LONG TERM CONSISTENT GLOBAL GEOV1 AVHRR BIOPHYSICAL PRODUCTS

Aleixandre Verger[1], Frédéric Baret[2], Marie Weiss[3], Roselyne Lacaze[4], Hassan Makhmara[5] and Eric Vermote[6]

1. INRA, UMR1114 EMMAH, Avignon, France; aleixandre.verger@avignon.inra.fr
2. INRA, UMR1114 EMMAH, Avignon, France; baret@avignon.inra.fr
3. INRA, UMR1114 EMMAH, Avignon, France; marie.weiss@paca.inra.fr
4. HYGEOS, Toulouse, France; rl@hygeos.com
5. CNES, Toulouse, France; hassan.makhmara@cnes.fr
6. University Maryland, Washington, USA; eric@ltdri.org

ABSTRACT

Long term global terrestrial vegetation monitoring from satellite Earth Observation system is a critical issue within global climate and earth science modeling applications. This contribution describes the algorithm and a first validation of the global LAI, FAPAR and FCOVER products derived from AVHRR Long Term Data Record (LTDR) for the 1981-2000 period within geoland2 project. GEOV1_AVHRR algorithm aims to ensure robustness and consistency of the derived biophysical products with the ones developed in the recent years, and particularly with the GEOV1_VGT products derived from SPOT/VGT sensor. The approach is based on the use of neural networks and dedicated temporal smoothing and gap filing techniques. The comparison of GEOV1_AVHRR with GEOV1_VGT and with similar products derived from AVHRR GIMMS shows that GEOV1_AVHRR outperforms the current existing products in terms of smoothness and continuity (fraction of valid data of 99.7%). At the same time, GEOV1_AVHRR shows high consistency with GEOV1_VGT and a similar accuracy (RMSE=0.80) as evaluated with ground-based observations. Combination of GEOV1_AVHRR and GEOV1_VGT realizes thus the objective of producing continuous and consistent time series of global observations of LAI, FAPAR and FCOVER for the last three decades. In the next months GEOV1_AVHRR products will be freely accessible at geoland2 portal at 0.05º spatial resolution and 10 days temporal sampling frequency.

INTRODUCTION

The geoland2 project (http://www.geoland2.eu) aims at developing pre-operational services for a range of land applications within GMES. Among the variables intervening in key land surface processes and accessible from remote sensing, LAI (Leaf Area Index) and FAPAR (Fraction of Absorbed Photosynthetic Active Radiation) have been recognized as Essential Climate Variables. The cover fraction (FCOVER) appears also a very pertinent variable for separating the contribution of the soil from that of the canopy. In a first stage GEOV1_VGT LAI, FAPAR and FCOVER variables were derived from VGT/SPOT data for the period 1999-2012 (1) and are accessible at http://www.geoland2.eu/core-mapping-services/biopar.html. Expanding VGT archive back in time is only possible using the Advanced Very High Resolution Radiometers (AVHRR) onboard the NOAA 7-14 series satellite platforms, which is operational since July 1981. In the next months, GEOV1_VGT time series will be completed with GEOV1_AVHRR products derived from AVHRR Long Term Data Record (LTDR) (2). GEOV1_AVHRR products will provide a global coverage at 0.05º spatial resolution and 10 days temporal sampling frequency covering the period from 1981 to 2000. In order to capitalize on the efforts accomplished and get a larger consensus from the user community, GEOV1_AVHRR products aim to ensure robustness and consistency with the ones developed in the recent years, and particularly with GEOV1_VGT. This contribution describes the principles of the GEOV1_AVHRR algorithm along with an evaluation of derived products which are compared with GEOV1_VGT (1), Boston University AVHRR products (3) and ground-based measurements.

ALGORITHM OUTLINE

The derivation of the GEOV1_AVHRR is based on the previous work of (4) demonstrating that neural networks could be trained to consistently estimate a given product from the reflectance measured by a different sensor. This principle is applied here to reproduce GEOV1_VGT from AVHRR LTDR reflectance observations. The resulting products were iteratively filtered to remove outliers due to possible cloud contamination or residual atmospheric effects. Then, filtered products were temporally smoothed and gap filled based on dedicated techniques and the use of the inter-annual climatology to provide consistent and continuous series of biophysical products at a decadal step. In the following, we schematically describe the main steps of the retrieval algorithm (further details are provided in (5)):

