knyazikhin, Y., & Myneni, R. B. (2009). Collection 5 MODIS LAI/FPAR Products, *4th Global Vegetation Workshop*, Numerical Terradynamic Simulation Group, Montana, USA.

Savitzky, A., & Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, *36(8)*, 1627−1639.

Schaaf, C. B., Gao, F., Strahler, A. H., Lucht, W., Li, X., Tsang, T., Strugnell, N. C., Zhang, X., Jin, Y., Muller, J.-P., Lewis, P., Barnsley, M., Hobson, P., Disney, M., Roberts, G., Dunderdale, M., Doll, C., d'Entremont, R., Hu, B., Liang, S., Privette, J. L., & Roy, D. P. (2002). First operational BRDF, albedo and nadir reflectance products from MODIS. *Remote Sensing of Environment*, *83*, 135−148.

Shabanov, N. V., Huang, D., Yang, W., Tan, B., Knyazikhin, Y., Myneni, R. B., Ahl, D. E., Gower, S. T., Huete, A. R., Aragao, L. E. O. C., & Shimabukuro, Y. E. (2005). Analysis and optimization of the MODIS leaf area index algorithm retrievals over broadleaf forests. *IEEE Transactions on Geoscience and Remote Sensing*, *43(8)*,1855−1865.

Verbesselt, J., Hyndman, R., Newnham, G., & Culvenor, D. (2010). Detecting trend and seasonal changes in satellite image time series. *Remote Sensing of Environment*, 114, 106-115.

Verger, A., Baret, F., & Weiss, M. (2008). The performances of neural networks for deriving LAI estimates from existing CYCLOPES and MODIS products. *Remote Sensing of Environment*, *112*, 2789–2803.

Verger, A., Baret, F., & Camacho, F. (2011). Optimal modalities for radiative transfer-neural network estimation of canopy biophysical characteristics: Evaluation over an agricultural area with CHRIS/PROBA observations, *Remote Sensing of Environment*, *115*, 415−426.

Vermote, E. F., Saleous, N. Z. E., Justice, C. O., Kaufman, Y. J., Privette, J., Remer, L., Roger, J. C., & Tanre, D. (1997). Atmospheric correction of visible to middle infrared EOS-MODIS data over land surface, background, operacional algoirthm and validation. *Journal of Geophysical Research*, *102(14)*, 17131–17141.

Viovy, N., Arino, O., & Belward, A. (1992). The Best INdex Slope Extraction (BISE): a method for reducing noise in NDVI timeseries. *International Journal of Remote Sensing*, *13(8)*, 1585–1590.

Weiss, M., Baret, B., Smith, G. J., Jonckheere, I., & Coppin, P. (2004). Review of methods for in situ leaf area index (LAI) determination. Part II. Estimation of LAI, errors and sampling. *Agricultural and Forest Meteorology*, *121*, 37−53.

Weiss, M., Baret, F., Garrigues, S., & Lacaze, R. (2007). LAI and fAPAR CYCLOPES global products derived from VEGETATION. Part 2: Validation and comparison with MODIS Collection 4 products. *Remote Sensing of Environment*, *110*, 317−331.

Yang, W., Tan, B., Huang, D., Rautiainen, M., Shabanov, N. V., Wang, Y., Privette, J. L., Huemmrich, K.F., Fensholt, R., Sandholt, I., Weiss, M., Ahl, D. E., Gower, S. T., Nemani, R. R., Knyazikhin, Y., & Myneni, R. B. (2006). MODIS leaf area index products: From validation to algorithm improvement. *IEEE Transactions on Geoscience and Remote Sensing*, *44(7)*, 1885–1898.

