Near real time vegetation monitoring at global scale

Aleixandre Verger[(1,2)*], Frédéric Baret[(2)] and Marie Weiss[(2)]

(1) CREAF, Cerdanyola del Vallès 08193, Catalonia, Spain

(2) INRA EMMAH UMR 1114, Avignon 84914, France

* Corresponding author: verger@creaf.uab.cat

Abstract

The NRT algorithm for near real time estimation of global LAI, FAPAR and FCOVER variables from VEGETATION (VGT) satellite data is here described. It consists of three steps (1) neural networks (one for each variable) to provide instantaneous estimates from daily VGT-P reflectances, (2) a multi-step filtering approach to eliminate data mainly affected by atmospheric effects and snow cover, and (3) Savitzky–Golay and climatology temporal smoothing and gap filling techniques to ensure consistency and continuity as well as short term projection of the product dynamics. Performances of NRT estimates were evaluated by comparison with other products over the 2005-2008 period: (1) the offline estimates from the application of the algorithm over historical time series (HIST), (2) the geoland2 version 1 products also issued from VGT (GEOV1/VGT) and (3) ground data. NRT rapidly converges closely to the HIST processing after 6 dekads with major improvement after 2 dekads. Successive reprocessing will therefore correct for some instabilities observed in the presence of noisy and missing data. The RMSE between NRT and HIST LAI is lower than 0.4 in all cases. It shows a rapid exponential decay with the number of observations in the composition window with convergence when 30 observations are available. NRT products are in good agreement with ground data (RMSE of 0.69 for LAI, 0.09 for FAPAR and 0.14 for FCOVER) and consistent with GEOV1/VGT products with a significant improvement in terms of continuity (only 1% of missing data) and smoothness, especially at high latitudes and Equatorial areas.

Index Terms— Near real time; continuity; consistency; biophysical variables; global scale; VEGETATION

1. Introduction

Near real time (NRT) estimation of global biophysical variables from moderate spatial resolution satellite sensors are of high interest in a range of application areas including numerical weather forecasting or monitoring of rapid land surface changes (e.g. droughts, hurricanes, forest fires, floods). These NRT variables are also required to support policies on environment and water management, agriculture and food security (White and Nemani 2006). The NRT concept refers to the minimum delay required to deliver the product. This delay corresponds to the time necessary to acquire and process the images for daily products corresponding to a single image acquisition. However, most products are derived after compositing the images acquired successively within a compositing window that extends generally symmetrically before and after the date of the product. The delay associated to the NRT product has thus to be extended by half the compositing window. To reduce this delay, the compositing window needs to be dissymmetric by reducing the part after the date of the product. This may be achieved by short term projection from the past values at the product date.

Although many studies point out the crucial need of NRT vegetation products (Fraser and Latifovic 2005; Ghulam et al. 2007; Running et al. 2004), only very few deal with near real time series (White and Nemani 2006; Xiao et al. 2011), and up to now, there is no delivery of NRT global vegetation product. Apart from this

NRT requirement, consistent and continuous long time series of global land surface variables are also essential to identify trends and anomalies and point out high risk areas.

Europe develops operational land monitoring services within the Copernicus initiative previously known as GMES (Global Monitoring of the Environment and Security). It will provide a series of bio-geophysical products describing the status and evolution of land surfaces at the global scale from long time series of remote sensing observations including near real time products. This project focuses on Leaf Area Index (LAI) and Fraction of Absorbed Photosynthetic Active Radiation (FAPAR) variables recognized as Essential Climate Variables by GCOS (GCOS 2011). It targets the vegetation cover fraction (FCOVER) variable as well. The Copernicus global land component benefits from the pre-operational geoland2 FP7 project (Lacaze et al. 2010), in which the GEOV1 products were developed.

GEOV1 products are derived from SPOT/VEGETATION data (hereafter called VGT) for the period 1999-present (Baret et al. 2013). GEOV1 has been demonstrated to outperform current existing products both in terms of accuracy and precision (Camacho et al. 2013). VGT time series archive was subsequently extended back in time using AVHRR/LTDR data for the period 1981-2000 (Verger et al. 2012). The second version of vegetation products (GEOV2) aims at being consistent with GEOV1/VGT in terms of accuracy while being produced in near real time. NRT estimates will be complemented with offline time series from 1999 to present which are expected to improve GEOV1/VGT in terms of continuity and consistency, especially at high latitudes and Equatorial areas. Similarly to the GEOV1/VGT previous version, global GEOV2/VGT products at 10-day time step and 1/112° spatial resolution will be freely delivered at Copernicus portal (land.copernicus.eu).

This paper focuses on the NRT aspect of forthcoming GEOV2/VGT products. The principles of the algorithm to generate near real time estimates of global biophysical variables from VGT data are first described. Then NRT estimates are evaluated based on the comparison with GEOV1/VGT and ground data. Particular attention is paid to the influence of noise and missing data on the estimation performances.

2. Algorithm outline

The NRT algorithm capitalizes on the efforts undertaken in the first version of GEOV1/VGT products (Baret et al. 2013) as well as in GEOV1/AVHRR (Verger et al. 2012) processing line. The main innovative relies in the near real time estimation achieved by performing short term projection of the product dynamics. It consists of using (1) neural networks to provide instantaneous variable estimates from VGT-P reflectances, (2) a multi-step filtering approach to eliminate data mainly affected by atmospheric effects and snow cover, and (3) temporal techniques to ensure consistency and continuity. The main steps of the NRT algorithm (Fig. 1) are summarized hereafter (further details are provided in Baret et al. (2012)).

[Fig. 1]

2.1. Instantaneous estimates from VGT-P data

The derivation of the instantaneous biophysical $P^{VGT-P}(d_1)$ 1-day products (LAI, FAPAR, FCOVER) is based on neural networks (NNT) trained using VGT-P reflectance data and fused MODIS and CYCLOPES LAI, FAPAR, FCOVER products similarly as in GEOV1/VGT (Baret et al. 2013). One NNT was calibrated for each of the three *P* variables (LAI, FAPAR, FCOVER) considered.

The inputs of the NNT are (1) the top of the atmosphere VGT-P reflectances in the four VGT bands (B0 (450 nm, Δλ=40 nm); B2 (645 nm, Δλ=70 nm); B3 (835 nm, Δλ=110 nm); SWIR (1165 nm, Δλ=170 nm)) , (2) acquisition geometry information: cosine of the view zenith, sun zenith and relative azimuth angles, and (3) atmospheric conditions: the ozone and water contents extracted from the VGT-P product. Correction for Rayleigh and aerosol effects is expected to be achieved implicitly through the direct training of the networks over the surface level biophysical products.

The output is the corresponding $P^{VGT-P}(d_1)$ instantaneous value of the biophysical variable (LAI, FAPAR or FCOVER). To be consistent with GEOV1/VGT algorithm, this output is computed similarly by fusing CYCLOPES version 3.1 (Baret et al. 2007) and MODIS collection 5 products (Yang et al. 2006). It consists in a weighted average of both products. The weighing, w, is designed to enhance the specific advantage of each product while limiting their deficiencies (Baret et al. 2013).

$$w = \frac{1}{0.982}(1 - \frac{1}{(1 + \exp(-2.LAI_{CYCV31} + 4))})$$

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

This smooth weighing function limits the brutal change observed for $LAI_{CYCv31} = 4$ in GEOV1 products (Fig. 2). Similarly to GEOV1, $w = 0.5$ when $LAI_{CYCv31} = 2$. Note that for FCOVER, no fusion was completed since CYCLOPES was the only existing product.

