Quality Assessment Report

GEOV2-AVHRR: Leaf Area Index (LAI), Fraction of Absorbed Photosynthetically Active Radiation (FAPAR) and Fraction of green Vegetation Cover (FCover) from LTDR AVHRR

THEIA-RP-44-0281-CSIC

Issue I2.00

Submission date: 21.03.2023

Organization name of lead contractor for this deliverable: CSIC

Author: Aleixandre Verger (CSIC)

Document Release Sheet

Book captain:	Aleixandre Verger		Sign Date 21.03.2023		
Approval:	Thérèse Barroso		Sign Date 12.05.2023		
Endorsement:			Sign Date		
Distribution:					

Change Record

Issue/Rev	Date	Page(s)	Description of Change	Release
	2.12.2015	All	First Issue: Validation of prototyped estimates	I1.00
I1.00	21.03.2023	All	Validation of the official GEOV2-AVHRR products for the period 1981-2021	I2.00

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

ANN	Artificial Neural Network
ATBD	Algorithm theoretical based Document
AVHRR	Advanced Very High Resolution Radiometer
BRDF	Bidirectional Reflectance Distribution Function
BS	Bare Soil
CACAO	Consistent Adjustment of Climatology to Actual Observations
CEOS	Committee for Earth Observation Satellite
CGLS	Copernicus Global Land Service
EBF	Evergreen Broadleaf Forest
ECV	Essential Climate Variable
FAPAR	Fraction of Absorbed Photosynthetically Active Radiation
FCOVER	Fraction of vegetation cover
GAC	Global Area Coverage
GAI	Green Area Index
GCOS	Global Climate Observation System
GEOV1	GEOLAND2 Version 1 product
GEOV2	GEOLAND2 Version 2 product
GIMMS3g	Third generation Global Inventory Monitoring and Modeling System
GLASS	Global LAnd Surface Satellite
GTOS	Global Terrestrial Observation System
HIST	Offline processing of historical time series
L2	Level 2 product
L3	Level 3 product
LAC	Local Area Coverage
LAI	Leaf Area Index
LTDR	Long Time Data Record
LTS	Long Time Series
MetOp	Meteorological Operational satellite
MODIS	aging Spectrometer
NDVI	Normalized Difference Vegetation Index
NIR	Near Infrared
NNT	Neural Network Technique
NRT	Near real Time
RMSE	Root Mean Square Error
RTM	Radiative Transfer Model
SPOT	Satellite Pour l'Observation de la Terre
SWIR	Short Wave Infrared
TOA	Top of Atmosphere
TOC	Top of Canopy
VI	Vegetation Index

1 BACKGROUND OF THE DOCUMENT

1.1 EXECUTIVE SUMMARY

GEOV2-AVHRR global vegetation products of leaf area index (LAI), fraction of absorbed photosynthetic active radiation (FAPAR) and vegetation cover fraction (FCover) were produced and distributed by CNES based on the algorithm developed by CREAF-CSIC and INRAE [THEIA_ATBD_GEOV2_AVHRR I2.30] in the framework of the Theia Land Data Centre. They were derived from advanced very high-resolution radiometer (AVHRR) long term data record (LTDR) version 4 reflectances for the period 1981-2018. These time series of biophysical products provide information on the status and evolution of land surface at global scale every ten days. The GEOV2-AVHRR products were designed to ensure (i) Global Climate Observing System (GCOS 2010) requirements and (ii) high consistency with the biophysical products developed in the recent years, and particularly with Version 2 of Copernicus Global Land Service 1km, GEOV2-CGLS, LAI, FAPAR and FCover products derived from VEGETATION and PROBA-V sensors.
Quality Assessment constitutes the only means of guaranteeing the compliance of generated products with user requirements. It concerns new products which must pass an exhaustive scientific evaluation before to be implemented operationally. The procedure follows, as much as possible, the guidelines, protocols and metrics defined by the Land Product Validation (LPV) group of the Committee on Earth Observation Satellite (CEOS) for the validation of satellite-derived land products.
This document presents a Quality Assessment of GEOV2-AVHRR LAI, FAPAR and FCover time series for the period 1981-2018. The analysis is mainly focused on the inter-comparison of GEOV2-AVHRR with similar existing and validated datasets as well as with ground data. In particular, GEOV2-AVHRR will be compared with GEOV2-CGLS (Verger et al. 2014; Verger et al. 2023) products for the period 1999 - 2018. We also compared GEOV2-AVHRR with existing long term GIMMS3g (Zhu et al. 2013), and GLASS (Liang et al. 2013) products derived from AVHRR time series since 1981. We assessed the spatio-temporal continuity and consistency of the long time series across products, variables and sensors. We evaluated their capability for discerning anomalies and trends. Finally, we performed a direct comparison of GEOV2-AVHRR estimates with ground data over a limited number of samples available. The assessment of GEOV2-AVHRR time series over an extended period for 2019-2021 is presented in annex, section 6.