WWW Sites

WWW1: Committee on Earth Observation Satellites (CEOS) Working group on calibration and validation (WGCV) Land Product Validation (LPV) Subgroup. http://lpvs.gsfc.nasa.gov/lai_intercomp.php

WWW2 : Pôle d'Observation des Surfaces continentales par TELédétection. http://postel.mediasfrance.org

WWW3: Land processes distributed active archive center. http://edcdaac.usgs.gov

WWW4: Oak Ridge National Laboratory Distributed Active Archive Center. http://daac.ornl.gov

WWW5: Validation of Land European Remote sensing Instruments. http://www.avignon.inra.fr/valeri

Table 1. Temporal sampling interval and resolution (length of the compositing window) of the considered LAI products. TSGF stands for temporal smoothing and gap filling.

	Sampling	Resolution	Description
MODIS	8 days	8 days	Original MODIS LAI product (MOD15A2)
SmoothedMOD	8 days	≤128 days	MOD15A2 LAI product processed using TSGF
MODrhoMODlai	16 days	16 days	MODIS LAI estimated from MOD43B4 product
CYCrhoMODlai	10 days	30 days	MODIS LAI estimated from CYCLOPES product
FUSION	8 days	≤128 days	Fusion of MODrhoMODlai and CYCrhoMODlai using TSGF

Table 2. Number of cases (pixel×date) used per biome class in the training data set and associated RMSE values as compared to the original MODIS products.

Biome class	Nb. data	RMSE	
		MODrhoMODlai	CYCrhoMODlai
Shrubs	784	0.12	0.11
Savanna	926	0.48	0.40
Grasses & Cereal Crops	1476	0.19	0.17
Broadleaf Crops	797	0.17	0.15
Needleleaf F.	1380	0.68	0.63
Deciduous Broadleaf F.	413	0.73	0.72
Evergreen Broadleaf F.	375	0.89	0.85

Table 3. Fraction of invalid data per biome for the several LAI products investigated (Table 1) over the BELMANIP sites during the 2001–2003 period.

Biome	MODIS	SmoothedMOD	CYCrhoMODlai	MODrhoMODlai	FUSION
Shrubs	0.35	0.21	0.12	0.16	0.017
Savanna	0.23	0.07	0.12	0.15	0.025
Grasses & Cereal Crops	0.28	0.07	0.07	0.11	0.003
Broadleaf Crops	0.26	0.05	0.11	0.13	0.006
Needleleaf F.	0.45	0.20	0.34	0.14	0.031
Deciduous Broadleaf F.	0.34	0.07	0.23	0.15	0.016
Evergreen Broadleaf F.	0.43	0.18	0.41	0.34	0.17
TOTAL	0.34	0.13	0.20	0.16	0.037

Table 4. Description of the sites selected for the temporal profiles (Fig. 7).

Site name	Location	Lat. (º)	Lon. (º)	MODIS land cover	Actual land cover	Reference for ground data
Mongu	Zambia	-15.44	23.25	Savanna	Kalahari woodland	Privette et al. (2005), WWW4
Dahra	Senegal	15.41	-15.43	Grasses & Cereal Crops	Savanna	Fensholt et al. (2004)
Hyytiälä	Finland	61.85	24.29	Evergreen Needleleaf F.	Mixed boreal forest	WWW5
Soignes	Belgium	50.78	4.42	Deciduous Broadleaf F.	Deciduous broadleaf forest	WWW5
Counami	French Guiana	5.34	-53.24	Evergreen Broadleaf F.	Tropical rain forest	WWW5

Table 5. Root mean square error (RMSE) computed between MODIS, SmoothedMOD, MODrhoMODlai, CYCrhoMODlai and FUSION products for forest (top triangular matrix) and non-forest (bottom triangular matrix) BELMANIP sites during the 2001-2003 period.

