



Quality assessment and improvement of temporally composited products of remotely sensed imagery by combination of VEGETATION 1 and 2 images

Olivier Hagolle, A. Lobo, P. Maisongrande, F. Cabot, B. Duchemin, A. De Pereyra

► To cite this version:

Olivier Hagolle, A. Lobo, P. Maisongrande, F. Cabot, B. Duchemin, et al.. Quality assessment and improvement of temporally composited products of remotely sensed imagery by combination of VEGETATION 1 and 2 images. *Remote Sensing of Environment*, Elsevier, 2005, 94 (2), pp.172-186. <10.1016/j.rse.2004.09.008>. <ird-00389597>

HAL Id: ird-00389597

<http://hal.ird.fr/ird-00389597>

Submitted on 29 May 2009

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.



Quality assessment and improvement of temporally composited products of remotely sensed imagery by combination of VEGETATION 1 and 2 images

O. Hagolle^{a,*}, A. Lobo^b, P. Maisongrande^c, F. Cabot^a, B. Duchemin^c, A. De Pereyra^a

^aCentre National d'Etudes Spatiales (CNES), bpi 811, 18, avenue Edouard Belin, 31401 Toulouse Cedex 9, France

^bInstitut de Ciències de la Terra "Jaume Almera" (CSIC), Lluís Solé Sabaris S/N, 08028 Barcelona, Spain

^cCentre d'Etudes Spatiales de la Biosphère (CESBIO), bpi 2801, 18, avenue Edouard Belin, 31401 Toulouse Cedex 9, France

Received 19 December 2003; received in revised form 16 September 2004; accepted 16 September 2004

Abstract

Global temporal composites of surface reflectances are among the most commonly used products of wide field-of-view satellite-borne instruments such as the Advanced Very High Resolution Radiometer (AVHRR), the MODerate resolution Imaging Spectroradiometer (MODIS) and VEGETATION. The multi-temporal and spatial consistencies of these composites are key elements for their usefulness. In this paper, we use two different criteria to evaluate the quality of existing and new temporal composite products in SPOT-VEGETATION imagery. The first criterion, based on variograms, analyses the spatial characteristics of composite images, and the second one evaluates the quality of the time series based on the analysis of simultaneous imagery from VEGETATION 1 and VEGETATION 2. Thanks to these criteria, we show that the standard deviation of the errors that affect the surface reflectances of current composite products can be reduced by a factor greater than 2 using improved algorithms detailed in this paper. Finally, we produce multi-instrument composites by integrating images from both VEGETATION instruments to further improve the composite products.

© 2004 Elsevier B.V. All rights reserved.

Keywords: VEGETATION; AVHRR; MODIS

1. Introduction

The Earth observation satellite SPOT 5 was successfully launched on May 4th, 2002, with VEGETATION 2 instrument onboard. VEGETATION 2 joined up SPOT4/VEGETATION 1 (launched in April 1998) on the same orbit. Both VEGETATION instruments are of multi-spectral push-broom type with a very wide field of view (more than 2000 km), and a geometrical resolution of about 1 km at Nadir. Each VEGETATION instrument provides an almost daily acquisition of the whole continental surfaces at four wavelengths termed B0 (blue), B2 (red), B3 (near-infrared), Medium InfraRed (MIR) and centred around 460, 670, 840

and 1640 nm, respectively. Like many wide field of view earth observation systems such as the Advanced Very High Resolution Radiometer (AVHRR), the MODerate resolution Imaging Spectroradiometer (MODIS), and the MEdium Resolution Imaging Spectrometer (MERIS), VEGETATION offers a broad range of products that correspond to an increasing degree of data processing. These products are classified in three levels, according to the definition by the Committee on Earth Observation Satellites (CEOS) (Kingwell et al., 1996).

Level 1 products provide users with geolocated and calibrated top of atmosphere (TOA) reflectances acquired during a time period that does not exceed the duration of one orbit. The Level 1 processing does not make any assumption on the physical nature of the observed target, and only corrects for sensor artefacts. Level 1 products from VEGETATION are called VGT-P products (P for Physical).

* Corresponding author. Tel.: +33 5 6128 2135; fax: +33 5 6127 3167.

E-mail address: Olivier.Hagolle@cnes.fr (O. Hagolle).

To derive variables describing the vegetation cover, it is necessary to convert the TOA reflectances into land surface reflectances. This is done during the Level 2 processing of VEGETATION data. Level 2 products are not distributed to users by the VEGETATION ground segment, but are a step towards the production of level 3 products.

VEGETATION Level 3 surface reflectance products, often called composite products or syntheses, are obtained using all the level 2 data acquired during periods of 10 days. Level 3 products are useful for applications that do not require daily observations: compared to level 2 products, they provide global or regional maps that minimise data volume and cloud cover. Currently, two different compositing methods are used, resulting in products named S10 and D10. The differences between the algorithms that deliver S10 and D10 products are described in Section 2.1.

Many error sources degrade the estimates of level 3 surface reflectances: instrumental noise, calibration or geometric registration errors, interpolation errors during geometrical projection, cloud screening, atmospheric correction errors, directional normalisation errors, or surface reflectance variation during the compositing period. Measuring errors in level 3 products is complicated because field measurements of reflectance at a 1-km resolution are not available except for some rare occasions and for very uniform sites such as deserts or polar regions, and rarely for vegetated areas. In situ measurements with a few meters resolution are available at some very rare sites (Privette et al., 2002; Weiss et al., 2001), and for very few dates, but generally, biophysical variables such as Leaf Area Index are measured instead of reflectance. It is also worth noting that the accuracy of field measurements of reflectance is not perfect, and that errors might often be added when scaling up data acquired at a few meter resolution to at least 1 km². Although the comparison of temporally composited reflectance to field measurements of reflectance is useful to detect biases, it is not an efficient way to characterise the errors that affect temporally composited products.

A number of authors (Duchemin & Maisongrande, 2002; Goward et al., 1993 for instance) have used the smoothness of the variation with time of reflectance (or vegetation index) as a criterion to judge the quality of products, but it is often uneasy to decide whether a given variation represents an artefact or is due to an actual change in the surface. For this reason, the above authors use regional averages, and do not try to assess the quality on a pixel by pixel basis.

Nowadays, both VEGETATION sensors can observe the same point twice a day with a 30' time lag (Fig. 1). This “extra acquisition” has a twofold interest: first, the comparison of products issued from each sensor gives the opportunity to quantify the noise that affects time series of temporally composited products, using the quality

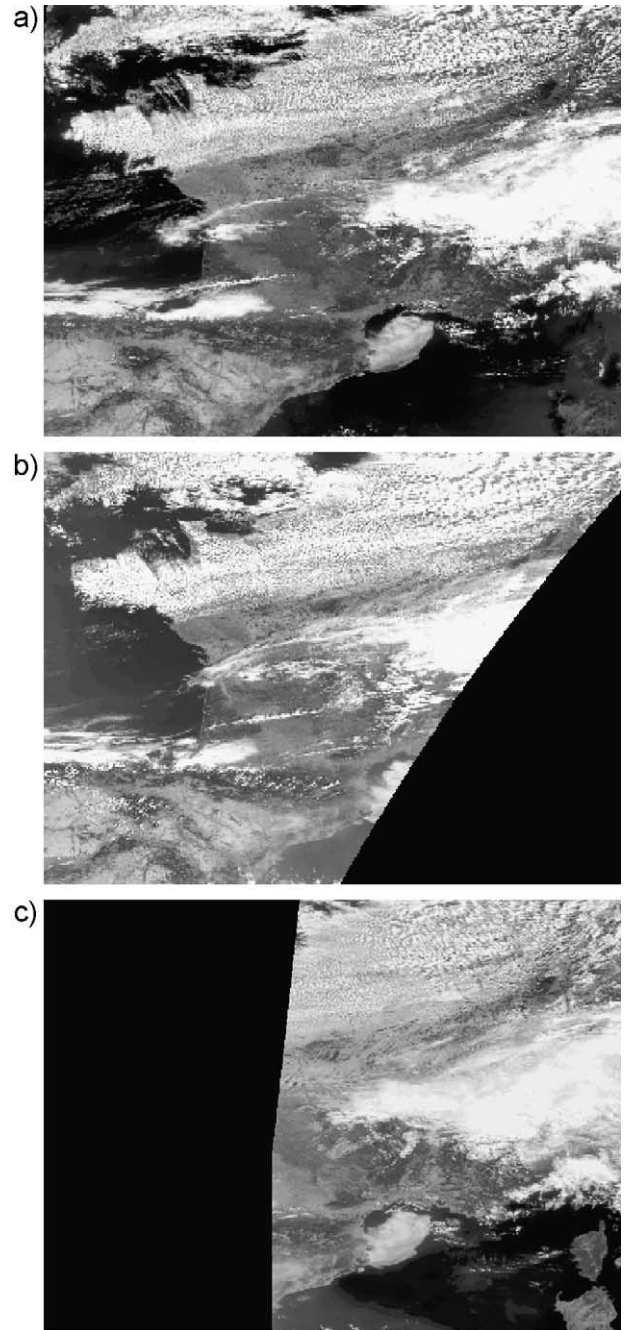


Fig. 1. Images acquired above France on July 21st, 2002 by Vegetation 1 (a) and 2 (b, c), at approximately (a) 10.30 h, (b) 10.00 h (c) 11.40 h.

criterion defined below. Second, having two satellites increases the amount of available data and consequently adds more information to the directional variations of reflectance.