1. *Spectral correction of AVHRR LTDR daily reflectances.* This accounts for the slightly differences in spectral sensitivity of the four AVHRR instruments covering the 1981-2000 period.
2. *Daily estimates of the products*. The corrected reflectances are then transformed into daily products using dedicated neural network trained over GEOV1_VGT products. The training process is achieved over the 1999-2000 period in which AVHRR LTDR and GEOV1_VGT are coexisting. The learning data base (input reflectance and output variable) was manually filtered to improve the NNET performances. Inputs are the spectrally corrected AVHRR reflectance values in the red and near infrared (and the cosine of the sun zenith angle at 10h00 only for fAPAR). This will provide time series of daily products.
3. *Scaling the output values* to improve the consistency with GEOV1_VGT products.
4. *First outlier rejection.* Data outside the definition domain of the training process are removed.
5. *Computing the climatology* for every decade over a 20-year period. The median of interannual accumulated values is computed over a ±15-day time window. The gap filling (GF) and temporal smoothing (TS) techniques as defined by (6) are then applied to get more continuous and robust climatology values.
6. *Temporal smoothing and gap filling (TSGF)* the data by applying a local Savitzky-Golay polynomial fitting within an adaptive window (with 5 observations at each side of the date being smoothed) (6), with (large) gaps filled with the climatology values.
7. *Second outlier rejection* based on the Savitzky-Golay filter and a specific upper envelope approach for very noisy data.
8. *Repeat steps 5 and 6* to get cleaner climatology and TSGF series.
9. *Consistent Adjustment of the Climatology to Actual Observations (CACAO)*. The climatology is first decomposed into sub-seasons defined by the lowest and highest seasonal values. For each sub-season, the climatology is then adjusted to the data allowing time shift and magnitude flexibility (7).
10. *Fusing TSGF and CACAO*. The GEOV1_AVHRR products are finally computed as a weighted average between TSGF and CACAO values at each decade, the weights being governed mainly by the distance to actual available observations in a 30-day composition period.

VALIDATION OF GEOV1_AVHRR PRODUCTS

GEOV1_AVHRR products were evaluated though the comparison with GEOV1_VGT (1) and BU_AVHRR products (3). First, the spatial and temporal continuity of products was analyzed through the fraction of valid data. Second, temporal consistency was evaluated from the temporal profiles of a set of control areas and from the smoothness level of the time series. Third, the statistical distributions of the product values were compared. This analysis was undertook over the CEOS BELMANIP2 (BEnchmark Land Multisite Analysis and Intercomparison of Products) ensemble of sites that represent the possible range of surface types and conditions over the Earth (8), and across different biome classes based on the GLOBCOVER land cover map (9). Finally, the accuracy of GEOV1_AVHRR was evaluated through the comparison with ground-based data (10). For brevity, the results presented here are focused on the LAI product.

Fraction of valid data

The fraction of invalid data across biomes (Figure 1) is around 20% for BU_AVHRR and GEOV1_VGT while it is less than 0.3% for GEOV1_AVHRR. For BU_AVHRR most of gaps are observed in Shrubs/Savana/Bare Soil class since for bare surfaces such as deserts, BU_AVHRR LAI product is not computed. The fraction of valid pixels reaches up to 98% if BU_AVHRR pixels flagged as 'bare surfaces' are considered as valid pixels with null LAI. Note however that BU_AVHRR product has a lower spatial and temporal resolution as compared to GEOV1 products since it is generated monthly at 8 km resolution using the average of the two 15-day maximum value AVHRR NDVI composites from the NASA GIMMS dataset as inputs (3). Conversely GEOV1_AVHRR and GEOV1_VGT products are provided every 10 days at a spatial resolution of 4 km and 1 km, respectively. GEOV1_VGT products are not available only when there are less than 2 clear observations in the compositing temporal window (30 days) or when the product value is out of range (LAI less than 0 higher than 7) (1). GEOV1_AVHRR benefits from the use of the TSGF filling and the climatology resulting from the compilation of data over an extended temporal period (20 years). GEOV1_AVHRR products are missing only if the climatology is not available. For GEOV1_AVHRR products, gaps were only detected for Evergreen Broadleaf Forests when discontinuities in the data prevent the computation of the climatology which requires at least 3 over 30 days x 20 years observations for each decade. The GEOV1_AVHRR quality flag indicates the number of available observations for the product composition. Low quality is associated to the products derived from gap filling with no clear observations in the 30-day composition period. This represents between 12% and 48% of the GEOV1_AVHRR data depending on the biome class. Restricting the analysis to the high quality products, the fraction of valid data is relatively similar between GEOV1_VGT and GEOV1_AVHRR which indicates the consistencies in compositing between VGT and AVHRR data. For both sensors, higher fractions of high quality data are observed over Shrubs/Savana/Bare soil, Crops/Grassland and Deciduous Broadleaf Forests. They correspond to locations where cloud occurrence is relatively low. Conversely, Needleleaf and Evergreen Broadleaf Forests achieved the poorest scores due to much higher cloud occurrence locations.