Forest / Non-Forest	MODIS	SmoothedMOD	MODrhoMODlai	CYCrhoMODlai	FUSION
MODIS	0.00	0.79	1.32	1.31	1.29
SmoothedMOD	0.35	0.00	1.10	1.21	1.03
MODrhoMODlai	0.86	0.80	0.00	0.31	0.40
CYCrhoMODlai	0.82	0.76	0.67	0.00	0.39
FUSION	0.84	0.78	0.19	0.18	0.00

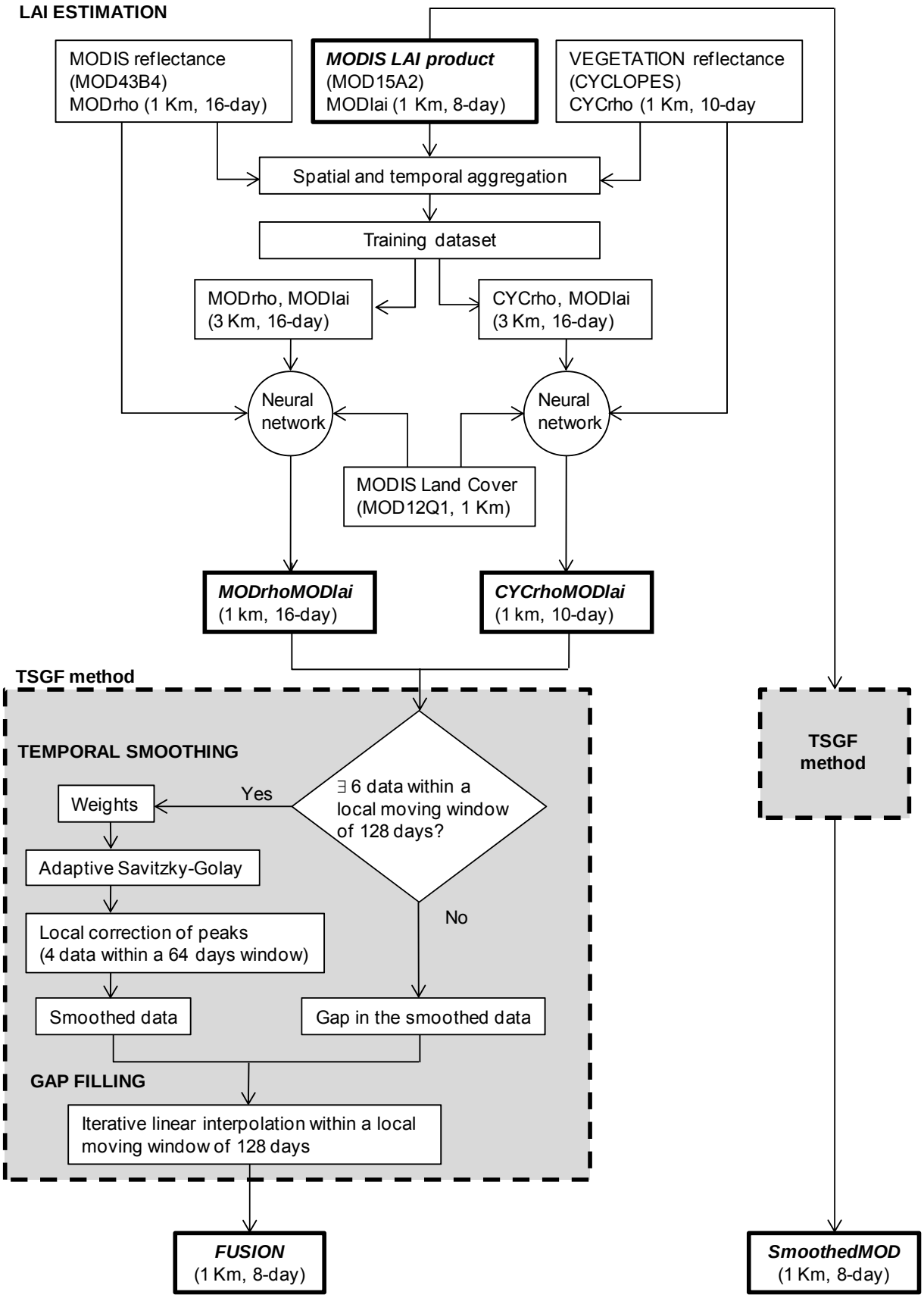

Fig. 1. Flow chart of the algorithm for the estimation of the LAI products (Table 1).