1.2 SCOPE AND OBJECTIVES

The scope of this paper is to present the quality assessment of GEOV2-AVHRR LAI, FAPAR and FCover products for the period 1981-2018. The assessment over an extended period for 2019-2021 is presented in annex, section 6.

The objective is to validate GEOV2-AVHRR LAI, FAPAR and FCover products based on the comparison with existing GEOV2-CGLS, GIMMS3g and GLASS products and ground data in terms of consistency, continuity and accuracy.

1.3 CONTENT OF THE DOCUMENT

This document is structured as follows:

Chapter 2. Data and methods

Chapter 3. Quality assessment of GEOV2-AVHRR prototyped estimates:

- Spatial consistency
- Comparison with GEOV2-CGLS, GIMMS3g and GLASS products
- Temporal consistency
- Accuracy assessment over the DIRECT2 sites

Chapter 4. Conclusions

Annex: assessment of extended time series for 2019-2021

1.4 RELATED DOCUMENTS

1.4.1 Applicable documents

Document ID	Descriptor
CNES-CSIC contract Nº 5700007172	CNES contracts scientific support of CSIC for the validation of the GEOV2-AVHRR products
THEIA-CT-44-0163-CNES	Scientific support requested
THEIA-SP-44-0207-CREAF, THEIA_ATBD_GEOV2_AVHRR_I2.30	Algorithm theoretical baseline description (ATBD) of GEOV2-AVHRR
THEIA-MU-44-381-CNES	Product user manual of GEOV2-AVHRR products

2 DATA AND METHODS

2.1 BIOPHYSICAL PRODUCTS

The main characteristics of the vegetation biophysical products here investigated are provided in Table 1.

2.1.1 GEOV2-AVHRR

The GEOV2-AVHRR are derived from version 4 LTDR surface reflectance data (https://ltdr.modaps.eosdis.nasa.gov/cgi-bin/ltdr/ltdrPage.cgi) at 10-day frequency, 0.05° (~5 km at the Equator) spatial resolution and temporal span from July 1981 to December 2018. The retrieval algorithm is described in [THEIA_ATBD_GEOV2_AVHRR I2.30]. It relies on neural networks trained with CYCLOPES version 3.1 and MODIS Collection 5 products. The input data are AVHRR top of the canopy directionally normalized daily reflectances in the red and near infrared bands. For the derivation of FAPAR, the cosine of the sun zenith angle at 10:00 solar time is also used as input since we defined it as the black-sky fAPAR at 10:00 solar time which is very close to the daily integrated FAPAR (Baret et al. 2003; Fensholt et al. 2004). The outputs of neural networks are daily LAI, FAPAR and FCover values. Daily estimates are then composited, smoothed and gap filled using a multistep Savitzky–Golay (SG) filtering procedure combined with a climatological gap-filling approach to get continuous and more robust products every 10 days. In particular, the GEOCLIM climatology (Verger et al. 2015) and CACAO (Verger et al. 2013) method were used for this purpose. Several quality indicators are associated to each product for a given pixel and dekadal date. They include the number of valid daily estimates in the compositing window, the length in days of the compositing window, the pathway used to fill the gaps in the time series and a proxy of the uncertainty computed as the root mean square error (RMSE) between the daily LAI, FAPAR and FCover estimates in the compositing window and the final 10-day GEOV2-AVHRR products. Further details on the algorithm and product characteristics are provided in [THEIA_ATBD_GEOV2_AVHRR I2.30] and [THEIA-MU-44-381-CNES], respectively.