[Fig. 2]

To make the training process computationally tractable, it was achieved for the 2005-2008 period over the Benchmark Land Multisite Analysis and Intercomparison of Products (BELMANIP2) (Baret et al. 2006) sub-sample of sites which are representative of surface types and conditions over the Earth. To improve the NNT performances, the learning data base was filtered to remove input data contaminated by large atmospheric BRDF effects (very high sun and view zenith angles) and by clouds, snow or water-bodies (blue reflectance values B0>0.25 and points lying bellow the soil line in the B2, B3 and SWIR bands). The 84789 samples retained after the filtering process were used to train the NNT.

2.2. Multi-step outlier rejection process

Despite the fact that water vapor and ozone were explicitly introduced as inputs in the NNT, that MODIS and CYCLOPES products are derived from atmospherically corrected reflectances, and that the training data set was filtered, the resulting instantaneous product estimates are still contaminated by residual cloud and atmospheric effects (Fig. 3). The remaining outliers are filtered using a three-step process: (1) data are excluded if they do not belong to the definition domain of reflectances defined by the convex hull formed by the training dataset or if the product is out of the physical range of variation of the variable (0-7 for LAI, 0-0.94 for FAPAR and 0-1 for FCOVER), (2) a 3-iteration Savitzky-Golay filter (Verger et al. 2011) which leads to a smoothed curve fitted to the upper envelope of values in the time series (Fig. 3) and (3) a specific procedure for very noisy data (high latitude and equatorial forests) based on prior knowledge of the expected seasonality (further details are provided in Baret et al. (2012)).

[Fig. 3]

2.3. Temporal composition

A temporal composition was finally applied over the filtered daily $P_F^{VGT-P}(d_1)$ estimates to generate the biophysical $P(d_{10})$ products at 10-day step. It combines TSGF (Verger et al. 2011) and CACAO (Verger et al. 2013) techniques. TSGF (Temporal Smoothing Gap Filling) fits a second-degree polynomial over an asymmetric temporal window. The window is made of past and future semi-windows of adaptive length varying between 30 and 60 days. The length of the semi-window is determined by the availability of 6 valid observations the closest to the date of the dekad at which the product is estimated (Verger et al. 2011). If less than 6 observations exist in a 60 day semi-window, CACAO values evenly distributed every 10-days are used to fill gaps before the application of TSGF.

CACAO (Consistent adjustment of Climatology to Actual Observations) consists in fitting the climatology to actual observations for each growth season by scaling the magnitude and shifting the phenology. CACAO allows to better cope with missing and noise contaminated data as compared to standard methods as found in Verger et al. (2013) and Kandasamy et al. (2013). The climatology is computed as the inter-annual average of GEOV1/VGT time series over the 1999-2012 period. If it is available for a given pixel, the CACAO method allows filling all the gaps in the time series, even for missing data during long periods. Indeed, $P^{CACAO}(d_1)$ is closer to the $P_F^{VGT-P}(d_1)$ data than the original climatology $P^{CLIM}(d_1)$ (Fig. 4). However, the main limitation of CACAO reconstruction method is its inability to capture underlying atypical modes of seasonality including rapid natural and human induced disturbances in the time series that strongly differ from the average climatology (e.g. flood or fire events, changes in the land cover) (Verger et al. 2013). To prevent from such drawback, priority is given to TSGF smoothing since it is closer than CACAO to the actual $P_F^{VGT-P}(d_1)$ observations, while CACAO is only used to fill large gaps in the time series before the application of TSGF.

The combination of the TSGF local fitting and the projection capacity of CACAO allows to process in near real time (NRT) when only past observations are available. It allows also to process in offline mode the historical time series (HIST) when observations are available before and after the considered date. In the NRT case, CACAO is applied systematically to provide data every 10-days in the 60-day period after the NRT date. TSGF is then applied using this 60 days semi-window in the future, while the past semi-window spreads over 30 to 60 days length depending on the availability of 6 valid observations.

The HIST and NRT processing is illustrated for LAI estimation in Fig. 4. For the NRT situation, although fitted only with past data, CACAO (dashed blue line) estimates are generally closer to the instantaneous valid LAI estimates from VGT-P (black circles) than the original climatology (dotted blue line). However, it can differ from the historical processing (HIST continuous blue line) for periods with rapid LAI variations (from senescence to dormancy in site #188) or when significant noise is observed in the data (e.g. second growing season in site #233). In this case, the climatology fitting benefits from the availability of data before and after the date at which LAI is estimated. The TSGF application partially mitigates the problems of CACAO and the resulting NRT estimates (green line) shows a better local adaptation to the data than the original CACAO (dashed blue line). Note that NRT (green line) and HIST (black line) time series are very similar although some instabilities in NRT solution are found for very noisy data (e.g. January-April in site #233).

To avoid the instability in NRT estimates and improve the consistency with HIST time series, the products are updated each time a new dekad is available and processed (real time estimates). This results in the delivery of *n* successive updates of the *n* recent past values of the products in the convergence (CONV) period. To determine *n*, the difference between the CONV-n product and the HIST processing was computed for the year 2008 over the BELMANIP2 sites. It was found that after 6 dekads the CONV-6 converges closely towards the HIST processing (Fig. 5). Therefore, products will be updated during 6 dekads and then remain stable.

[Fig. 4]

[Fig. 5]

3. Evaluation of near real time estimates

The NRT algorithm was applied over time series of VGT-P data by considering only observations in the past period of the date being processed. The performances of NRT and CONV-n estimates are assessed by comparison both with (1) the HIST solution resulting from the application of the algorithm in offline mode with observations before and after the date being evaluated; (2) and GEOV1/VGT products as well as the few ground data available. This was achieved over the BELMANIP2 (Baret et al. 2006) and DIRECT sites (Garrigues et al. 2008) were the variables were measured at the ground level. The temporal profiles over a sample of sites are first discussed. Then, the performance of NRT and CONV-n estimates is assessed focusing on the influence of noise and missing data. Finally, the accuracy of NRT estimates is assessed as compared to ground data. For the sake of brevity, results focus on LAI which, among the three derived products (LAI, FAPAR and FCOVER), is the most used by the scientific community and the most sensitive to uncertainties in the satellite data.

3.1. Temporal profiles for a sample of sites

Few BELMANIP2 and DIRECT sites showing typical features have been selected to illustrate the performance of NRT estimates as compared to HIST processing, GEOV1 product and ground based measurements (Garrigues et al. 2008).

For regular sites having enough high quality data (Fig. 6), NRT and HIST solutions are very close and show a good agreement with GEOV1 product and ground measurements. Reasonable performances and no significant differences were found between the NRT estimates, the intermediate solutions in the convergence period (CONV-3) and the consolidated one (CONV-6). The NRT estimates show more instability due to the non-availability of actual observations in the semi compositing window after the date being processed (climatology values are used in this case).

For sites near the equator, having a significant fraction of missing data and noise due to persistent clouds (Fig. 7), NRT still provides reliable solutions and clearly outperforms GEOV1 in terms of consistency and continuity. The background information provided by the climatology fitting allows to efficiently fill the gaps: NRT estimates show less than 1% of missing data over the BELMANIP2 sites for the 2003-2010 period as compared to the 20% (up to 40% for needleleaf and evergreen broadleaf forests) of gaps in GEOV1 product (Verger et al.

2014). Further, the outlier rejection applied for very noisy situations (e.g. Tapajos site) allows eliminating contaminated data while keeping the expected level of LAI as indicated by the good agreement with ground data.