This paper begins with a description of VEGETATION operational level 3 algorithms, followed by several proposed improvements. Then, two criteria to assess the quality of products are presented, and finally used to assess the performances of current and enhanced VEGETATION composite products.

2. Operational level 3 algorithm and possible improvements

2.1. Operational algorithms

Currently available S10 and D10 products differ in the compositing strategy. Up to level 2, the same processing line is used to perform cloud screening and atmospheric correction. The cloud screening for VEGETATION products is based mainly on a threshold on B0 TOA reflectance because the contrast between clouds and Earth surface is the greatest in the blue band: if reflectance is above 0.246, the pixel is not cloud-free (Kempeneers et al., 2000). The atmospheric correction is done using the SMAC method (Rahman & Dedieu, 1994) with METEO FRANCE analyses as input for the pressure and water vapour content, and with a month average total column of ozone. The aerosol content is determined following a method described in Duchemin et al. (2002) and Maisongrande et al. (2004). This method is based on an empirical relationship between the ratio of blue and MIR bands and the Normalised Difference Vegetation Index NDVI (see Eq. (1)). The algorithm searches the aerosol optical depth for which the atmospherically corrected reflectances agree at best to this relationship.

2.1.1. S10 compositing method

The most common technique to produce reflectance level 3 products over lands is the Maximum Value Composite (MVC) (Holben, 1986; Tarpley et al., 1984). This method has been applied for years to AVHRR, and is also used to produce VEGETATION 10-day (S10) composite products. For a given time window, the MVC selects the measurements of the date when the Normalised Difference Vegetation Index (NDVI) is maximum. This method is only applicable over land, and helps rejecting pixels affected by clouds because their NDVI is always lower than that of cloud free pixels. The NDVI definition is recalled in Eq. (1), where ρ_{red} (resp ρ_{NIR}) is the TOA reflectance in the red part of spectrum (resp Near-Infrared). For VEGETATION, B2 band is the red band and B3 the NIR band:

$$\text{NDVI} = \frac{\rho_{\text{NIR}} - \rho_{\text{red}}}{\rho_{\text{NIR}} + \rho_{\text{red}}} \quad (1)$$

Among the many vegetation indexes, NDVI was designed to reduce the noise that affects reflectance. If the same error occurs in the NIR and red channels, the error is cancelled. This is quite true for cloud shadows or thin cirrus clouds, and, since the anisotropy of reflectances varies slowly with the spectral bands, the NDVI varies also slowly with the viewing direction. While NDVI is not too sensitive to directional effects, the reflectance composites derived from the MVC methods are very noisy. Selecting the maximum values of the NDVI minimises the selection of cloudy and/or heavy aerosol pixels but it does not take directional effects into account. When the observation and

illumination geometries vary, it is fairly common to observe reflectance variations above 50% (Roujean et al., 1992), that have a deleterious impact on reflectance time series.

2.1.2. D10 compositing method

To cope with directional effects, a better composite product, the 10-day Directional composite (D10), has been developed by the VEGETATION project to enhance reflectance composites. The MODIS production entity (Schaaf et al., 2002) consider a similar approach. S10 production was maintained to preserve data continuity. The D10 algorithm is based on four elementary steps as detailed in Duchemin et al. (2002). First, for each pixel, the algorithm selects the 10 most recent cloud free acquisitions, even if the older ones are not within the 10-day compositing period. Second, the series of ten clear observations is used to fit a Bidirectional Reflectance Distribution Function (BRDF) that specifies surface scattering as a function of illumination and view angles. Third, the fitted BRDF is used to normalise the non-cloudy reflectances observed during the 10-day period, to Nadir viewing direction and to the solar elevation observed at 10.30-h local time for the median day of the compositing period. Finally, the D10 value is obtained by averaging the normalised, cloud free reflectances found in the last 10 days (it is computed even if only one observation is available, resulting in a high sensitivity to the quality of this observation).

The model from Roujean et al. (1992) describes the BRDF as a linear combination of three terms.

$$\rho(\theta_s, \theta_v, \phi) = k_0 + k_1 f_1(\theta_s, \theta_v, \phi) + k_2 f_2(\theta_s, \theta_v, \phi) \quad (2)$$

θ_s, θ_v, ϕ are, respectively, the solar zenith, view zenith and relative azimuth angles, while f_1 and f_2 are the geometric and volume scattering functions respectively, and k_i are the weighting parameters of the f_i functions. The method tries to minimise the cost function C defined as:

$$C = \sum_{i=1}^N (\rho_i - \hat{\rho}_i) \quad (3)$$

where N is the number of observations available during the synthesis period, ρ_i are the observations, and $\hat{\rho}_i$ are the best-fit values of the Roujean model.

2.2. Improvements

Although directional syntheses represent a considerable advance in the art of optical data processing, they still remain very sensitive to the quality of cloud filtering and atmospheric corrections. As shown in Section 4.2, the performances of operational S10 and D10 products are far from being perfect. To test new algorithms, a qualification line has been implemented at CNES. The general framework and main algorithms are inspired by the VEGETATION official processing lines, but we also have the possibility to plug in new and enhanced algorithms.

2.2.1. Improved cloud screening

Defining a good cloud threshold, valid for the entire world, is a very difficult task (Stowe et al., 1991). The official cloud screening for VEGETATION uses a constant threshold on B0 TOA reflectance, regardless of the surface cover or of the pixel viewing geometry. To avoid discarding pixels when the atmospheric path radiance is high, the threshold is loose, and many clouds remain undetected.

The cloud threshold that we propose is based on the surface reflectance in the B0 spectral band: ρ_S (B0). It is computed after correcting for Rayleigh scattering and molecular absorption using the SMAC method (Rahman & Dedieu, 1994). Our threshold also depends on surface cover: a land-cover classification (De Fries et al., 1998) is used as input to determine if a given pixel lies on a desertic or dry savanna region. In such a case, the pixel is considered cloudy if ρ_S (B0) > 0.23. As some regions are dry during part of the year only, and since the B0 reflectance is higher during the dry season, our threshold value also depends on the NDVI. If the NDVI is below 0.2, a threshold of 0.14 is used; if the NDVI is above 0.2, the threshold is set to 0.09. Cloud shadows are also discarded taking into account viewing and solar angles, under the assumption that clouds are at a 5-km altitude. As an additional precaution, pixels distant by less than 3 km to a cloudy pixel or to a shadow are classified as cloudy as well. For the time being, our algorithm considers snow covered pixels as cloudy pixels, although VEGETATION enables to discriminate snow and clouds thanks to the MIR band. The above thresholds have been tuned and tested with several images, and the resulting improvement for composite products has been checked using the temporal criterion exposed in next section (see the results in Section 4.2.2).

2.2.2. Improved compositing method

As it will be seen Section 4, the D10 compositing method is very sensitive to residual clouds that are still present despite our strict screening method. For a given pixel and a given 10-day period, the distribution of the viewing angles inside the field of view is irregular, because of the cloud cover. If an undetected low cloud affects a viewing direction at a high viewing angle and if no more measurements are available near that viewing angle, the directional model fitting such a series will be biased (see Fig. 2). We have studied two ways of avoiding such errors:

- (i) using a priori information: a confidence interval for the parameters of the Roujean model can be used to constrain the least-squares fit. Data leading to model parameters too far from the constrained value are discarded, using an iterative process;
- (ii) increasing the number of observations with longer compositing periods or simultaneous VEGETATION 1 and 2 data.

2.2.2.1. Using a priori information. In order to constrain the Roujean model, we add two conditions to the least-squares minimisation system:

$$k_1 = C1(\lambda), \quad k_2 = C2(\lambda)$$

As a result, the cost function is modified as follows:

$$C = \left(\sum_{i=1}^N \frac{(\rho_i - \hat{\rho}_i)^2}{\sigma_i^2} \right) + \frac{(k_1 - C1(\lambda))^2}{\sigma_{k_1}^2} + \frac{(k_2 - C2(\lambda))^2}{\sigma_{k_2}^2} \quad (4)$$

where $1/\sigma_i^2$ are the weights of the observations (related to the standard deviation of errors), $1/\sigma_{k_1}^2$ and $1/\sigma_{k_2}^2$ are the weights of the constraints.

Evidently, the choice of the constraint values has some influence on the results. In our study, a very simple option was used: $C1(\lambda)$ and $C2(\lambda)$ do not depend either on the pixel or on the date. $C1$ and $C2$ are determined by running the unconstrained algorithm for various types of landscapes and by averaging the retrieved k_1 and k_2 values. To discard erroneous values before computing the average, we selected only the pixels for which k_1 and k_2 values derived from VEGETATION 1 and 2 products gave very close values. Of course, using only one single value for surface covers ranging from deserts to dense forest is a bit hazardous: therefore, we have chosen to give a small weight to the constraint equations by choosing $\sigma_{k_1}^2 = \sigma_{k_2}^2 = 4\sigma_i^2$. In the near future, we intend to introduce pixel-dependent constraints: to produce the synthesis for a given period, we could use as a constraint the values of k_1 and k_2 obtained for the previous one. But the main issue is to ensure that the process does not diverge.