2.1.2 GEOV2-CGLS

Version 2 of Copernicus Global Land Service 1km products of LAI, FAPAR and FCover (called hereafter GEOV2-CGLS) were derived from SPOT/VEGETATION and PROBA-V data at 10-day frequency for the period 1999 - June 2020 (Verger et al. 2014; Verger et al. 2023). GEOV2-CGLS products were designed to balance the known pros and cons of already existing CYCLOPES (Baret et al. 2007), MODIS (Myneni et al. 2002) and GEOV1-CGLS (Baret et al. 2013) products. The retrieval algorithm of GEOV2-CGLS is very similar to GEOV2-AVHRR one. It also includes two steps. First, neural networks are trained with CYCLOPES version 3.1 and MODIS Collection 5 products similarly as in GEOV2-AVHRR. They were trained using VEGETATION reflectances in the red, near infrared and shortwave infrared bands as well as the zenith and azimuth angles of the associated sun-view geometry. Differently to GEOV2-AVHRR, specific neural networks are trained for evergreen broadleaf forests (EBFs) and non EBFs. The second step of the GEOV2-CGLS retrieval algorithm is the smoothing, gap filling and 10-day composition. This step is identical to the one applied in GEOV2-AVHRR algorithm with only some parameter adaptations to the particularities of the data.

2.1.3 GIMMS3g

The GIMMS3g LAI and FAPAR products were derived from GIMMS3g normalized difference vegetation index (NDVI) product (Zhu et al. 2013). They have 15-day temporal fequency, 0.08° spatial resolution (~8 km at the Equator) and temporal span from July 1981 to December 2011. The LAI and FAPAR retrieval algorithms (Zhu et al. 2013) use Feed-Forward Neural Networks (FFNN) trained for each month with reprocessed MODIS LAI and FAPAR products (Yuan et al. 2011) and GIMMS3g NDVI data from 2000 to 2009. A set of twelve FFNN models were trained for LAI and another set of twelve FFNNT for FAPAR using the latitude, longitude, the MODIS land cover class (MCD12C1) and GIMMS3g NDVI as input data (Zhu et al. 2013). The maximum NDVI value during the 15-day compositing period is used because residual atmospheric effects decrease the magnitude of NDVI (Zhu et al. 2013).

2.1.4 GLASS

The GLASS AVHRR LAI and FAPAR products were derived from version 4 LTDR surface reflectance data, and FCover was derived from version 5 LTDR (https://ltdr.modaps.eosdis.nasa.gov/cgi-bin/ltdr/ltdrPage.cgi). They have 8-day temporal resolution and 0.05° (~5 km) spatial resolution. The GLASS LAI time series span from 1981 to 2018 while FAPAR and FCover products start in 1982. We used version 4 of GLASS AVHRR products (http://www.glass.umd.edu/Download.html). Different retrieval algorithms are used to generate each variable:

The GLASS AVHRR LAI algorithm (Xiao et al. 2016) used a general regression neural networks (GRNNs) trained with MODIS Collection 5 and CYCLOPES version 3.1 LAI products over the BELMANIP1 (Baret et al. 2006) for 2003 and 2004. MODIS and CYCLOPES products were fused based on their uncertainties. The GLASS LAI values correspond to the true LAI. The CYCLOPES LAI defined as effective LAI was first transformed into true LAI using a biome specific clumping index map (Pisek et al. 2010) and the MODIS land cover (MOD12Q1, layer 3). GRNNs were trained for each MODIS biome type using the fused CYCLOPES and MODIS products and LTDR reflectance data from an entire year. Unlike other existing methods such as GEOV2-AVHRR or GIMMS3g that use input data acquired only at a specific time, the GLASS GRNNs ingest reprocessed AVHRR reflectances from an entire year to estimate the one-year LAI profile. The AVHRR reflectance data were preprocessed to remove cloud contamination and fill any missing data using the time-series cloud detection (TSCD) algorithm (Tang et al. 2013).