For sites located at very high latitudes (Lat>50º) (Fig. 8), GEOV1 shows some artifacts and anomalous seasonality in winter time (e.g. unexpected increase of LAI in October-November for Tundra and site #93). These problems are probably due to the instabilities in the Bidirectional Reflectance Distribution Function (BRDF) correction in extreme illumination conditions as well as possible residual snow pixel contamination (Baret et al. 2007). These artifacts were corrected in NRT estimates because (1) the sun angle was explicitly considered in the NNT, (2) outliers were rejected (e.g. site #93) and (3) climatology background information regularizes the estimation leading to smooth and continuous temporal profiles. Some underestimation problems also occurred in GEOV1 for these high latitude sites in the summer period partly due to the significant amount of noise in the data. Although mostly corrected in the HIST and the consolidated CONV-6 products, they can still be observed for NRT and CONV-3 (e.g. site #107, #93 and #418). This is further investigated by evaluating the impact of noise and missing data on NRT product in the next section.

[Fig. 6]

[Fig. 7]

[Fig. 8]

3.2. Sensitivity analysis to the number of observations and data noise

To better assess the performance of NRT estimates and their expected convergence towards HIST, the root mean square error (RMSE) between HIST and CONV-n for *n* varying from *n*=0 (initial solution corresponding to the NRT case with no observations after the date of estimation) to *n*=6 (consolidated product) is investigated as a function of the noise in the data and the number of available valid observations before the date of the estimate in the compositing window. The RMSE between the daily P_F^{VGT-P} observations and the HIST values is used as an estimate of the noise.

The RMSE between CONV-n and HIST LAI linearly increases with the amount of noise in the data (Fig. 9a). The RMSE slope of CONV-n versus RMSE of P_F^{VGT-P} relationship (both RMSE computed as compared to HIST) is higher for the lower *n* order of CONV-n solutions: the initial NRT solution (*n*=0) is the most affected by noisy data conversely to the consolidated solution (CONV-6). A rapid convergence is observed for the intermediate solutions which show a similar pattern after 2 dekads (CONV-2). In all the cases including the initial NRT (CONV-0), the slope of RMSE as a function of the noise is lower than 1 which indicates an improvement in the performance of NRT estimates as compared to the original instantaneous P_F^{VGT-P}.

Regarding the relationship with the number of available observations (Fig. 9b), the RMSE between CONV-n and HIST shows a rapid exponential decay. It is almost zero for 30 observations, i.e. the maximum number of observations before the date of the estimate corresponding to the length of the half compositing window. The RMSE values when no data is acquired before the date of the estimation (Fig. 9b) are very similar to the RMSE values for a 0.5 noise level (Fig. 9a). In all cases, including the situations when few observations are available and/or data are associated to a high level of noise, the discrepancies between NRT and HIST are reasonably low (RMSE<0.4 for LAI).

The number of available observations and the RMSE values of P_F^{VGT-P}appear to be pertinent indicators of the quality of the NRT, CONV-n and HIST products and may help the user exploiting the time series.

[Fig. 9]

The spatio-temporal distribution of the average number of valid P_F^{VGT-P} observations before the date of estimation (Fig. 10a) shows obvious patterns, with fewer valid observations around the equator and in winter for the higher northern and southern latitudes. The distribution of RMSE between NRT (Fig. 10b), CONV-3 (Fig. 10c), CONV-6 (Fig. 10d) and HIST show consistent spatio-temporal patterns with the number of valid observations: higher RMSE values are observed at locations and periods corresponding to the lower number of available data. It must be noticed that due attention to the phenology and the associated expected LAI value is required for a correct interpretation of the RMSE distributions. For example, the observed difference between NRT and HIST for high latitude (lat>50) (Fig. 10b) are lower for the winter time than in the growing season consistently with the expected lower LAI values in winter (Fig. 8). The comparison of NRT and CONV-3 distributions of RMSE shows that a clear improvement of CONV-3 performances (compare Fig. 10b and Fig. 10c). Marginal differences exist between RMSE distributions for CONV-3 (Fig. 10c) and CONV-6 (Fig. 10d). These results are consistent with the previous findings (Fig. 9b) indicating a rapid convergence of the solution after 2 dekads.

[Fig. 10]

3.3. Accuracy assessment

Both NRT and HIST estimates show a relatively good agreement with the available ground measurements of LAI as observed over different dates along the phenological cycle of sites displayed in Fig. 6, Fig. 7 and Fig. 8. More quantitative assessment was achieved using the 19 available ground-based measurements acquired over 15 different 3 km x 3 km sites in the 2003-2007 period and compiled by Garrigues et al. (2008). In addition to LAI, validation was achieved for FAPAR and FCOVER variables. For comparison purposes, GEOV1/VGT products were also validated over the same ground dataset. Each product was interpolated at the date of the ground measurements if two valid dekadal data exist within a maximum period of ±30 days. The comparison of NRT with the ground-based observations of LAI, FAPAR and FCOVER variables shows respectively an overall RMSE of 0.69, 0.09, 0.14 (Table 1, Fig. 11). Similar performances are found for HIST products over the same ground dataset (Table 1). NRT and HIST slightly outperform GEOV1 (lower RMSE, higher correlation and slopes and offset of the linear regression respectively closer to 1 and 0). However, this validation is limited by the low number of available ground-based measurements that were mostly achieved in non-problematic conditions close to the maximum peak of vegetation. Further confrontation with ground based data is required, particularly over sites located at equatorial regions (Fig. 7) or very high latitude (Fig. 8) where higher noise and occurrence of missing data is expected in the satellite surface reflectance data.

[Fig. 11]

[Table 1]

4. Conclusions

This paper presents the NRT algorithm to derive dekadal biophysical products both for the near real time conditions, as well as for the processing of the historical archive of VGT data at global scale. It capitalizes on the efforts undertaken in the first version of geoland2 products (GEOV1/VGT and GEOV1/AVHRR) through a 3-step procedure including (1) neural networks to provide instantaneous estimates from VGT-P reflectances, (2) a multi-step filtering approach to eliminate data mainly affected by atmospheric effects and snow cover, and (3) temporal techniques (TSGF Savitzky-Golay adaptive local filter and CACAO climatology fitting) to ensure consistency and continuity as well as short term projection of the product dynamics.

The NRT method was applied over actual satellite daily VGT-P observations for the 2005-2008 period over the BELMANIP2 and DIRECT ensemble of sites. Performances were evaluated by comparison of NRT estimates (i.e. when no available observation is available after the date of estimation) with historical (HIST) processing resulting from the application of the algorithm over the archive of VGT data (i.e. using observations before and after the considered date), GEOV1/VGT products and available ground data. Results show the potential of the NRT algorithm for continuous, consistent and near real time estimation of global biophysical products from satellite observations.

NRT estimates show reasonable performances to reproduce the expected seasonality over a variety of vegetation conditions although some instabilities in the solution were identified in presence of noise and gaps in the data. To increase the robustness of the solution and improve the consistency with HIST time series, the products are updated each time a new dekad is processed until reaching convergence. NRT rapidly converges closely towards the HIST processing after 6 dekads (consolidated solution) with major improvements in the patterns of intermediate solutions after 2 dekads. The performances of the algorithm are closely linked to the number of observations available before the date of estimation and the noise in the data. Both of them are provided along with the product values as quality indicators.