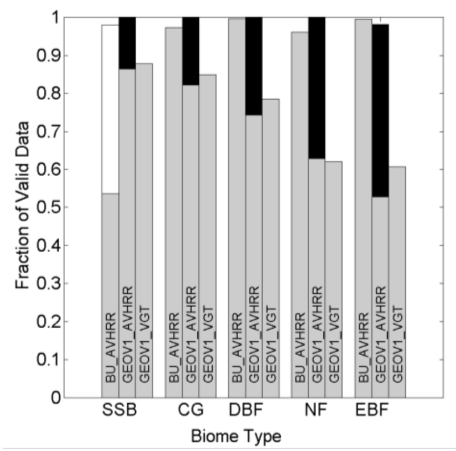

Figure 1. Fraction of valid data per biome over the BELMANIP2 sites for BU_AVHRR (July 1981 to December 2006), GEOV1_AVHRR (July 1981 to December 2000), GEOV1_VGT (December 1998 to December 2010) products. The biome classes are Shrubs/Savana/Bare soil (SSB), Crops and Grassland (CG), Deciduous Broadleaf Forests (DBF), Needleleaf Forest (NF), and Evergreen Broadleaf Forest (EBF). For GEOV1_AVHRR, high quality products (grey) and gap filled products with non valid observations in the compositing period (black) are distinguished. For BU_AVHRR, pixels flagged as 'Bare surfaces' are also separated (white).

Temporal consistency

Since the LAI results from incremental bio-physical processes it is expected to evolve in a relatively smooth manner except in extreme situations such as human intervention. The smoothness level of LAI temporal series was evaluated based on the difference, δ, between *LAI*(*t*) product value at date *t* and the mean value between the two bracketing dates in a maximum period of 60 days (6). The smoother the temporal evolution, the smaller the δ difference should be. The histogram of δ over the whole dataset of BELMANIP2 sites in the 1981-2010 period (Figure 2) shows that both GEOV1_AVHRR and GEOV1_VGT products are smoother than BU_AVHRR with differences lower than 0.5 for most of cases. The GEOV1_AVHRR shows an improvement in the smoothness due to the use of the climatology and the TSGF process which play a major regularization role.

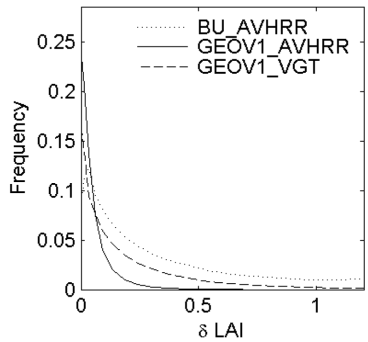

Figure 2. Histogram of the δLAI absolute difference representing temporal smoothness for BU_AVHRR, GEOV1_AVHRR and GEOV1_VGT LAI products.

The performance of the different products to reproduce consistently and continuously the temporal evolution of vegetated surfaces is illustrated in the temporal profiles of Figure 3. High temporal consistency is observed between GEOV1_AVHRR and GEOV1_VGT products with a significant good agreement between concomitant data in the years 1999 and 2000. BU_AVHRR shows higher differences both in terms of temporal consistency as well as the magnitude of the LAI time series. GEOV1_AVHRR product generally reproduces the expected seasonality and corrects most of the artifacts observed in BU_AVHRR (e.g. anomalous seasonally and very noisy time series for the site #435 corresponding to Evergreen Broadleaf Forest) and GEOV1_VGT (e.g. artificial peak in winter time in site #98). Further, GEOV1_AVHRR provides continuous and robust estimates in the periods and locations where GEOV1_VGT presents gaps. The benefits of GEOV1_AVHRR in terms of consistency and continuity as compared to BU_AVHRR and GEOV1_VGT are particularly significant in problematic areas such as the Equator (e.g. site #435) where data is missing for large periods or of bad quality due to persistent clouds, as well as at very high latitudes (e.g. site #98) where problems are mainly due to snow cover and directional effects.