Version définitive du manuscrit publié dans / Final version of the manuscript published in : Remote Sensing of Environment, 2011, In Press, DOI: 10.1016/j.rse.2011.05.

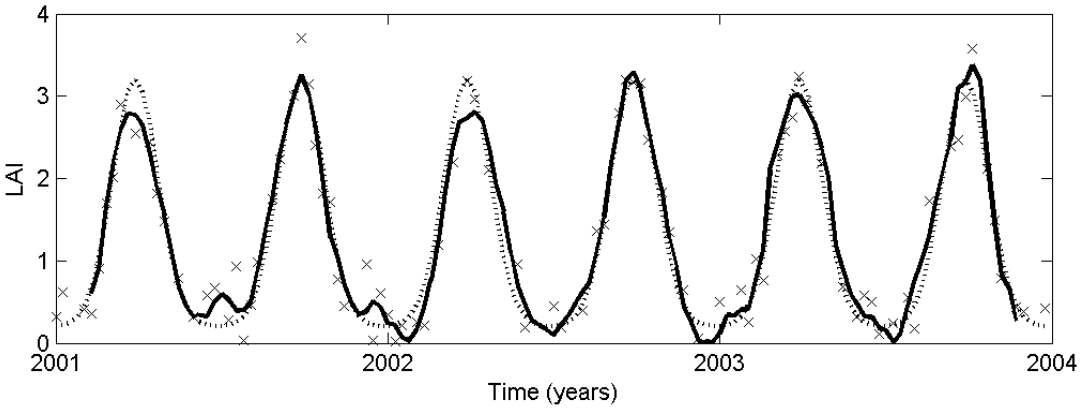

Fig. 2. Simulated 8-day LAI time series of a tropical cropland (dotted line). Simulated observations by introducing a white noise of 0.3 standard deviation and a percentage of gaps of 0.3 to the original time series (crosses). Reconstructed time series by applying the TSGF method to the observations (solid line).

Version définitive du manuscrit publié dans / Final version of the manuscript published in : Remote Sensing of Environment, 2011, In Press, DOI: 10.1016/j.rse.2011.05.

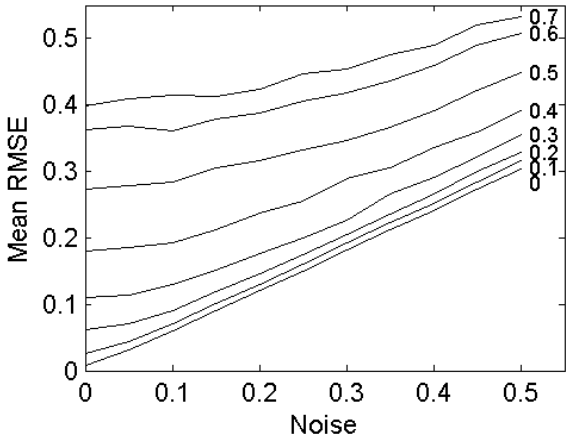

Fig. 3. Mean of 500 RMSEs for the TSGF reconstructed LAI time series as a function of the σ noise level ($LAI_{noisy} = LAI + \aleph(0,\sigma)$) for a fraction of missing data varying between 0 and 0.7.