First validation results indicate that NRT products have high consistency with GEOV1/VGT products with a significant improvement in terms of continuity (less than 1% of missing data over the BELMANIP2 sites as compared to the 20% of gaps in GEOV1 products) and temporal consistency (smoother products less affected by noise in the data), especially at high latitudes and Equatorial areas. Indeed, the use of the climatology allows filling gaps and improves robustness in time series. Note however that for long periods of missing data, NRT estimation is very challenging and solution may be affected by the inability of the climatology to capture underlying atypical modes of seasonality. The user is advised to use quality metrics and quality flags associated with the products. Finally, good agreement with the available ground measurements was observed (root mean square error of 0.69 for LAI, 0.09 for FAPAR and 0.14 for FCOVER) although this validation needs to be extended to other sites and years. Further, confrontations with other existing temporal methods for NRT estimation (e.g. Jiang et al. 2010) and assimilation techniques including Kalman Filters (e.g. Xiao et al. 2011) should also be conducted.

The proposed method will be implemented in the Copernicus GIO land programme and GEOV2/VGT global biophysical products will be freely delivered through the Copernicus portal (land.copernicus.eu) at 1/112° spatial resolution every 10 days in NRT as well as in offline mode (time series from 1999 to present). NRT products are expected to support a wide range of applications and policies on the environment requiring information of the status and evolution of land surface at global scale. Since GEOV2/VGT products are based on the same principles as for the GEOV1/AVHRR (1981-2000) products (Verger et al. 2012), the combination of both datasets is expected to provide continuous and consistent long time series of global LAI, FAPAR and FCOVER variables for the last three decades. These long-term data records are expected to contribute to global climate monitoring and earth science modeling applications.

Acknowledgement

This research was partially supported by the Copernicus European Earth monitoring program, FP7 geoland2 (218795) and GIOBIO (32-566) projects. Aleixandre Verger is the recipient of a *Juan de la Cierva* postdoctoral fellowship from the Spanish Ministry of Science and Innovation.

References

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

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

Baret, F., Weiss, M., Lacaze, R., Camacho, F., Makhmara, H., Pacholcyzk, P., & Smets, B. (2013). GEOV1: LAI, FAPAR Essential Climate Variables and FCOVER global time series capitalizing over existing products. Part1: Principles of development and production. *Remote Sensing of Environment, 137*, 299-309

Baret , F., Weiss, M., Verger, A., & Kandasamy, S. (2012). BioPar Methods Compendium - LAI, FAPAR and FCOVER from VEGETATION P products series. In (p. 43). Avignon: INRA-EMMAH

Camacho, F., Cernicharo, J., Lacaze, R., Baret, F., & Weiss, M. (2013). GEOV1: LAI, FAPAR Essential Climate Variables and FCOVER global time series capitalizing over existing products. Part 2: Validation and intercomparison with reference products. *Remote Sensing of Environment, 137*, 310-329

Defourny, P., Bicheron, P., Brockmann, C., Bontemps, S., Van Bogaert, E., Vancutsem, C., Pekel, J.F., Huc, M., Henry, C., Ranera, F., Achard, F., di Gregorio, A., Herold, M., Leroy, M., & Arino, O. (2009). The first 300 m global land cover map for 2005 using ENVISAT MERIS time series: a product of the GlobCover system,. In, *Proceedings of the 33rd International Symposium on Remote Sensing of Environment*. Stresa (Italy)

Fraser, R.H., & Latifovic, R. (2005). Mapping insect-induced tree defoliation and mortality using coarse spatial resolution satellite imagery. *International Journal of Remote Sensing, 26*, 193-200

Garrigues, S., Lacaze, R., Baret, F., Morisette, J., Weiss, M., Nickeson, J., 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*

GCOS (2011). -Global Climate Observing System - Systematic Observation Requirements for Satellite-Based Products for Climate - 2011 Update, Supplemental Details to the Satellite Based Component of the Implementation Plan for the Global Observing System for Climate in Support of the UNFCCC (2010 Update), . In (p. 138). Geneva, Switzerland: World Meteorological Organization

Ghulam, A., Qin, Q., Teyip, T., & Li, Z.-L. (2007). Modified perpendicular drought index (MPDI): a real-time drought monitoring method. *ISPRS Journal of Photogrammetry and Remote Sensing, 62*, 150-164

Jiang, B., Liang, S., Wang, J., & Xiao, Z. (2010). Modeling MODIS LAI time series using three statistical methods. *Remote Sensing of Environment, 114*, 1432-1444

Kandasamy, S., Baret, F., Verger, A., Neveux, P., & Weiss, M. (2013). A comparison of methods for smoothing and gap filling time series of remote sensing observations. Application to MODIS LAI products. *Biogeosciences, 9*, 17053-17097

Lacaze, R., Balsamo, G., Baret, F., Bradley, A., Calvet, J.C., Camacho, F., D'Andrimont, R., Freitash, S.C., Makhmara, H., Naeimij, V., Pacholczyk, P., Poilvé, H., Smets, B., Tansey, K., I.F., T., Wagner, W., & Weiss, M. (2010). GEOLAND2 - Towards an operational GMES land monitoring core service : first results of the biogeophysical parameter core mapping service. In W. Wagner, & B. Székely (Eds.), *ISPRS TC VII Symposium – 100 Years ISPRS,* (pp. 354-358). Vienna (Austria): IAPRS

Running, S.W., Nemani, R.R., Heinsch, F.A., Zhao, M., Reeves, M., & Hashimoto, H. (2004). A continuous satellite-derived measure of global terrestrial primary production. *Bioscience, 54*, 547-560

Verger, A., Baret , F., & Weiss, M. (2011). A multisensor fusion approach to improve LAI time series. *Remote Sensing of Enviroment, 115*, 2460-2470

Verger, A., Baret , F., Weiss, M., Kandasamy, S., & Vermote, E. (2013). The CACAO method for smoothing, gap filling and characterizing seasonal anomalies in satellite time series. *IEEE transactions on Geoscience and Remote Sensing, 51*, 1963-1972

Verger, A., Baret , F., Weiss, M., Lacaze, R., Makhmara, H., & Vermote, E. (2012). Long term consistent global GEOV1 AVHRR biophysical products. In, *Proceedings of 1st EARSeL Workshop on Temporal Analysis of Satellite Images* (pp. 28-33)

Verger, A., Baret, F., Weiss, M., Smets, B., Lacaze, R., & Camacho, F. (2014). Near real time estimation of biophysical variables within Copernicus global land service. In, *Global vegetation monitoring and modeling (available at https://colloque.inra.fr/gv2m/Poster-Sessions/Poster-S7)*. Avignon (France)

White, M.A., & Nemani, R.R. (2006). Real-time monitoring and short-term forecasting of land surface phenology. *Remote Sensing of Environment, 104*, 43-49

Xiao, Z., Liang, S., Wang, J., Jiang, B., & Li, X. (2011). Real-time retrieval of Leaf Area Index from MODIS time series data. *Remote Sensing of Environment, 115*, 97-106

Yang, W., Shabanov, N.V., Huang, D., Wang, W., Dickinson, R.E., Nemani, R.R., Knyazikhin, Y., & Myneni, R.B. (2006). Analysis of leaf area index products from combination of MODIS Terra and Aqua data. *Remote Sensing of Environment, 104*, 297-312

Table 1. Statistics of the comparison of NRT, HIST and GEOV1/VGT with ground measurements for LAI, FAPAR and FCOVER variables over the DIRECT sites for the 2003-2007 years: number of sites, number of samples (sites x dates), percentage of samples which meet GCOS requirements in terms of accuracy (max(20%, 0.5) for LAI, max(10%, 0.05) for FAPAR (and FCOVER) (GCOS 2011)), root mean square error (RMSE), correlation coefficient (R), slope and offset of the linear regression.