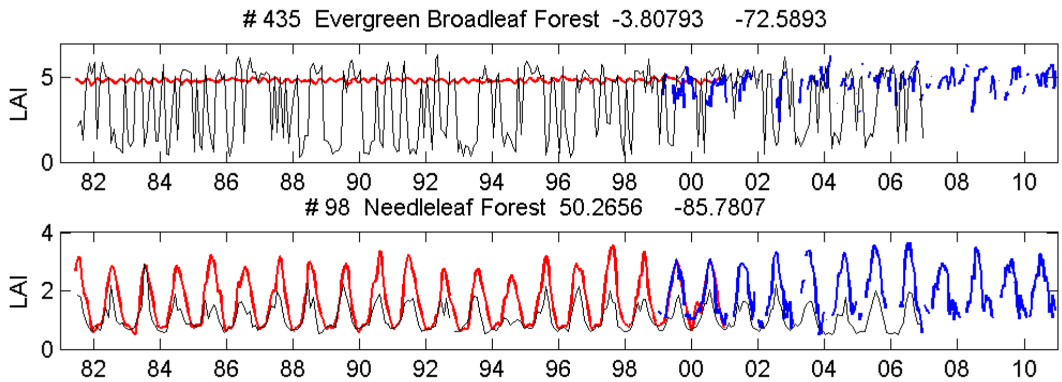

Figure 3. Temporal profiles of the GEOV1_AVHRR (red line), GEOV1_VGT (blue line) and BU_AVHRR (black line) LAI products.

Statistical distributions

Histograms of LAI values (Figure 4) show very consistent distributions between GEOV1_AVHRR and GEOV1_VGT products which indicates qualitative agreement between the two datasets. Indeed, GEOV1_AVHRR product is learned from GEOV1_VGT, using the BELMANIP2 database over 1999-2000 period. However, the statistical distributions remain very consistent through the other years. Higher discrepancies are found for BU_AVHRR product that shows bimodal distributions with a first peak for low LAI values around 0.2 and a second peak for high values around 5 but with very low frequencies for intermediate LAI values from 3 to 5.

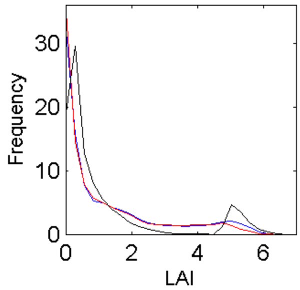

Figure 4. Distributions of GEOV1_AVHRR (red line), GEOV1_VGT (blue line) and BU_AVHRR (black line) LAI products.

Validation with ground-based observations

A direct validation was performed over a limited dataset of 28 available ground-based measurements (10) acquired over 13 different 3 km x 3 km sites in the period from 1999 to 2000 where the different analyzed satellite products exist. The value of each product was interpolated at the date of the ground measurements if two valid values exist within a maximum period of ±30 days. The comparison of GEOV1_AVHRR with the ground-based observations shows an overall accuracy of 0.80 in terms of RMSE and a linear correlation coefficient R=0.84 (Figure 5). A similar performance is found for GEOV1_AVHRR (RMSE=0.73 and R=0.83) while a lower accuracy is observed for BU_AVHRR (RMSE=1.00 and R=0.69) over the same dataset.