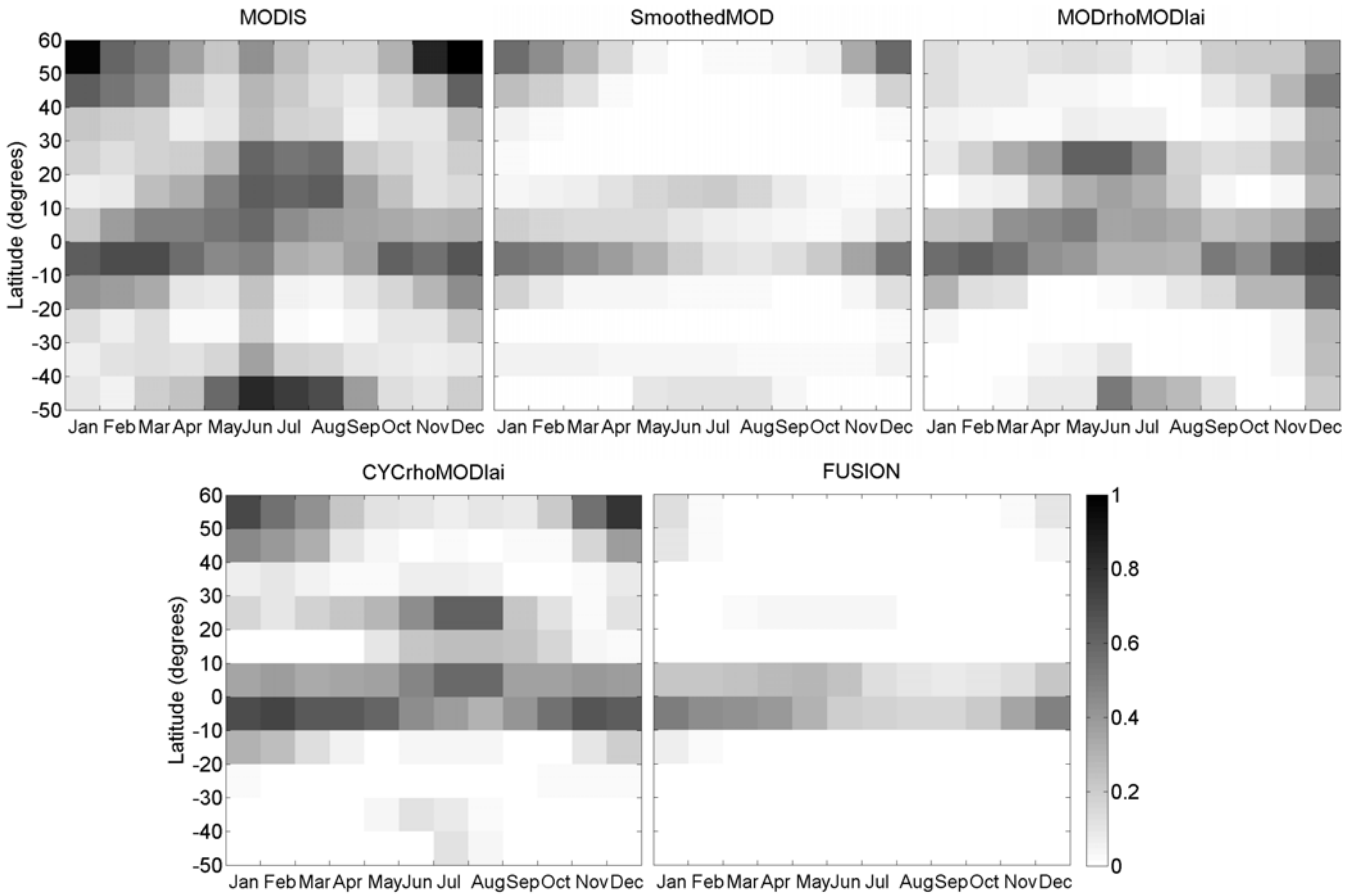

Fig. 4. Fraction of invalid data as a function of the latitude and the date of acquisition for the several LAI products investigated (Table 1). Evaluation over the BELMANIP sites during the 2001–2003 period in 10º × 1-month cells.