To mitigate the problem of undetected clouds, it is also possible to use a priori knowledge on the statistics of errors in the observations. As a matter of fact, the distribution of errors that affect surface reflectances differs considerably from a gaussian curve. The instrumental noise, and mainly the atmospheric correction errors contribute to a gaussian with a standard deviation ranging from 2% to 10%, depending on the bands. But the main noise contributor is the cloud detection error: positive errors ranging from 10% to 50% may be observed, corresponding to thin or broken clouds, while negative errors, related to cloud shadows, are much less frequent.

Having constrained the BRDF model, it is easier, in most cases, to detect and discard cloud-contaminated pixels (Fig. 2). The detection is done using the B0 band because it enables the best distinction between clouds and continental surfaces. If the standard deviation of errors σ of the model fit is above a threshold, our algorithm discards all the observations that are greater than the adjusted model values plus σ . At this stage, many previously undetected cloudy pixels are discarded. Then a second fit and a new σ value is computed (smaller than

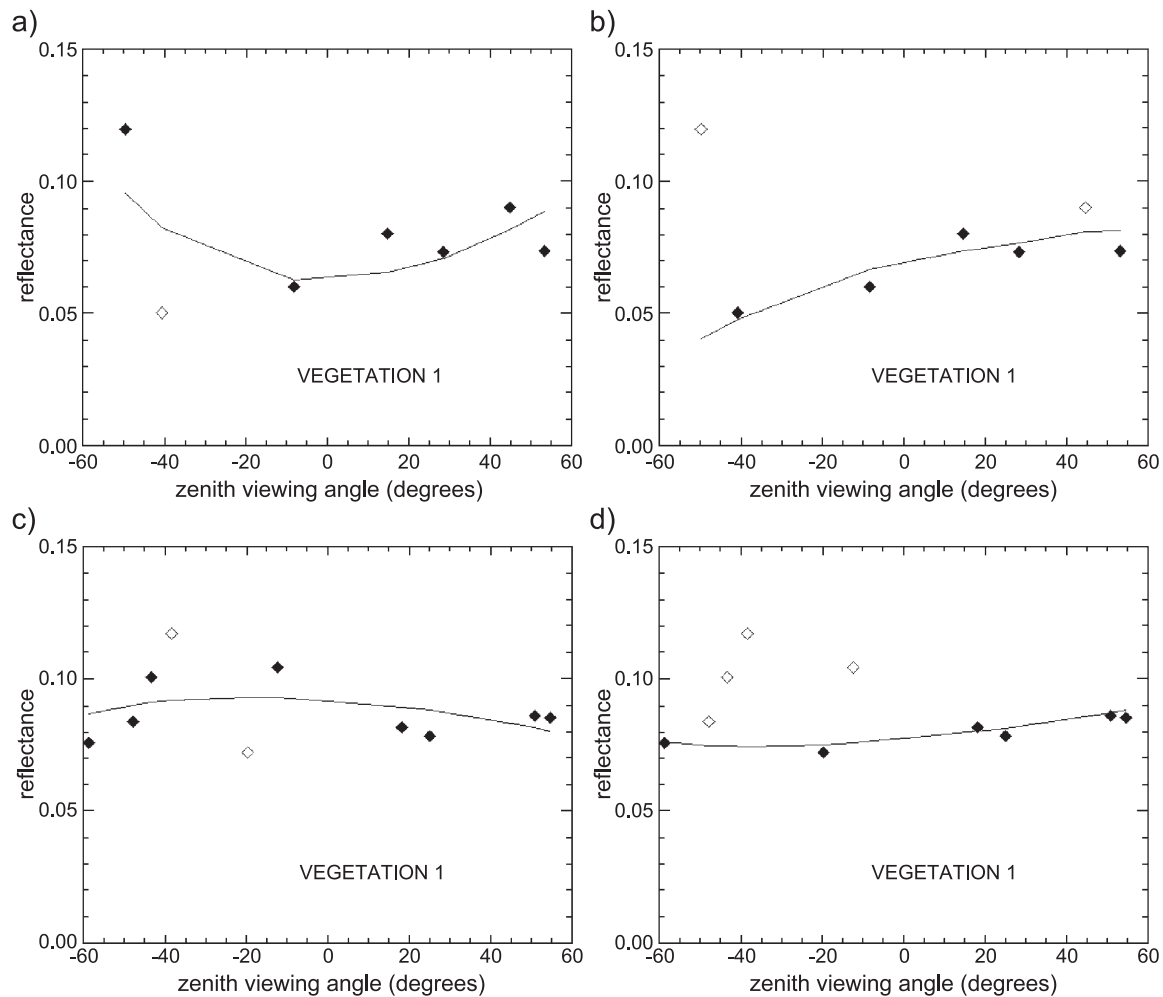


Fig. 2. Left column, reflectances (diamonds) and the retrieved BRDF model (solid lines) for two different pixels (Côte d’Ivoire, December 2002) in spectral band B2, during a period of 15 days, using the D10 method. Right column, retrieved model using the new proposed method that uses a priori information. In all plots, unfilled symbols correspond to outliers that were discarded by the iterative process.

the previous one); all the pixels with a difference to the fitted model greater than $1.5 \cdot \sigma$ are discarded whatever the sign of the difference. This last stage is iterated until no pixel is discarded anymore, or until the number of remaining pixels is lower than 3.

In the D10 method, pixels with a strong deviation from the model fit were also discarded, but with no distinction between higher and lower values. The improvement due to this new method is illustrated for an example pixel in Fig. 2 and the impact on composited products is evaluated in Section 4.2.2.

2.2.2.2. Increasing the number of observations. To better constrain the BRDF model, it is wise to increase the amount of available data, since the number of valid observations acquired during 10 days is often insufficient. The compositing period duration can be increased, keeping in mind that if the time period is too long, surface variation will be observed within the time window. We tested here 15- and 30-day periods instead of 10 (see Section 4).

The simultaneous availability of VEGETATION 1 and 2 data also gives the opportunity to double the number of observations to enhance the directional fit, without increasing the compositing period. Images from both sensors are used in the compositing just as if they came from the same instrument.

3. Performance criteria

In this study, we used two complementary criteria to estimate the quality of level 3 products. The first one, called “temporal criterion” uses both VEGETATION instruments to estimate the noise on time series. The second one, called “spatial criterion” estimates the spatial noise added to images.

3.1. Temporal criterion

SPOT 4 and SPOT 5 are in the same geo-synchronous polar orbit at an altitude of 830 km, SPOT 5

crossing the equator 30 min before SPOT 4. Because of the Earth rotation, the track of SPOT 4 is about 750 km westward (at the equator) from the SPOT5. Since VEGETATION instruments have a 2200 km field of view, the footprints of VEGETATION 1 and 2 overlap largely, as shown in Fig. 1. Moreover, the VEGETATION 2 footprint acquired during the subsequent orbit (110 min after the first one) also overlaps the VEGETATION 1 footprint. Therefore, a given point on the Earth surface may be observed up to three times a day (up to twice near the equator), with observation angles that may differ by more than 30°.

Fig. 3 shows the complementarity of VEGETATION 1 and 2 data acquired for a given day, where VEGETATION 1 data fill a gap between the observation angles provided by VEGETATION 2. Depending on the instrument, the errors on the surface reflectance due to unfiltered clouds or aerosols affect different parts of the observation angle range and induce different errors on the BRDF model fit. As a consequence, the errors on level 2 products will have a different impact on the composite reflectance for VEGETATION 1 or 2, and the errors on composite products have a quite good statistical independence, even if not perfect (of course, for a given day, if a pixel is cloudy for VEGETATION 1, it is fairly likely that it will also be cloudy for VEGETATION 2).

Therefore, it is possible to estimate the errors on composite reflectances (noted ρ) from the comparison of products coming from both satellites. As values provided in the composite images of each instrument are estimates of the mean value of the Nadir normalised reflectances during the compositing period, they should be identical in the absence of artefacts. The quality estimation is based

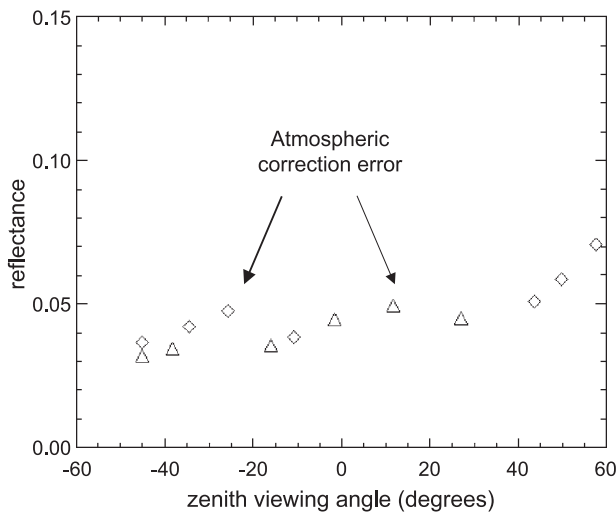


Fig. 3. Surface reflectances measured for B2 band by VEGETATION 1 (triangles) and VEGETATION 2 (diamonds), for an example pixel. Note that VEGETATION 1 data fill a gap in the observation angle range and also that atmospheric correction errors that occurred during the same day correspond to different viewing angles for the two instruments.

on what we call the Normalised Reflectance Difference (NRD) index:

$$\text{NRD} = 2 \frac{\rho_{\text{VGT2}} - \rho_{\text{VGT1}}}{\rho_{\text{VGT2}} + \rho_{\text{VGT1}}} \quad (5)$$

The bias between VEGETATION 1 and 2 syntheses is evaluated with the mean NRD over all the points that are valid for both VEGETATION images. Assuming that the errors affecting both data sets are statistically independent, the local (in time and space) standard deviation of the noise may be estimated by the standard deviation of NRD divided by the square root of 2. Since this criterion is an estimator of the standard deviation of the noise added to the time series of individual pixels, it will be known as “temporal criterion”.