		Nb. site	Nb. sample	%OK GCOS	RMSE	R	slope	offset
LAI	NRT	15	19	89	0.69	0.94	0.89	-0.01
	HIST	15	19	80	0.78	0.92	0.86	0.03
	GEOV1	15	19	75	0.76	0.92	0.89	0.01
FAPAR	NRT	13	13	62	0.09	0.97	0.97	0.06
	HIST	13	13	54	0.09	0.95	0.90	0.07
	GEOV1	13	13	38	0.12	0.91	0.91	0.07
FCOVER	NRT	17	24	33	0.14	0.88	1.27	-0.10
	HIST	17	24	29	0.14	0.87	1.28	-0.10
	GEOV1	17	24	33	0.15	0.86	1.34	-0.14

Fig. 1. Flow chart showing the general principles of the NRT algorithm used to derive biophysical products from VGT-P reflectances. The three main steps identified in the text are highlighted in grey. The white ellipses correspond to the applied methods. The original, final and intermediate products with specification of their temporal sampling (d_1 for daily and d_{10} for dekadal) are indicated in rectangular white boxes. For step 1, fused MODIS +CYCLOPES data are used only in the NNT training process.

Fig. 2. The weighing function used for the fusion between CYCLOPES and MODIS LAI and FAPAR products. The dashed line corresponds to the weight used for generating GEOV1 products. The dotted line corresponds to $w = 0.5$.

Fig. 3. Illustration of the 3-iterations of TSGF filtering (continuous line) to eliminate contaminated data (filled circles). Empty circles correspond to valid data. The number of the BELMANIP2 site, the biome class (following the GLOBCOVER map, (Defourny et al. 2009)) and the latitude and longitude are indicated.

Fig. 4. Illustration of the HIST and NRT LAI estimation over two BELMANIP2 sites. The number of each site, the biome class (following GLOBCOVER map, (Defourny et al. 2009), the latitude and longitude are indicated.

Fig. 5. Evaluation of the differences between CONV-n and HIST processing over the BELMANIP2 sites for year 2008 as a function of the number of dekads after the date being processed, *n*. Zero dekads ($\boldsymbol{n} = \mathbf{0}$) corresponds to the NRT case with data available only for the past. The several gray values correspond to 75% (dark gray), 90% (medium gray) and 95% (light gray) of the population, and the dots to 5% percentile of residual outliers. The bold back solid line (close to 0:0 line) corresponds to the median value of the differences.

Fig. 6. Temporal profiles of the NRT, CONV-n for n=3,6 dekads after the date being processed, HIST and GEOV1 LAI products for regular sites. The valid and filtered VGT-P LAI estimates as well as the available ground measurements are shown. The title of each plot indicates the DIRECT site name or BELMANIP2 site number, the GLOBCOVER (Defourny et al. 2009) biome class, the latitude and longitude in degrees.

Fig. 7. Same as Fig. 6 but for sites near the Equator.

Fig. 8. Same as Fig. 6 but for very high latitude sites.

Fig. 9. RMSE between CONV-n and HIST estimates as a function of (a) noise in the data computed as the RMSE between the daily P_F^{VGT-P} observations and the HIST product and (b) number of valid observations in the left semi-period of composition (i.e. before the date of estimation). n=0, 1…6 corresponds to the dekad number after the date of estimation results shown for the LAI over the BELMANIP2 sites for the year 2008. Zero dekads ($n = 0$) corresponds to the NRT case.

Fig. 10. Spatio-temporal distribution as a function of the latitude (10° steps) and the date of acquisition (monthly step) of (a) the average number of valid daily VGT-P observations before the date of the NRT estimates, (b) the RMSE between NRT and HIST, CONV-n and HIST for (c) n=3 and (d) n=6 where n is the number of dekads after the date of estimation. Evaluation for LAI estimates over the 445 BELMANIP2 sites for the year 2008.

Fig. 11. Comparison of NRT estimates with scaled ground measurements for LAI, FAPAR and FCOVER. The different symbols correspond to the five biome classes as derived from the GLOBCOVER (Defourny et al. 2009) global landcover: Shrubs/Savana/Bare soil (SSB), Crops and Grassland (CG), Deciduous Broadleaf Forests (DBF), Needleleaf Forest (NF), and Evergreen Broadleaf Forest (EBF). The dotted line corresponds to the 1:1 line. The solid lines represent the GCOS accuracy criteria: max(20%, 0.5) for LAI, max(10%, 0.05) for FAPAR (and FCOVER) (GCOS 2011). The statistics of the comparison are provided in Table 1.

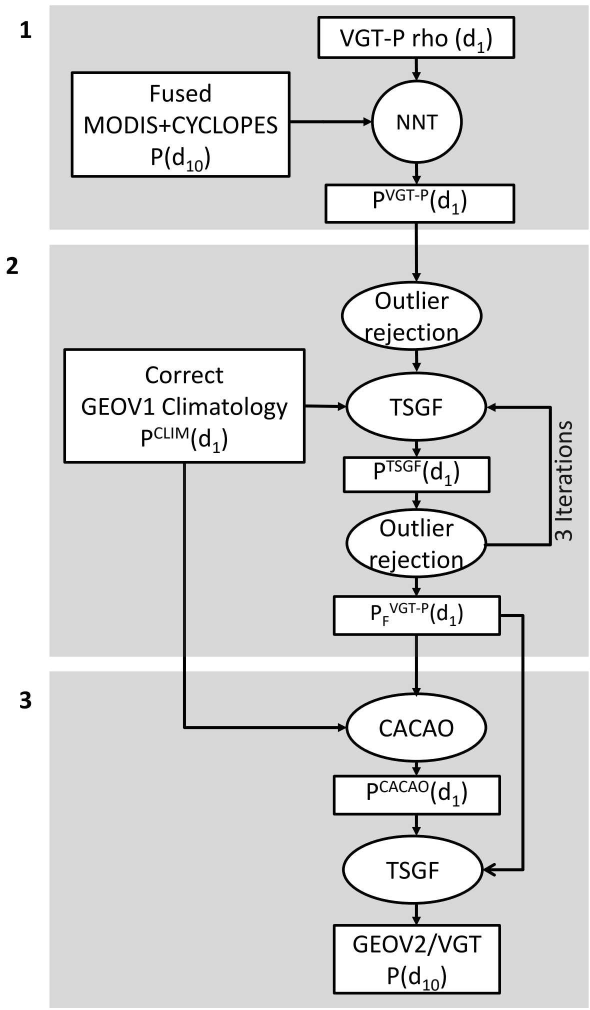

1
VGT-P rho (d1)
Fused MODIS+CYCLOPES P(d10)
NNT
PVGT-P(d1)
2
Outlier rejection
Correct GEOV1 Climatology PCLIM(d1)
TSGF
PTSGF(d1)
Outlier rejection
3 Iterations
PFVGT-P(d1)
3
CACAO
PCACAO(d1)
TSGF
GEOV2/VGT P(d10)