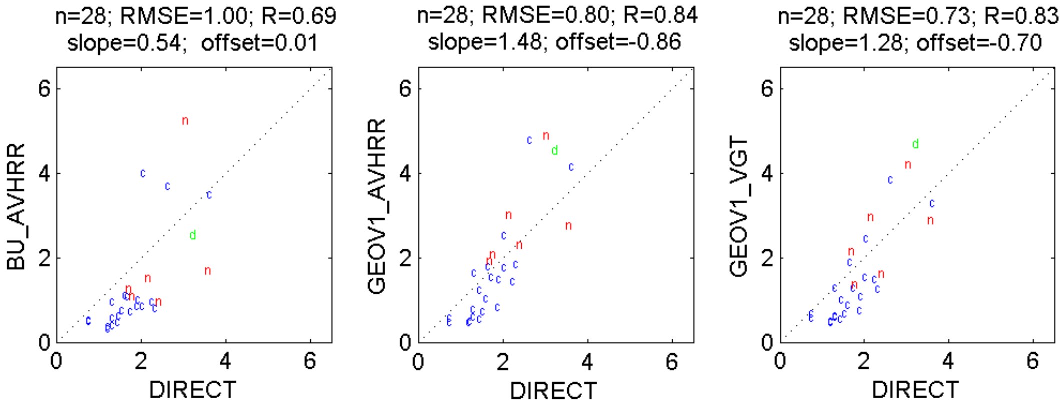

Figure 5. Comparison of BU_AVHRR, GEOV1_AVHRR and GEOV1_VGT LAI product with direct ground measurements of true LAI. Letter markers correspond to different biome classes: crops and grassland (c), needleleaf forest (n) and deciduous broadleaf forest (d). Slope and offset were calculated with robust linear regression.

CONCLUSIONS

The GEOV1_AVHRR LAI, FAPAR and FCOVER products capitalize on the efforts undertaken to pre-process the AVHRR temporal series, resulting in the LTDR data (2), and the recent development of improved processing of biophysical products resulting in the GEOV1_VGT data (1). The GEOV1_AVHRR algorithm was designed to provide smooth and continuous time series and special emphasis was put on achieving consistency with GEOV1_VGT products. The accuracy assessment of GEOV1_AVHRR products and the comparison with GEOV1_VGT indicate a very good agreement between the two datasets with a similar level of accuracy as evaluated through the comparison with ground-based measurements. Further a significant improvement in terms of smoothness and continuity is achieved in GEOV1_AVHRR products as compared to GEOV1_VGT and current existing products based on the AVHRR GIMMS dataset. The combination of GEOV1_AVHRR (1981-2000) and GEOV1_VGT (1999-2012) products results in consistent time series of LAI, FAPAR and FCOVER global products for the last three decades. GEOV1_VGT products are available at http://www.geoland2.eu/core-mapping-services/biopar.html and GEOV1_AVHRR are being processed and, in mid 2012, they will be freely delivered at 0.05º (around 4 km) spatial resolution and every 10 days through the geoland2 website. The proposed generic algorithm will be also applied to other sensors (e.g. AVHRR/METOP, PROBA-V, Sentinel-3) for continuation after 2012.

ACKNOWLEDGEMENTS

This research was supported by FP7 GEOLAND2 project. The first author was funded by the VALi+d postdoctoral program (FUSAT, GV-20100270). We thank S. Ganguly and R. Myneni for making the BU_AVHRR data available.

REFERENCES

1 Baret, F., Weiss, M., Lacaze, R., Camacho-de Coca, F., Pacholcyzk, P., Smets, B., 2010. *Consistent and accurate LAI, FAPAR and FCOVER global products: principles and evaluation of GEOLAND2 products*. In Third International Symposium on Recent Advances in Quantitative Remote Sensing. Torrent (Spain), 208-213.

2 Vermote, E., Justice, C., Csiszar, I., Eidenshink, J., Myneni, R., Baret, F., Masuoka, E., Wolfe, R., 2009. A Terrestrial Surface Climate Data Record for Global Change Studies. In American Geophysical Union Fall Meeting. San-Francisco (USA), EOS-Trans AGU 90(52), IN42A-08.

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

4 Verger, A., F. Baret , M. Weiss, 2008. Performances of neural networks for deriving LAI estimates from existing CYCLOPES and MODIS products. Remote Sensing of Environment, 112, 2789-2803.

5 Baret, F., Weiss, M., Verger, A., Kandasamy, S., 2011. BioPar Methods Compendium - LAI, FAPAR and FCOVER from LTDR AVHRR series. INRA-EMMAH: Avignon. p. 50.

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

7 Verger, A., Baret, F., Weiss, M., Kandasamy, S., Vermote, E. 2011. Quantification of LAI interannual anomalies by adjusting climatological patterns. In Proceedings of 6th International Workshop on the Analysis of Multi-Temporal Remote Sensing Images, Trento (Italy), 113-116.

8 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 (7), 1794-1803.

9 Bicheron, P., Leroy, M., Brockmann, C., Krämer, U., Miras, B., Huc, M., Ninô, F., Defourny, P., Vancutsem, C., Arino, O., Ranéra, F., Petit, D., Amberg, V., Berthelot, B., Gross, D, 2006, GLOBCOVER: A 300 m global land cover product for 2005 using ENVISAT MERIS time series. In Proceedings of the 2nd Recent Advances in Quantitative Remote Sensing, Torren (Spain), 538-542.

10 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 (G02028).