3.2. Spatial criterion

Current temporally composited S10 images have a very obvious characteristic when compared with daily images of the same scene acquired under clear atmospheric conditions (Fig. 5). They have a different grain despite the fact that both types of images have the same resolution. The reason for this difference is that daily images can be considered as snapshots of the scene, in which all elements are simultaneously acquired, while temporally composited images are synthetic products in which information from different dates is mixed. The nature of landscape changes smoothly with the location and as a result, pixels tend to be more similar to their neighbours than to those located at larger distances. The particular form of such spatial dependency for the different elements of the scene is an important characteristic that becomes distorted by the temporal compositing.

Semivariograms are a very simple and commonly used statistical tool to analyse spatial dependency. These estimators, which have been applied to remotely sensed imagery on numerous occasions are plots of half of the average squared difference between m pairs of pixel values (I) against the distance (h) between the elements of the pairs (I and I_h):

$$\gamma(h) = \frac{1}{2m} \sum_1^m (I - I_h)^2 \quad (6)$$

Note that if n sites ($n=2m$) are used instead of m , the same plot is known as a variogram.

To evaluate the distortion of the spatial dependency that is introduced by the different compositing methods (the spatial criterion), we compare variograms of VEGETATION operational products and those of our newer composites in several windows. The spatial variability of reflectances observed in the variograms results from the interaction between the errors introduced by the compositing methods and the spatial heterogeneity of the scene itself. To try to extract the errors of the compositing process, we compare the variograms of the various composite products to those of a reference image: this image is a single day image of the same zone acquired under very clear atmospheric condi-

tions, within the compositing period. It is assumed that this reference is closest to a perfect composited product in terms of spatial quality. In order to make a fair comparison, we combine cloud masks of all products and calculate all variograms with their reunion. Only pixels that are valid (i.e. declared cloud free) for all the methods are considered.

4. Assessment of composite products quality

4.1. Data sets

Owing to operational reasons, a global coverage of the continents by both instruments is possible but quite costly. During the commissioning phase (May to December 2002), VEGETATION 2 was acquiring data over the whole Europe and Africa, whereas VEGETATION 1 was covering all continents. During the first 10 days of December 2002, both instruments observed the whole continental surfaces. For this study, global S10 and D10 syntheses were produced for each instrument from the global data set acquired by both instruments during the first decade of December 2002. S10 and D10 data were extracted and their performances evaluated over four geographical zones depicted in Fig. 4.

We were not able to produce global data sets of our improved products for operational reasons (data availability, computer power). For these new products, we used a data set that covers two of the geographical zones mentioned above: Central West (CW) Africa and South West (SW) Europe

Table 1

Summary of criteria for evaluating the performance of each algorithm

	CW Africa 12/02	SW Europe 09/02	Other zones 12/02
S10	Spatial–Temporal	Spatial	Temporal
D10	Spatial–Temporal	Spatial	Temporal
D10*, E10	Temporal	Temporal	
E15	Spatial–Temporal	Spatial–Temporal	
F15	Spatial	Spatial	

Europe, from July 2002 to March 2003. For the first decade of December 2002, S10 and D10 products are also available for both instruments, and it is thus possible to compare the new products to the official ones using the temporal criterion. Unfortunately, this comparison is not possible above SW Europe in December because of a heavy cloud cover. For this zone, we used a different period, September 2002, but since the official products of VGT2 are not available for this date, the comparison of new and official products is only feasible with the spatial criterion that only requires one instrument.

In order to simplify the description of results, the enhanced products have been named:

- D10*, product obtained with the D10 compositing method, but with the enhanced cloud screening
- E10 (Enhanced, 10 days) for the enhanced product obtained with only one sensor over 10-day period
- E15 (Enhanced, 15 days) same as E10, with 15-day periods (E30, 30 days)

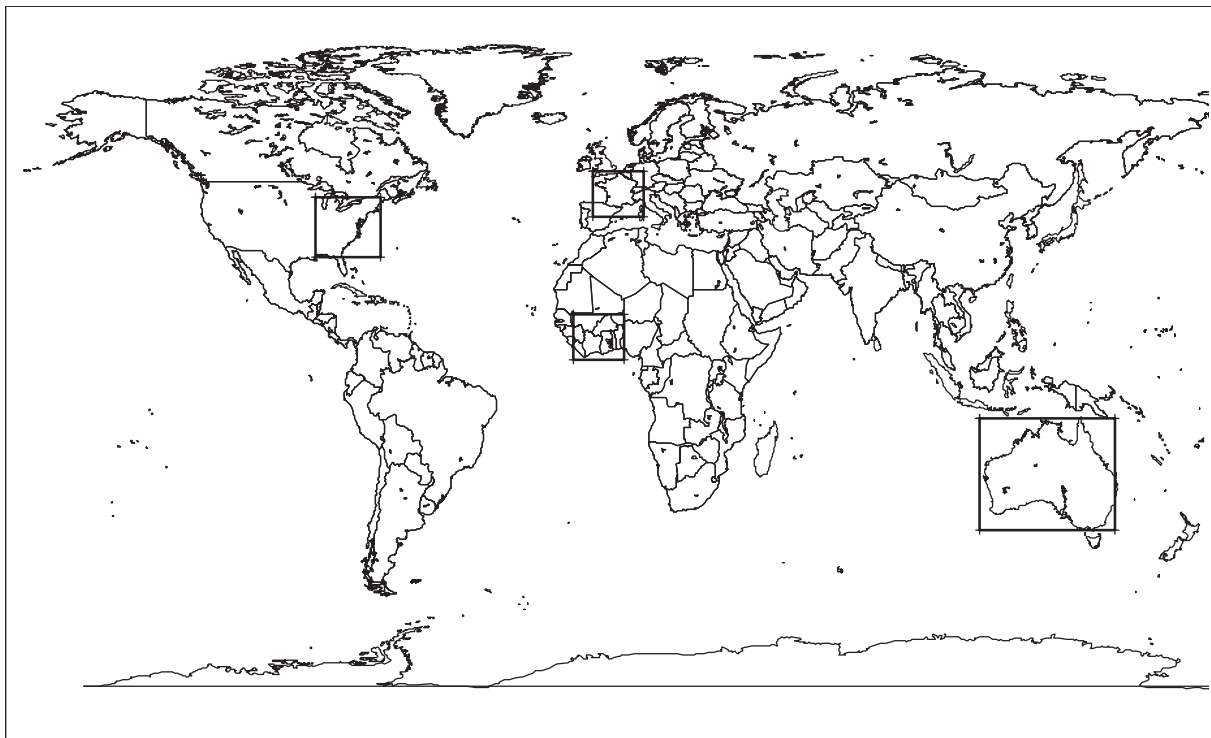


Fig. 4. Areas of study.

- F15 (Fusion, 15 days) for the enhanced products obtained with both VEGETATION sensors.

Table 1 summarises the data sets for which performances were evaluated.

4.2. Performance of operational and improved products

4.2.1. Visual inspection

Fig. 5 shows vignettes extracted from all S10, D10 E15 and D15 composite products for nine different zones: (rows