Fig.1

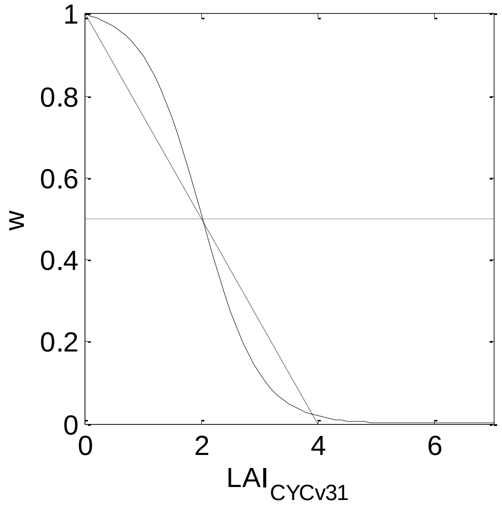
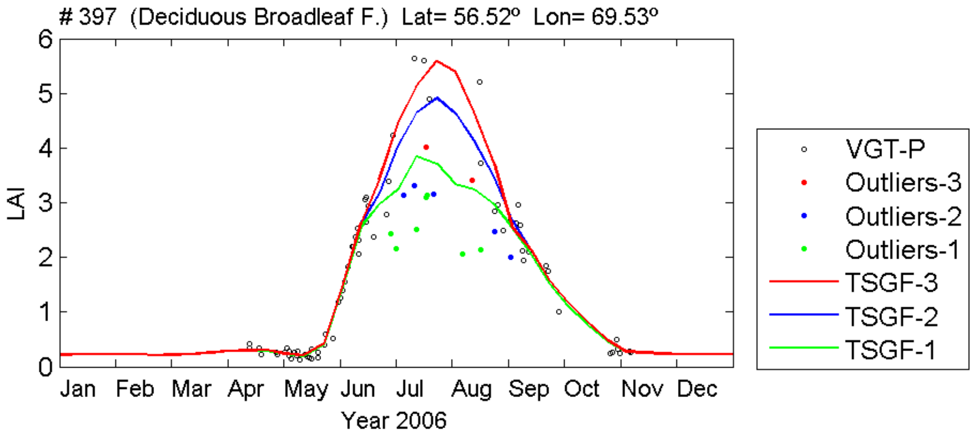
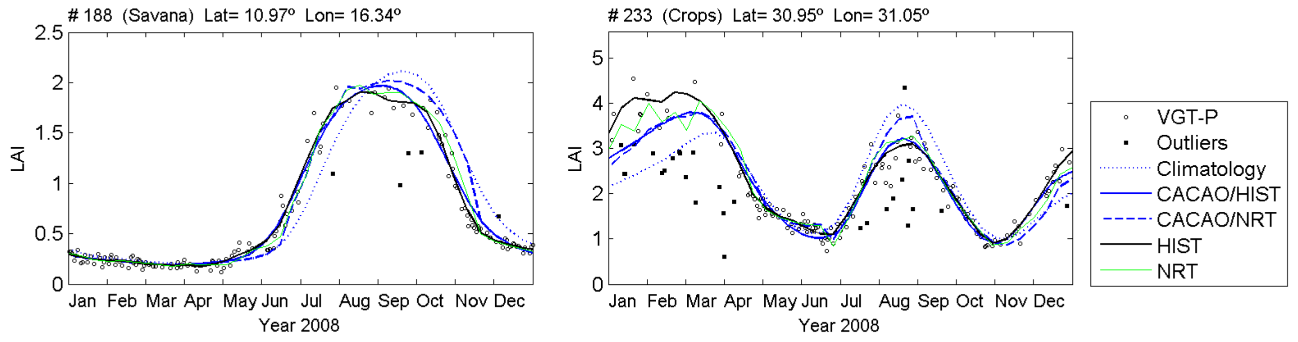


Fig. 2



397 (Deciduous Broadleaf F.) Lat= 56.52º Lon= 69.53º
LAI
Year 2006
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
0
1
2
3
4
5
6
VGT-P
Outliers-3
Outliers-2
Outliers-1
TSGF-3
TSGF-2
TSGF-1

Fig. 3



188 (Savana) Lat= 10.97º Lon= 16.34º
233 (Crops) Lat= 30.95º Lon= 31.05º
LAI
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
Year 2008
VGT-P
Outliers
Climatology
CACAO/HIST
CACAO/NRT
HIST
NRT

Fig. 4

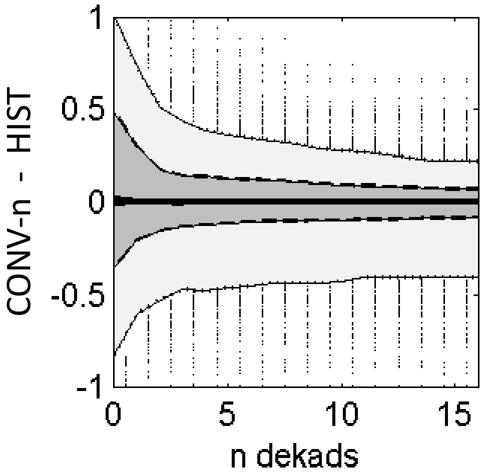


Fig. 5

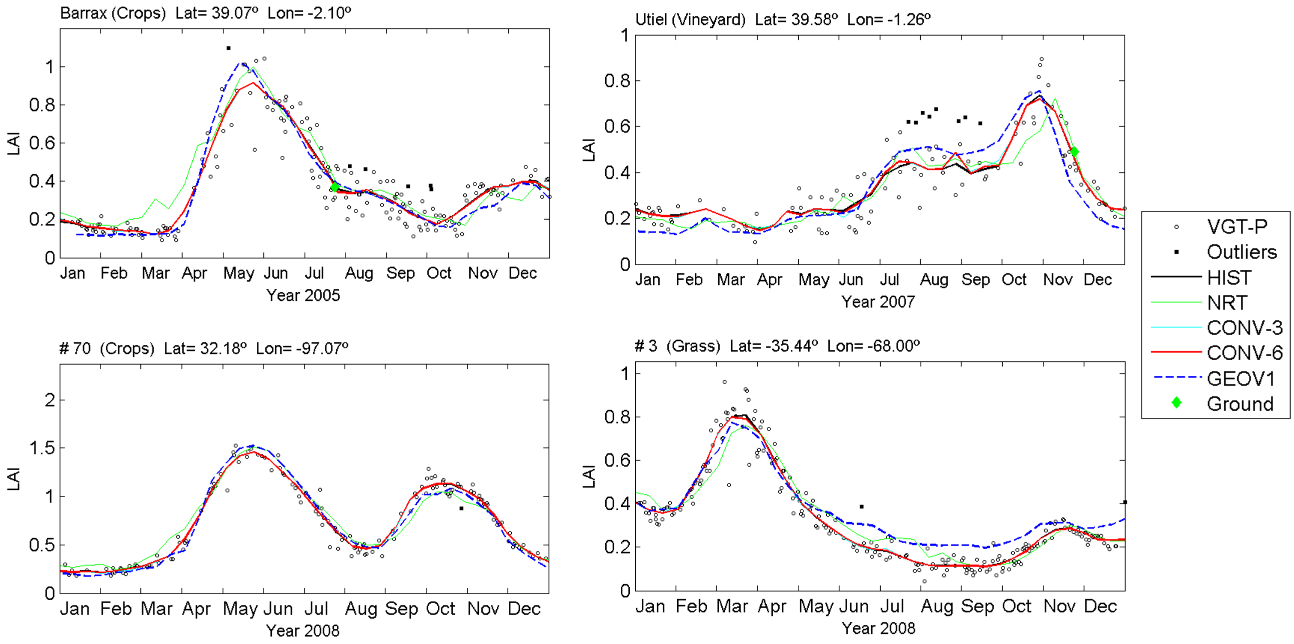

Barrax (Crops) Lat= 39.07º Lon= -2.10º
Utiel (Vineyard) Lat= 39.58º Lon= -1.26º
70 (Crops) Lat= 32.18º Lon= -97.07º
3 (Grass) Lat= -35.44º Lon= -68.00º
LAI
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
Year 2005
Year 2007
Year 2008
VGT-P
Outliers
HIST
NRT
CONV-3
CONV-6
GEOV1
Ground