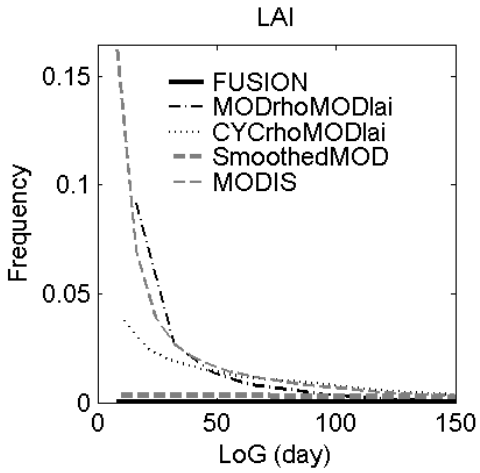

Fig. 5. Histogram of length of gaps (LoG).

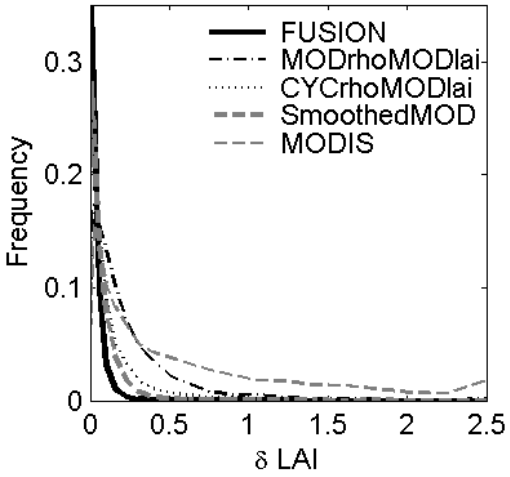

Fig. 6. Histogram of the δ absolute value for the original and the developed MODIS LAI products.