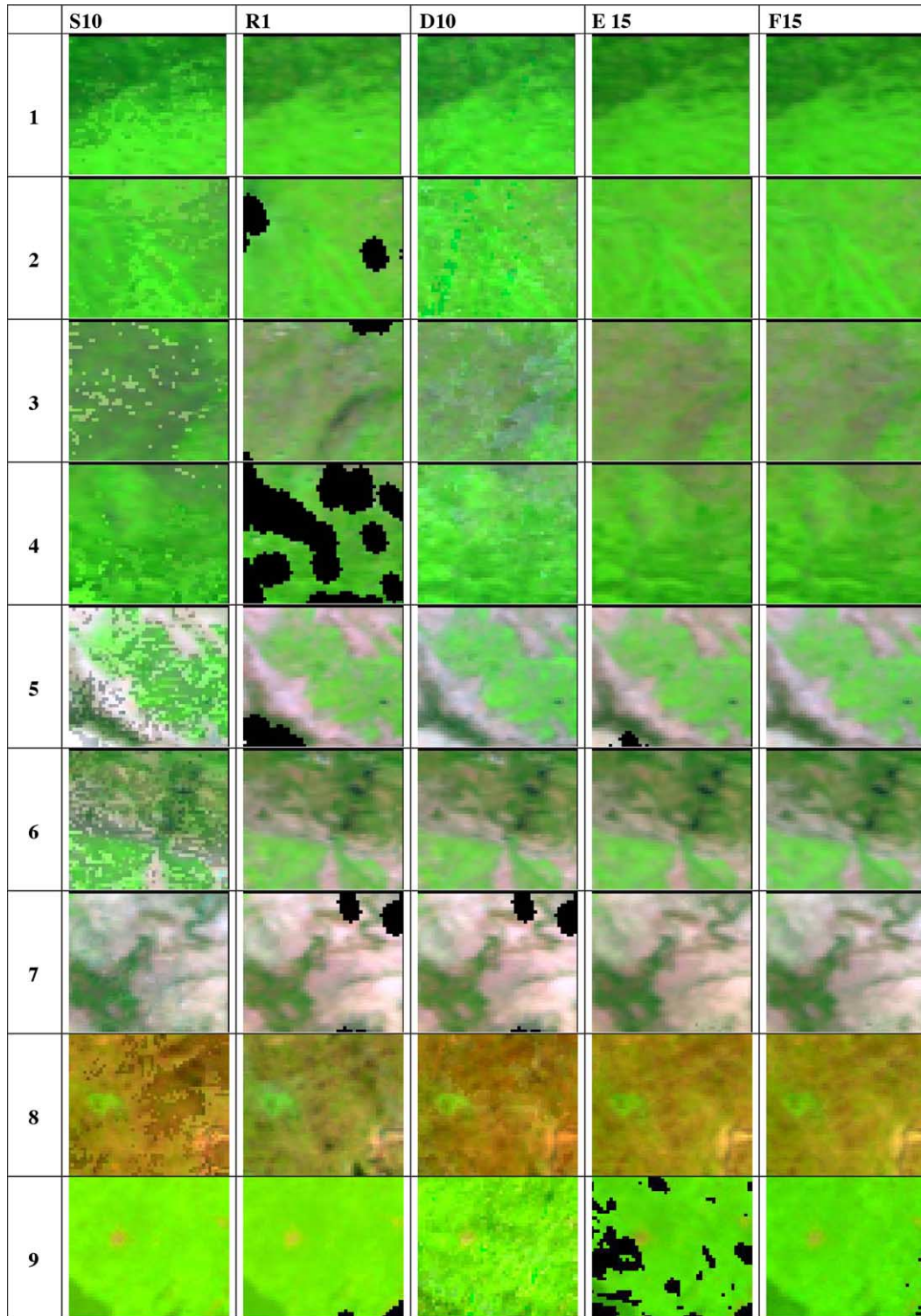


Fig. 5. Details of the different composite products for vignettes distributed in the SW Europe scene (rows 1–7) and in the CW Africa scene (rows 8 and 9). S10, currently available Maximum Value Composite (10-days period); R1 1-day image; D10, currently available Directional Composite (10-day period); E15, enhanced compositing method (15-day period) using one single instrument; F15, enhanced compositing method (15-day period) using both instruments.

1–7) are situated in the SW Europe scene and rows 8 and 9 are situated in the CW African scene. For nearly all the vignettes, S10 products look very noisy, and for a few of them, D10 is also quite noisy, whereas the enhanced algorithms seem to perform well when enough data (more than 3) are available to perform the BRDF model fit. Vignettes located in more arid zones (rows 5–8) show that D10, E15 and F15 are very similar under clear atmospheric conditions while S10 products are spatially corrupted. Vignettes located in more humid zones (rows 1–4, 9), prone to more hazy and cloudy conditions, show that D10 products approach the poor quality of S10 products under these circumstances, while E15 and F15 seem correct.

4.2.2. Temporal criterion results

4.2.2.1. Bias. Lack of continuity between successive instruments has been a source of serious trouble for Earth observers in the past. Many authors have had problems at processing time series derived from successive AVHRR instruments (Kaufmann et al., 2000). (Maisongrande et al., 1995; Malmström et al., 1997) show a 10% variation in the evaluation of Net Primary Production when AVHRR 9 is replaced by AVHRR 11. Biases may be caused by calibration errors, spectral band differences and/or angular differences related to different overpass times. In the case of VEGETATION, the overpass times are the same, and spectral bands differ by only a few nanometres. A special care has been given to the cross calibration of VEGETATION 1 and 2. This task was accomplished by using the desert sites method (Cabot et al., 2000), and the resulting uncertainty is believed to be better than 2% at the top of atmosphere. However, calibration biases are worsened after atmospheric correction, in particular for B0 and B2 bands.

Table 2 sums up the mean value of NRD over four geographical zones in December 2002. In order to make a fair comparison of the S10 and D10 products, statistics are computed only for those pixels that are simultaneously valid

for VEGETATION 1 and 2 S10 and D10 products. Biases between instruments are kept within $\pm 5\%$, except for B0 band, with some differences from one zone to the other, or from one compositing method to the other. The average biases for B2, B3 and MIR bands are lower than 2%. For the B0, the bias may be as high as 20% for the NE USA zone, and nearly 0 for CW Africa. Such a high variability for this band is due to the fact that the atmospheric contribution often represents more than 70% of the top of atmosphere signal, particularly in the Northern Hemisphere in winter. Moreover, a cloud that is detected for one instrument and not for the other may introduce a very large difference in the average reflectances.

According to our results, the continuity between VEGETATION 1 and 2 is correct (except for the B0 band) with a small bias between the two data sets. Therefore, it is possible to use time series of data that span before and after the change of the operational instrument for the temporal analysis of the dynamics of surface properties.

4.2.2.2. Standard deviation of NRD. We plot in Fig. 6 the estimations of the standard deviation of NRD (divided by square root of 2) measured on the official composites S10 and D10, for the four zones in December 2002. Two facts are worth mentioning: first, the standard deviations are very high in the B0 band, exceeding 30% in the Northern Hemisphere, but decreasing as wavelength increases.; second, for most cases, the D10 products performances are better than those of S10, except for B0 and B2 bands when cloud-free observations are very scarce. For B3 and MIR bands, the standard deviation is reduced by a factor greater than 2 for all zones, except on the cloudy SW Europe. As regards NDVI, S10 products still have better performances than D10 products except marginally in the dry Australian zone. The NDVI tends to cancel the errors added to the B2 and B3 bands when they are highly correlated (Tarpley et al., 1984). The D10 algorithm reduces the noise on reflectance but somewhat un-correlates the errors in the various bands, reducing the advantages of the NDVI formulation. To confirm this hypothesis, we have computed the correlation coefficient of the NRD of B2 and the NRD of B3 for the African scene: the correlation coefficient is 0.91 for the S10 algorithm, 0.83 for the D10 and is decreased to 0.64 for the E15.

The temporal criterion shows the very poor temporal performances of the MVC method to yield reflectances in the S10 products, whereas the benefit of taking directional effects into account in D10 products is evident. However, the D10 compositing method is quite sensitive to the quality of the cloud screening and produces more invalid pixels, it thus needs more cloud-free observations than the MVC method. But our main conclusion is that despite the enhancements, the standard deviation of the D10 products is still usually above 5% for B3 and MIR bands, and above 10% for B2 and B0 bands. There is therefore plenty of room for improvement.

Table 2

Bias between S10 and D10 products from instruments VEGETATION 1 and 2, computed for the first decade of December 2002, and for four geographic zones

D10	B0	B2	B3	MIR
SW Europe	-6.9	0.8	4.9	-0.4
Australia	-8.8	0.8	1.8	-2.0
NE USA	10.1	7.9	4.4	-0.3
CW Africa	-6.3	0.9	3.0	-2.0
S10	B0	B2	B3	MIR
SW Europe	-5.3	-4.1	-2.1	-6.5
Australia	-2.9	-1.2	-0.1	-3.4
NE USA	14.8	-4.0	0.9	-5.2
CW Africa	-4.4	2.8	3.5	-0.2

The bias estimated is the mean value of the Normalised Reflectance Difference (NRD), expressed in %.

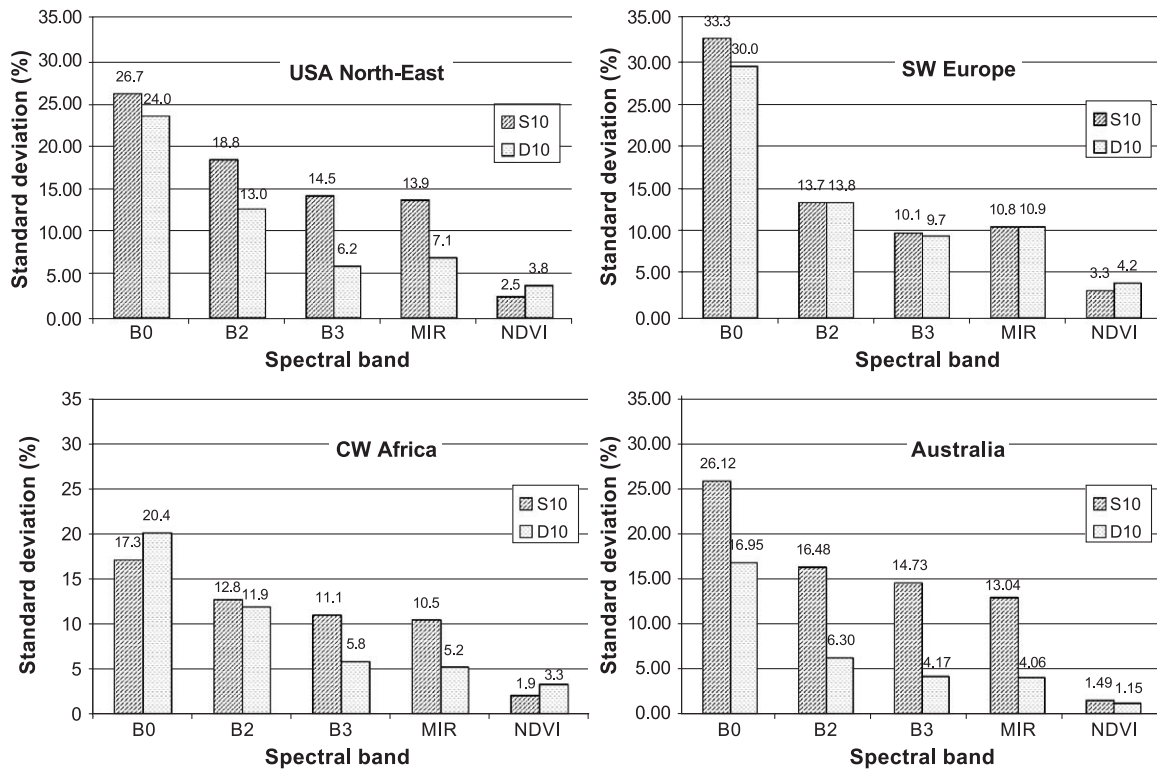


Fig. 6. Results of the temporal criterion (standard deviation of NRD divided by square root of two) over four different geographical zones, for S10 and D10 methods and for all spectral bands plus NDVI, in December 2002.