Fig. 6

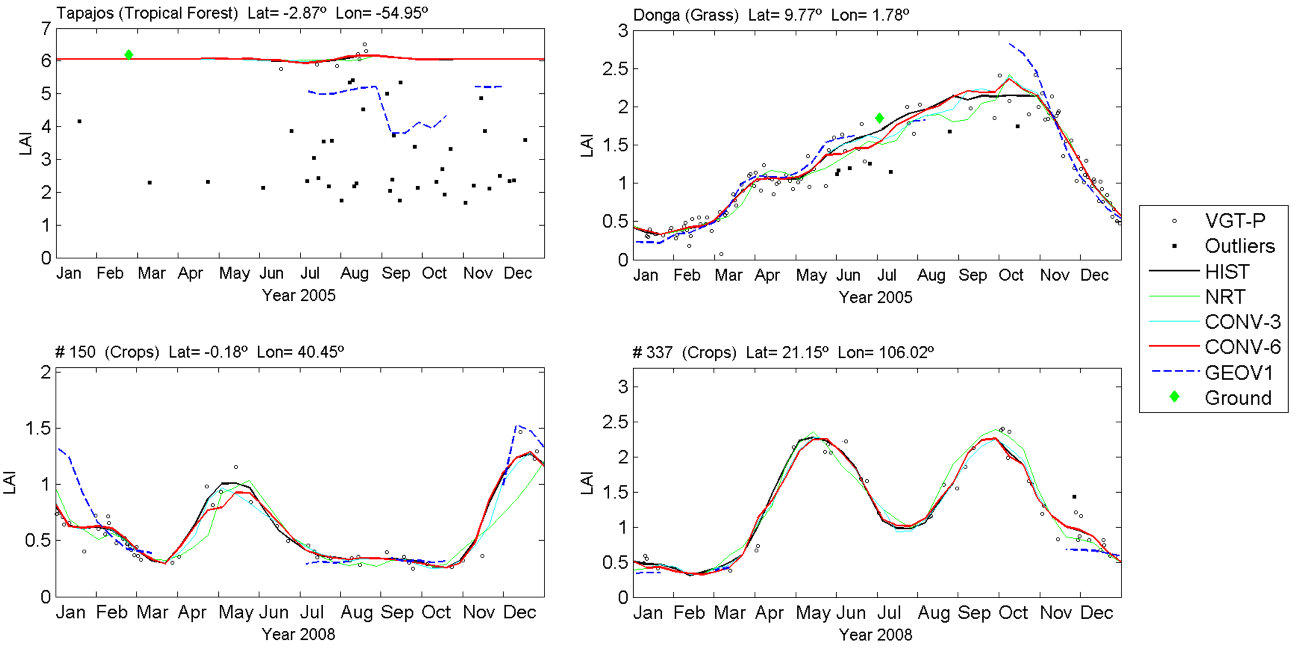
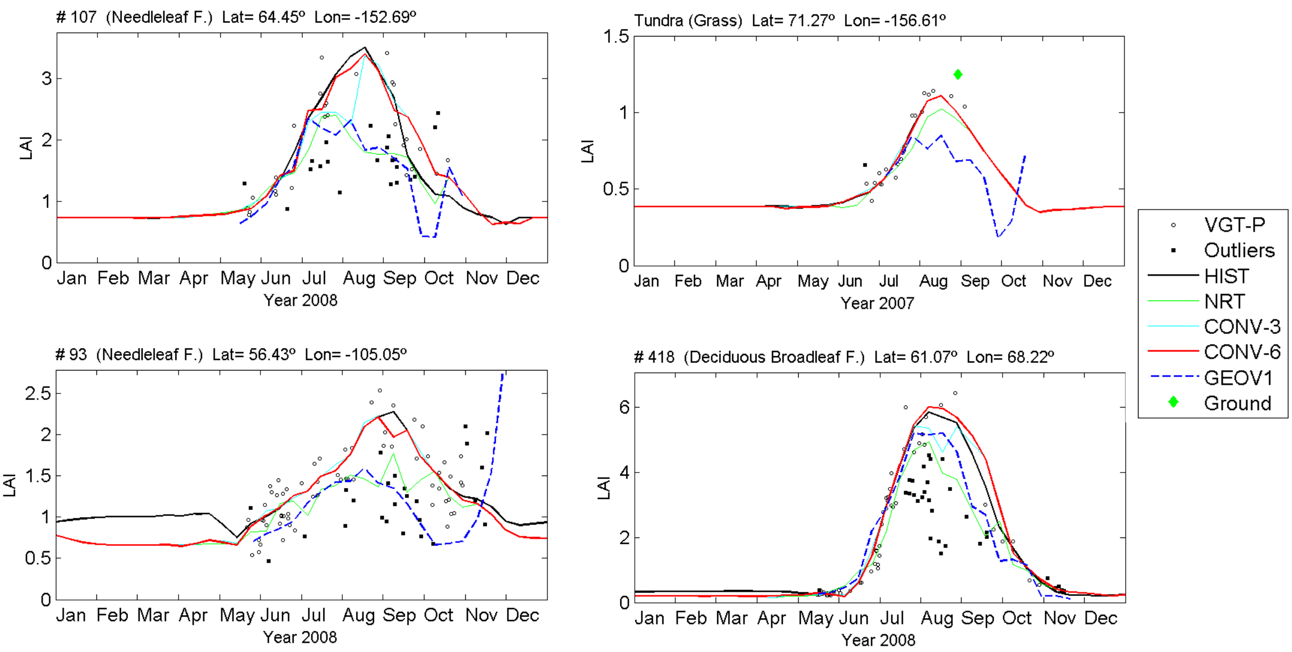
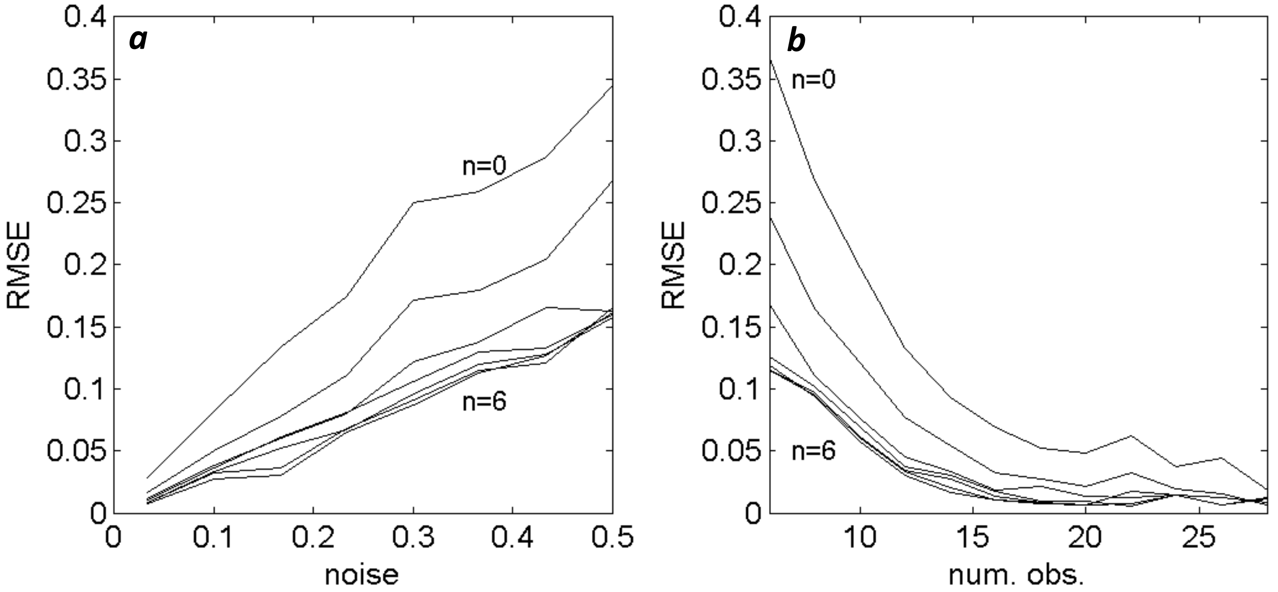