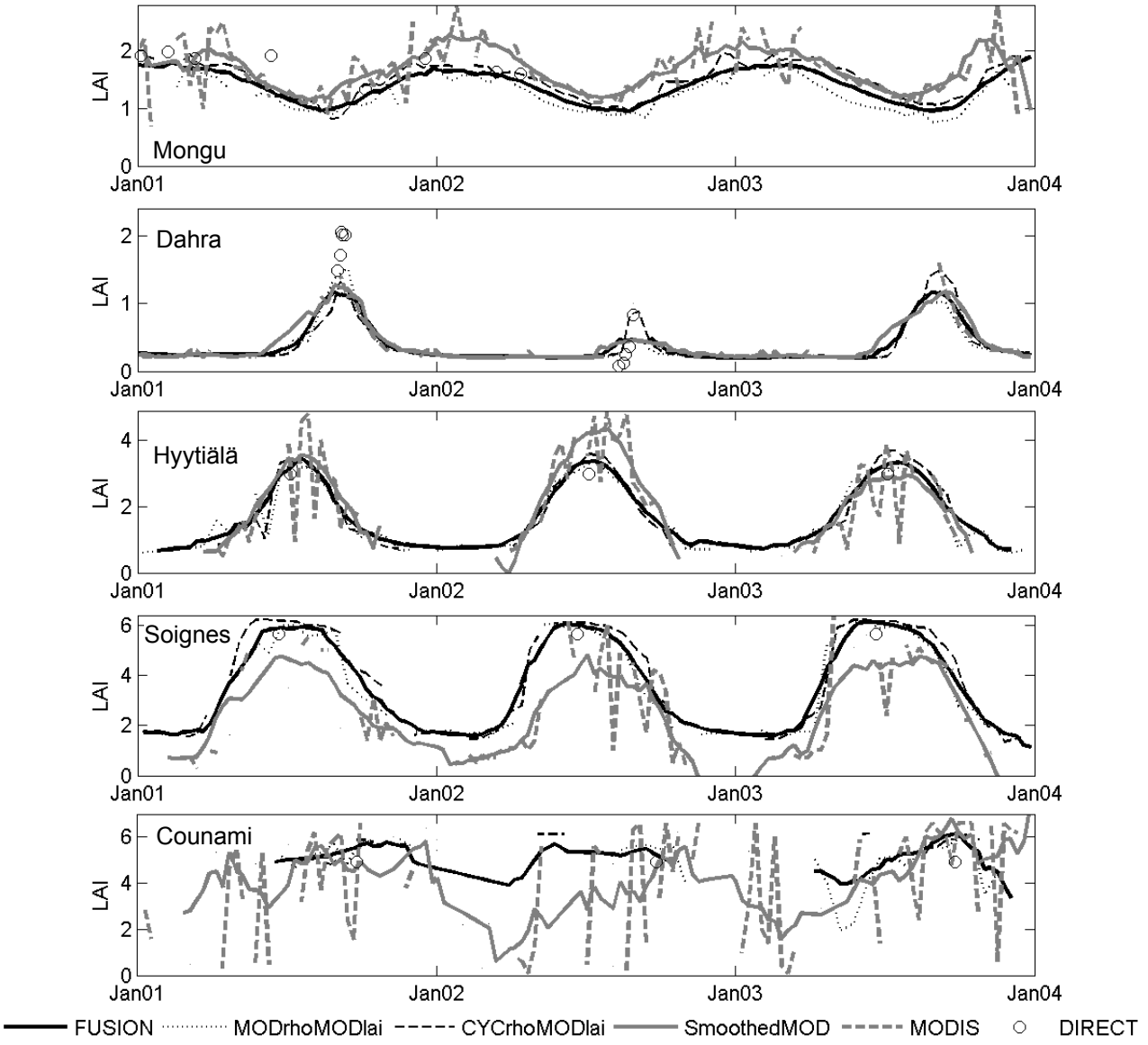

Fig. 7. Temporal evolution of MODIS, SmoothedMOD, MODrhoMODlai, CYCrhoMODlai and FUSION LAI products for the period 2001–2003 over the central pixel of five BELMANIP sites representing five biome classes (Table 4).

Version définitive du manuscrit publié dans / Final version of the manuscript published in : Remote Sensing of Environment, 2011, In Press, DOI: 10.1016/j.rse.2011.05.

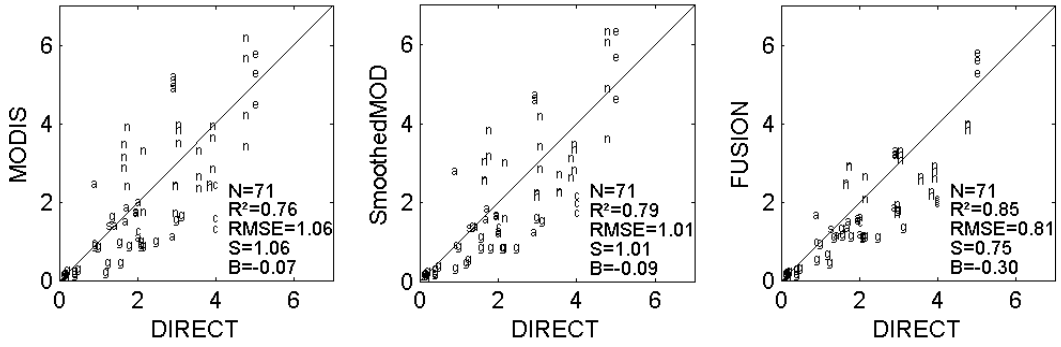

Fig. 8. Comparison of MODIS, SmoothedMOD and FUSION LAI products with direct ground measurements of true LAI. Letter markers correspond to the biome classes: shrubs (s), savanna (a), grasses and cereal crops (g), broadleaf crops (b), needleleaf forest (n), deciduous broadleaf forest (d) and evergreen broadleaf forest (e). The statistics are: number of samples (N), correlation coefficient (R^2), root mean square error (RMSE), standard deviation (S) and bias (B).