Fig. 7 shows the decrease of standard deviation of errors brought by the new improved products described in Section 2.2. The performances are evaluated with the temporal

criterion in the CW Africa region, in December 2002, considering only pixels that are valid for each method and each instrument. Performances are roughly enhanced by a

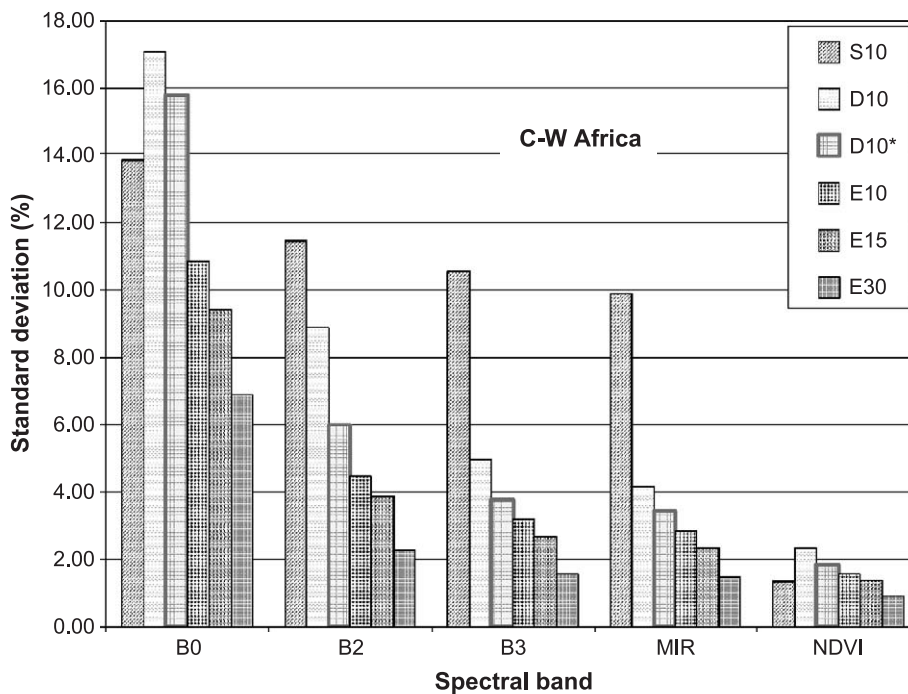


Fig. 7. Comparison of the standard deviation of NRD (temporal criterion) for all level 3 algorithms (current and enhanced), for the CW African zone, in December 2002.

factor 2 for the E15 compared to the D10. Standard deviations for the E15 product are now under 2% for B3 and MIR, under 5% for B2. After all the enhancements, the B0 band has still a standard deviation of about 10%, but this band is specifically sensitive to atmospheric effects and is not intended for vegetation monitoring. As far as NDVI is concerned, the performances are improved compared to D10 product, and the E15 product performs as well as the S10 product.

The same comparison has been done for a completely different landscape (SW Europe), in September 2002, but the results, presented in Fig. 8 are very close to those obtained for CW Africa, except standard deviations in B0 and B2 a bit higher for SW Europe than for CW Africa, probably because average reflectances in B0 and B2 are lower for the former zone than for the latter.

As shown in Fig. 9, the amount of valid pixels (for which a sufficient number of cloud free observations is available) decreases for the D10* and E10 since the cloud thresholds are stricter, but the E15 and E30 products increase again the coverage. The E30 product further enhances the performances, but of course considerably reduces data repetitivity. However, it is still possible to use 30-day composite products produced every 10 days with a shifting window. The E30 performance is also a good indicator of the performance of the F15 product, since it is roughly obtained with the same amount of data (the performances of F15 product cannot be estimated with the temporal criterion, since it would require four VEGETATION instruments to build two independent data sets). The degradation of performances brought by small differ-

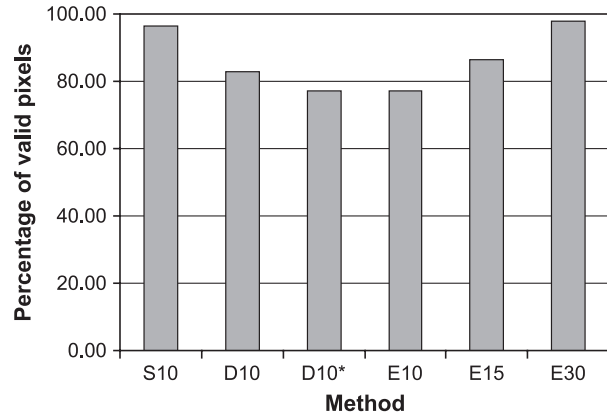


Fig. 9. Comparison of the percentage of valid pixels in the image obtained for each algorithm for the CW African zone in December 2002.

ences in the instruments should be compensated by the use of a shorter time period, with a lower evolution of the vegetation cover.

It is also interesting to study how performances evolve as a function of the number of observations used to inverse the BRDF model. This can be done with the temporal criterion, using only pixels for which the number of observations for VEGETATION 1 and 2 syntheses is above a minimum number. Fig. 10 shows the standard deviation of NRD as a function of the minimum number of observations: the standard deviation of errors is greatly reduced when the minimum number of observations is greater than 6. Of course, discarding pixels obtained with less than five observations would improve the global performances of products, but their coverage would be reduced. A better

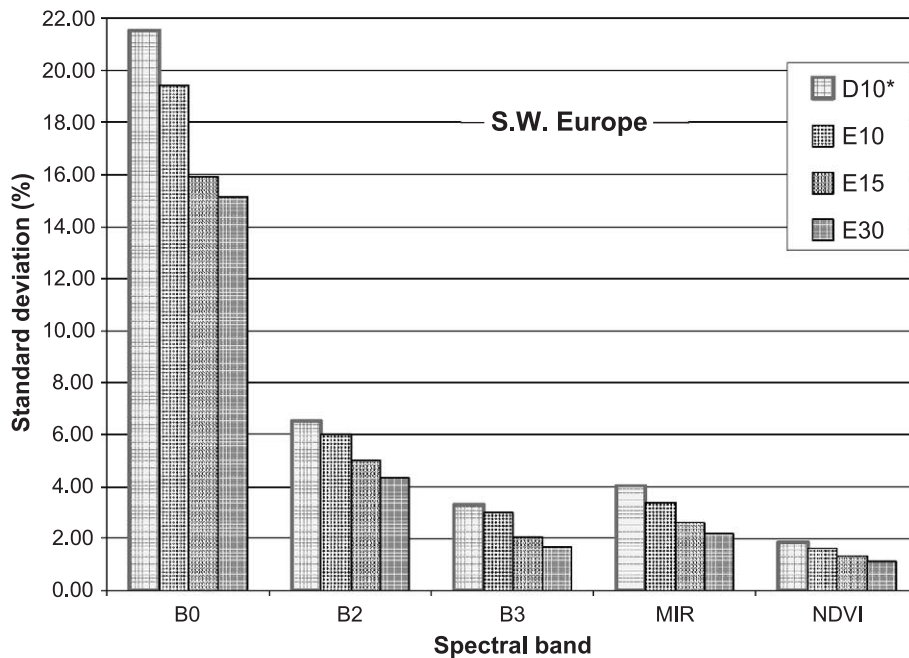


Fig. 8. Comparison of the standard deviation of NRD (Temporal criterion) for enhanced level 3 products, for the SW Europe zone in September 2002. Current composited products from both VEGETATION-1 and 2 were not available for this time period.