Tapajos (Tropical Forest) Lat= -2.87º Lon= -54.95º
Donga (Grass) Lat= 9.77º Lon= 1.78º
150 (Crops) Lat= -0.18º Lon= 40.45º
337 (Crops) Lat= 21.15º Lon= 106.02º
LAI
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
Year 2005
Year 2008
VGT-P
Outliers
HIST
NRT
CONV-3
CONV-6
GEOV1
Ground

Fig. 7



107 (Needleleaf F.) Lat= 64.45º Lon= -152.69º
Tundra (Grass) Lat= 71.27º Lon= -156.61º
93 (Needleleaf F.) Lat= 56.43º Lon= -105.05º
418 (Deciduous Broadleaf F.) Lat= 61.07º Lon= 68.22º
LAI
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
Year 2008
Year 2007
VGT-P
Outliers
HIST
NRT
CONV-3
CONV-6
GEOV1
Ground

Fig. 8



a
b
RMSE
noise
num. obs.
n=0
n=6
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.5
10
15
20
25

Fig. 9

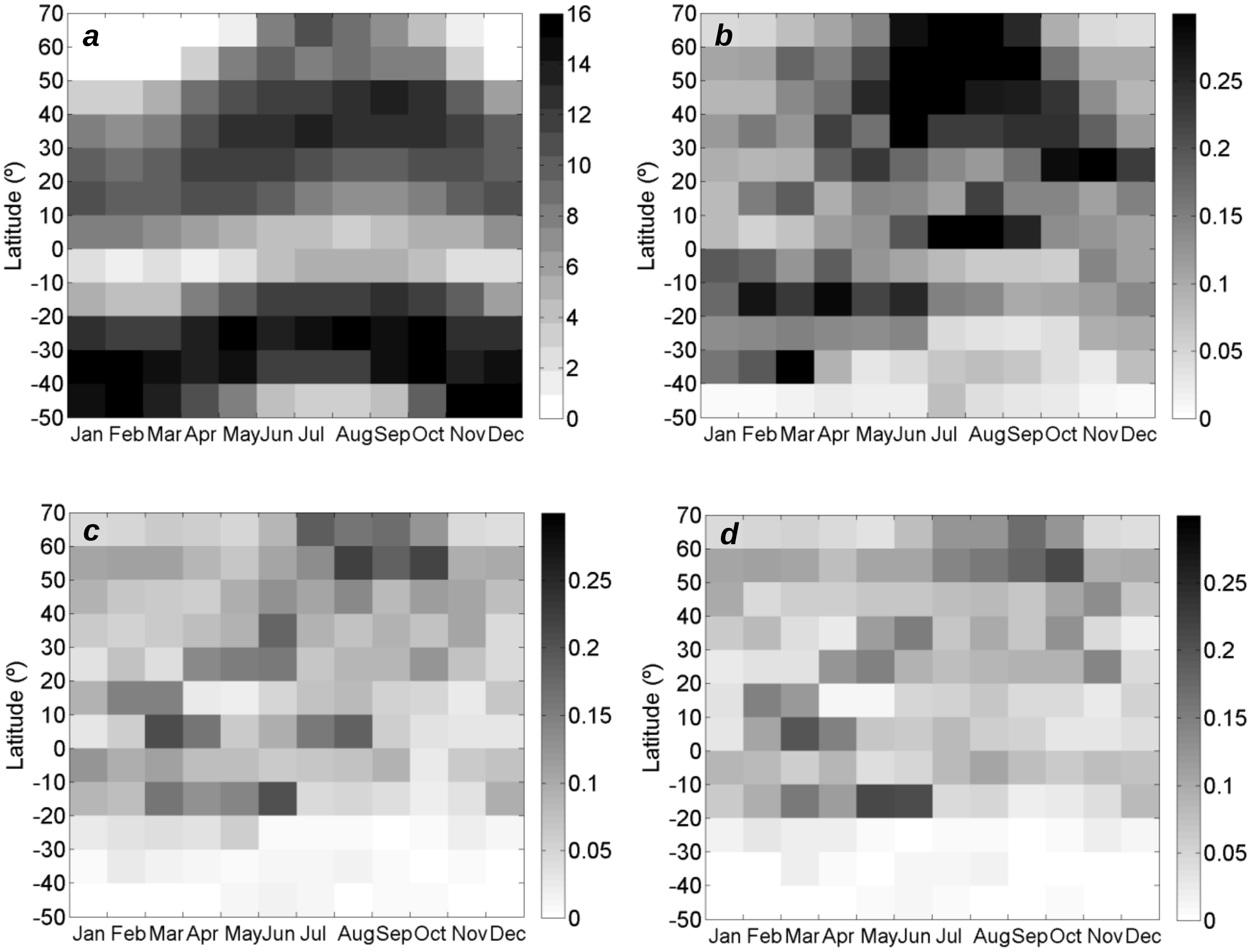

a
b
c
d
Latitude (°)
70
60
50
40
30
20
10
0
-10
-20
-30
-40
-50
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
0
2
4
6
8
10
12
14
16
0.05
0.1
0.15
0.2
0.25

Fig. 10

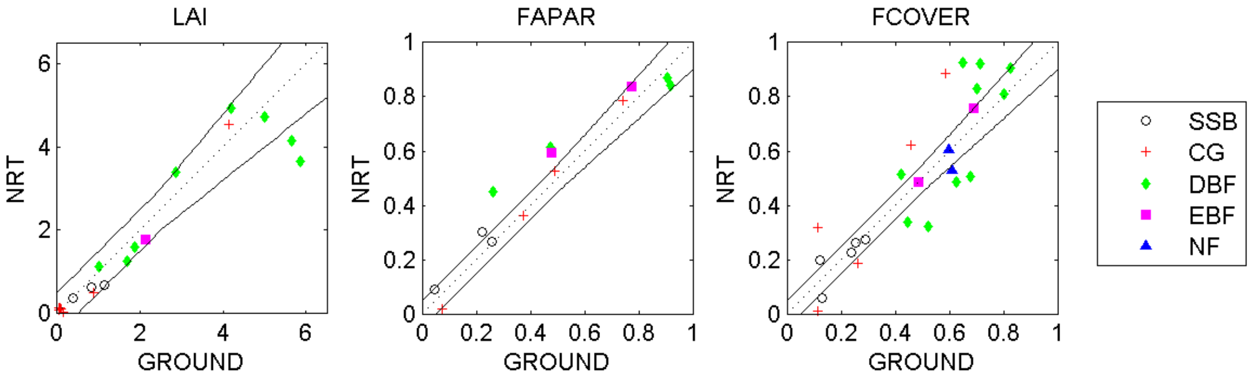


Fig. 11