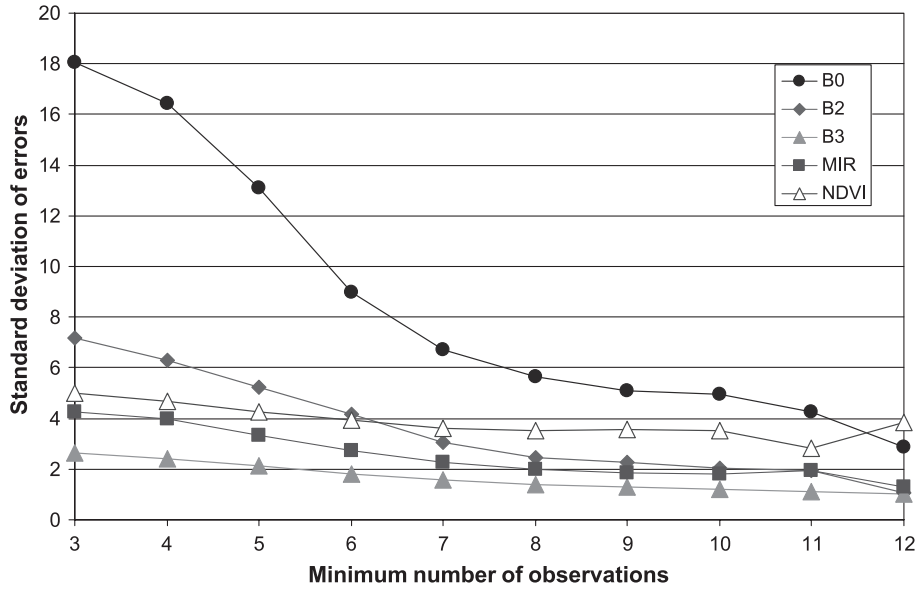


Fig. 10. Temporal criterion for the SW Europe zone for an E15 product. Standard deviation of NRD divided by square root of two, as a function of the minimum number of observations used in the BRDF model inversion.

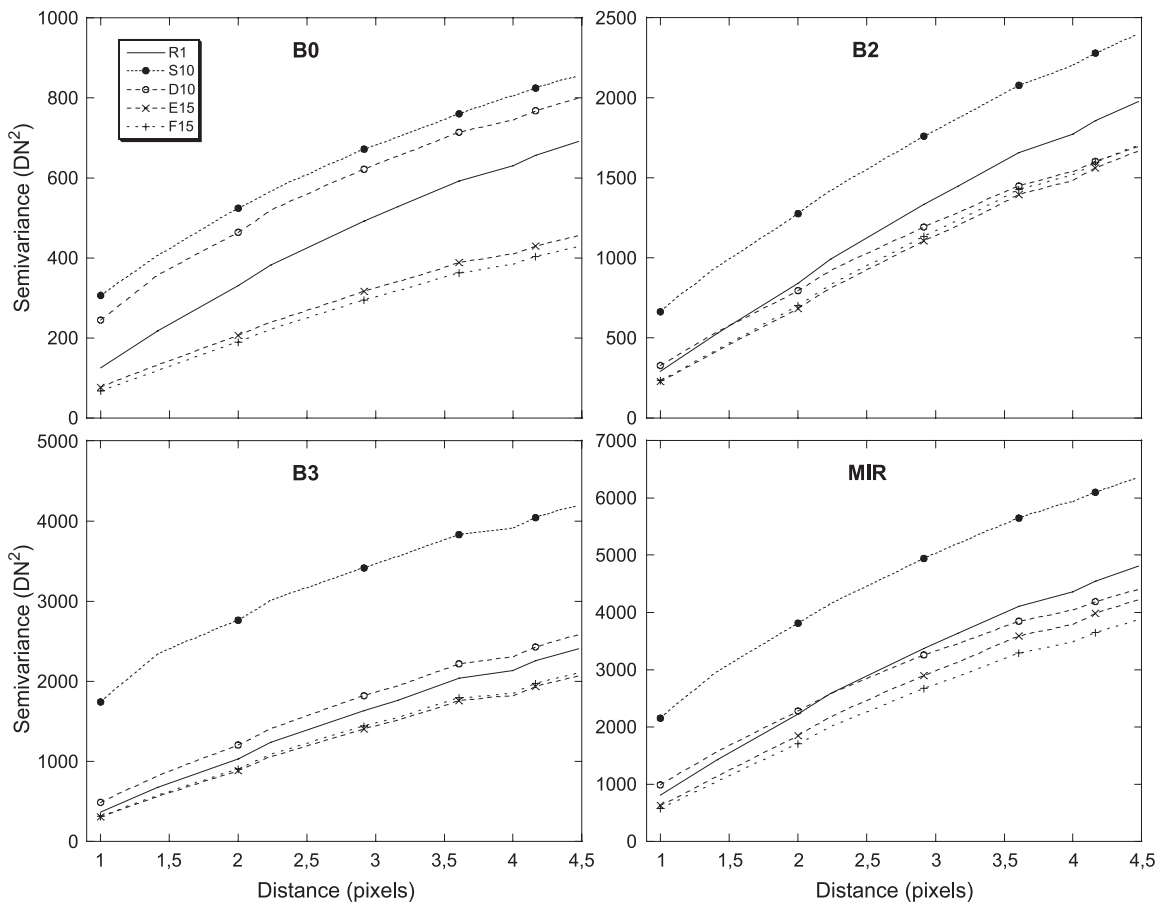


Fig. 11. Variograms of the four temporally composited products (S10, D10, E15 and F15) and the image acquired on 10-October-2002 (“R1”), for the region in SW Europe defined in Fig. 10. S10, currently available Maximum Value Composite (10-day period); R1, image acquired on 10-October-2002; D10, currently available Directional Composite (10-day period); E15, enhanced compositing method (15-day period) using one single instrument; F15, enhanced compositing method (15-days period) using both instruments. Note that the S10 product has a higher spatial variance at all scales and in all spectral bands. The higher the variogram, the more spatial noise.

approach is to use the number of observations used in the model inversion as a quality index for each pixel: many applications, in particular those involving data assimilation techniques, really need estimations of the uncertainty associated to each measurement.

4.2.3. Spatial criterion results

Fig. 11 presents the semivariograms of four of the composited products (S10, D10, E15 and F15) and of a clear-atmosphere single day image (R1) in all spectral bands of the SW Europe scene. Results from this figure can be summarized as follows:

- S10 products have a coarser grain in all bands, while the other composite products present a spatial quality that is similar to that of an image acquired in a clear-day (R1), except in the B0 band ;
- products E15 and F15 are very similar in terms of spatial characteristics and have a finer grain than D10.
- for B0, the band that is most affected by residual atmospheric noise, the variogram of D10 is almost as high as the variogram of the S10 product. For bands B3 and MIR, the variogram of D10 is only slightly grainier than the ones of E15 and F15.

Table 3

Percent of the surface with invalid values produced by each composite method

Product	S10	D10	E15	F15
SW Europe	0.0	2.9	6.4	2.8
W Africa	0.0	4.3	11.0	4.1
W Africa (500 km coastal fringe)	0.0	10.8	27.0	10.1

Invalid values occur when the number of cloud free observation is not sufficient.

Results from the CW Africa scene (Fig. 12) give further insight. While the ordering and shapes of R1, E15 and F15 are similar to the pattern found in the scene of SW Europe, the highest values of spatial variance are found in D10 images rather than in S10 images. A careful inspection of the CW African scene reveals that, in fact, there is no contradiction with results presented in the previous paragraph. As stated above, the quality of D10 products decreases as atmospheric conditions worsen at a higher rate than for remaining methods and the frequency of overcast conditions in CW Africa is much higher than in SW Europe for the considered time periods (Table 3).

For the NDVI (Fig. 13), the results obtained with the spatial criterion for the African scene are very close to those

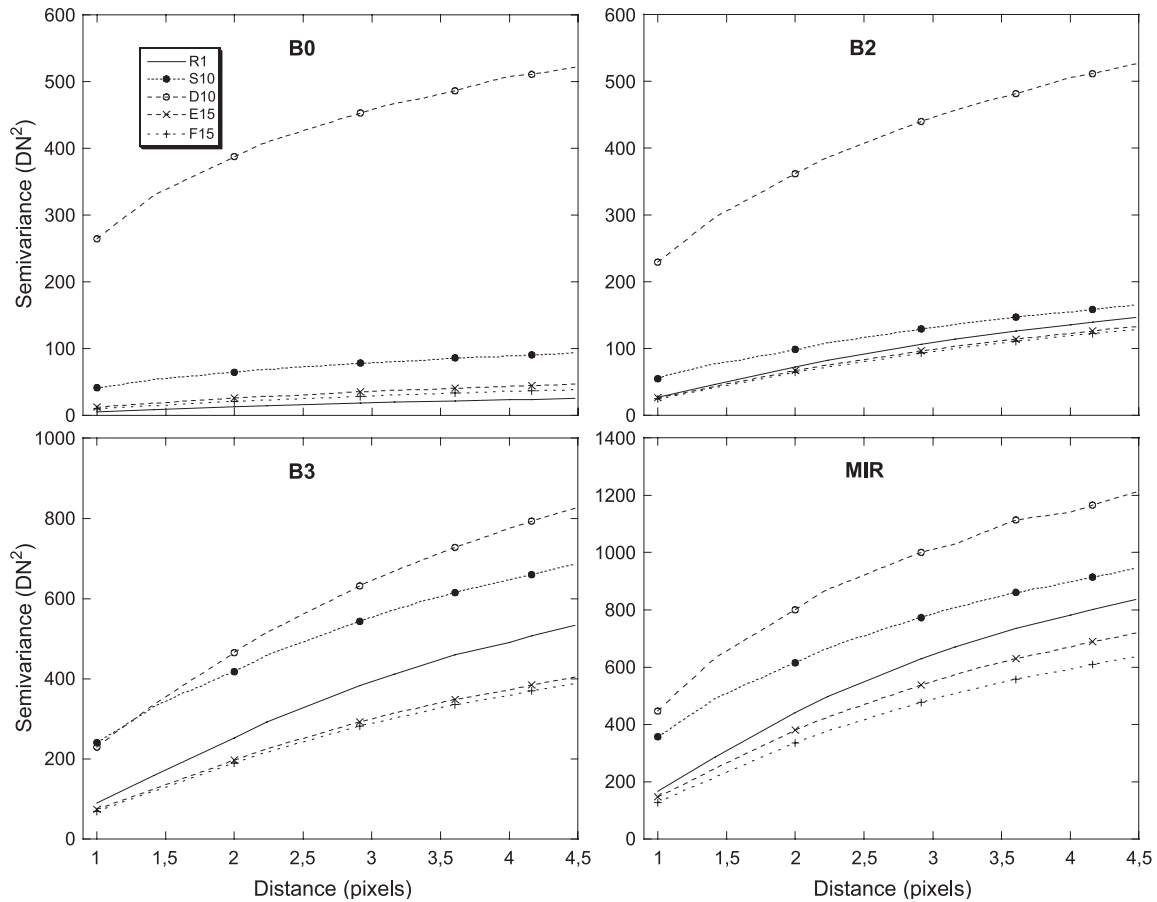


Fig. 12. Variograms of the four temporally composited products (S10, D10, E15 and F15) and the image acquired on 12-December-2002 (“R1”) for the CW Africa scene. See Fig. 10 for notation.

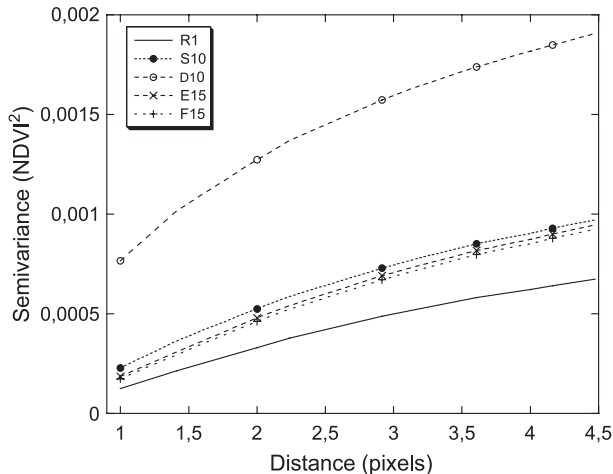


Fig. 13. Same plot as in Fig. 11, but for NDVI instead of reflectances.

of the temporal criterion, showing that the E and F products slightly improve the quality of NDVI compared to the S10, but that the D10 degrades it.

Considering results from variograms and from the visual inspection of the vignettes altogether, we can conclude that even if D10 products are a significant improvement over S10 products, this improvement decreases as atmospheric conditions worsen. The enhanced products E15 and F15 still perform well in the humid zones, but, conversely to the temporal criterion that shows a quality improvement between E15 and E30, the spatial criterion does not measure a quality improvement between E15 and F15: the remaining noise on those products must be low enough to be negligible compared to the spatial variability of the scene.

However, the essential fact is that the F15 method produces a much lower fraction of null values than E15 (see Table 3). The advantage of using both VEGETATION 1 and 2 instruments to compensate for the high probability of cloud cover is most evident here: while E15 produces a 27% of null values in the coastal fringe of the CW Africa scene, F15 reduces this proportion to 10%.

5. Conclusion

This paper addresses the problem of estimating the performances of global reflectance composited products over land surfaces. Two criteria were used to perform such estimations: a spatial criterion based on variograms to examine the spatial noise, and a new criterion based on the comparison of products issued from VEGETATION 1 and VEGETATION 2. The latter criterion has been termed the temporal criterion since it provides a way to estimate the standard deviation of the noise added to time series of reflectance data. These two criteria have been used to assess the quality of VEGETATION official products and to quantify the noise reduction brought by a few proposed

enhancements. Both criteria show consistent results even if the spatial criterion is less sensitive to small errors.

We show in this article that the S10 compositing method often produces images with very severe distortions of the spatial structure of the scene. For regions and periods of time in which good atmospheric conditions are not rare, D10 products are a good, simple and currently available alternative. Nevertheless, the quality of D10 products is very sensitive to the presence of undetected clouds and in extreme cases such as those in the equatorial zone, the improvement as compared to S10 data is reduced.

We propose improvements in the cloud screening and in the D10 compositing method, which are based on the use of a priori information on the directional model and on the noise statistics. As a result, the standard deviation of reflectance errors is divided by 2 when comparing E15 products with D10 products. The E15 product is more robust to the eventual occurrence of residual clouds in the time period to be composited. There are, however, regions and periods of time in which a significant fraction of the scene does not have a sufficient number of valid observations for E15 to produce an estimate. In these cases, the advantage of integrating images acquired from two instruments, VEGETATION 1 and VEGETATION 2, through the F15 method, becomes evident. Our work contributes to the idea that a significant improvement of global-scale products can be achieved through the operation of constellations of small satellites that acquire near-to-simultaneous images with equivalent instruments.

Acknowledgements

We are very grateful to the Centre de Traitement des Images Végétation at VITO in Belgium for having processed VEGETATION 1 and 2, and to the CNES EO/QI team for having ordered and downloaded all the VGT-P products. This work was partly funded by the European Union in the framework of CYCLOPES FP5 project.

References

- Cabot, F., Hagolle, O., & Henry, P. (2000). Relative and multitemporal calibration of AVHRR, SeaWiFS, and vegetation using POLDER characterization of desert sites. *International Geoscience and Remote Sensing Symposium*, 5, 2188–2190.
- De Fries, R. S., Hansen, M., Townshend, J. R. G., & Sohlberg, R. (1998). Global land cover classification at 8 km spatial resolution: The use of training data derived from Landsat imagery in decision tree classifiers. *International Journal of Remote Sensing*, 19, 3141–4168.
- Duchemin, B., Berthelot, B., Dedieu, G., Leroy, M., & Maisongrande, P. (2002). Normalisation of directional effects in 10-day global syntheses derived from VEGETATION/SPOT: II. Validation of an operational method on actual data sets. *Remote Sensing of Environment*, 81, 101–113.
- Duchemin, B., & Maisongrande, P. (2002). Normalisation of directional effects in 10-day global syntheses derived from VEGETATION/SPOT:

- I. Investigation of concepts based on simulation. *Remote Sensing of Environment*, 81, 90–100.
- Goward, S. N., Dye, D. G., Turner, S., & Yang, J. (1993). Objective assessment of the NOAA global vegetation index data product. *International Journal of Remote Sensing*, 14(18), 3365–3394.
- Holben, B. -N. (1986). Characteristics of maximum-value composite images from temporal AVHRR data. *International Journal of Remote Sensing*, 7, 1435–1445.
- Kempeneers, P., Lissens, G., Fierens, F., & Van Rensbergen, J. (2000). Detection of clouds and cloud-shadows for VEGETATION images. In G. Saint (Ed.), *Proceedings of VEGETATION 2000 symposium*, Lake Maggiore.
- Kaufmann, Y. J., Zhou, L., Knyazikhin, Y., Shabanov, N., Myneni, R. B., & Tucker, C. J. (2000). Effect of orbital drift and sensor changes on the time series of AVHRR vegetation index data. *IEEE Transactions on Geoscience and Remote Sensing*, 38(6), 2584–2597.
- Maisongrande, P., Duchemin, B., & Dedieu, G. (2004). VEGETATION/SPOT—an operational mission for the earth monitoring—presentation of new standard products. *International Journal of Remote Sensing*, 25, 9–14.
- Maisongrande, P., Ruimy, A., Dedieu, G., & Saugier, B. (1995). Monitoring seasonal and interannual variations of gross primary productivity, net primary productivity and net ecosystem productivity using a diagnostic model and remotely-sensed data. *Tellus. Series B, Chemical and Physical Meteorology*, 47(1–2), 178–190.
- Malmström, C. M., et al. (1997). Interannual variation in global-scale net primary production: Testing model estimates. *Global Biogeochemical Cycles*, 11(3), 367–392.
- Privette, J. L., Myneni, R. B., Knyazikhin, Y., Mukelabai, M., Roberts, G., Tian, Y., Wang, Y., & Leblanc, S. G. (2002). Early spatial and temporal validation of MODIS LAI product in the Southern Africa Kalahari. *Remote Sensing of Environment*, 83, 232–243.
- Rahman, H., & Dedieu, G. (1994). SMAC: A simplified method for the atmospheric correction of satellite measurements in the solar spectrum. *International Journal of Remote Sensing*, 15(1), 123–143.
- 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.
- Schaaf, C., 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. P., Hu, B., Liang, L., Privette, J. L., & Roy, D. (2002). First operational BRDF, Albedo and Nadir reflectance products from MODIS. *Remote Sensing of Environment*, 83(1–2), 134–148.
- Stowe, L. L., McClain, E. P., Carey, R., Pellegrino, P., Gutman, G., Davis, P., Long, C., & Hart, S. (1991). Global distribution of cloud cover derived from NOAA/AVHRR operational satellite data. *Advances in Space Research*, 3, 51–54.
- Tarpley, J. P., Schneider, S. R., & Money, R. L. (1984). Global vegetation indices from NOAA-7 meteorological satellite. *Journal of Climate and Applied Meteorology*, 23, 491–494.
- Weiss, M., et al. (2001). Mapping leaf area index measurements at different scales for the validation of large swath satellite sensors: first results of the VALERI project. *8th International symposium in physical measurements and remote sensing, Aussois (France)* (pp. 125–130).