The GLASS AVHRR FAPAR product is derived from GLASS LAI to ensure physical consistency between LAI and FAPAR retrievals assuming that FAPAR can be aproximated as the complementary of the transmitted PAR down to the soil and approximating the radiation transmitted through the canopy by an exponential model with LAI and the clumpling index (Xiao et al. 2015). The GLASS LAI product and the MODIS derived clumpling index map (He et al. 2012) were used as input data to calculate the canopy transmittance in GLASS FAPAR algortihm (Xiao et al. 2015). The

GLASS AVHRR FAPAR values correspond to the total FAPAR at 10:30 A.M. local time, which is a close approximation of the daily average FAPAR (Fensholt et al. 2004).
The GLASS AVHRR FCover product is generated using the multivariate adaptive regression splines method (MARS) trained with GLASS MODIS FCover (Jia et al. 2015) and reprocessed smoothed and gap filled AVHRR reflectance data (Jia et al. 2019). The AVHRR FCover values were rescaled to MODIS estimates using a linear correction calibrated at pixel level from the 2000 to 2015 overlapped years. The GLASS MODIS FCover product was also estimated with the MARS method, whereas the training samples were generated from finer resolution Landsat FCover estimates (Jia et al. 2015).

Table 1. Description of the used LAI, FAPAR and FCover products.

		GEOV2-AVHRR	GEOV2-CGLS	GIMMS3g	GLASSV4
Input data		NTOC-r red and NIR LTDR v4	TOC-r red, NIR and SWIR, and cosine of the 3 angles of sun and view directions	NDVI GIMMS3g	NTOC-r red and NIR LTDR v4, v5* *FCover
Sensor AVHRR (A) Vegetation (VGT) **Satellite** NOAA (N) SPOT (S) PROBA (P) **Start date**		A2 N7 1981/06 A2 N9 1985/01 A2 N11 1988/11 A2 N14 1995/01 A3 N16 2000/11 A3 N18 2005/07 A3 N19 2009/06	VGT S 1999/01 VGT P 2014/01	A2 N7 1981/06 A2 N9 1985/03 A2 N11 1988/11 A2 N9 1994/09 A2 N14 1995/01 A3 N16 2000/11 A3 N17 2004/01 A3 N18 2009/01	A2 N7 1981/06 A2 N9 1985/01 A2 N11 1988/11 A2 N14 1995/01 A3 N16 2000/11 A3 N18 2005/07 A3 N19 2009/06
Precision	**LAI**	0.033	0.033	0.1	0.01
	FAPAR	0.004	0.004	0.01	0.004
	FCover	0.004	0.004	-	0.004
Spatial sampling		0.05°	1/112°	1/12°	0.05°
Temporal sampling		10 days	10 days	15 days	8 days
Time period		1981-2018	1999-June 2020	1981-2011	1981/82*-2018 *FAPAR, FCover
Reference		Verger et al. 2021-This issue	Verger et al. 2023	Zhu et al. (2013)	Jia et al. (2019); Xiao et al. (2015); Xiao et al. (2016)

2.1 VALIDATION APPROACH

The quality assessment of GEOV2-AVHRR products was based on the cross-comparison with GEOV2-CGLS, GIMMS3g and GLASSV4 products as well as the comparison with available ground data. The validation approach follows CEOS WGCV LPV protocols and good practices (Fernandes et al. 2014):

- *Spatial consistency and product completeness:* We assessed the spatial consistency of GEOV2-AVHRR biophysical variables and their associated quality indicators. We also explored the fraction and distribution in space and time of the missing data of satellite products mainly due to cloud and snow cover.
- *Comparison with GEOV2-CGLS, GIMMS3g and GLASSV4:* We assessed the consistency of GEOV2-AVHRR with similar satellite products through the maps of mean differences for the overlapping period. For comparison purposes, we aggregated the satellite products at a spatial resolution of 0.5°. The 0.5° spatial resolution corresponds to the typical resolution of global models and reduces computation time. The GIMMS3g and GLASSV4 LAI, FAPAR and FCover time series were linearly interpolated at the dates of GEOV2-AVHRR products at 10-day time step. The GEOV2-CGLS were already available at 10-day step for the same dates, i.e. 5th, 15th and 25th of each month.
- *Statistical analysis*: We conducted a more quantitative statistical analysis over the BELMANIP2.1 sites for all samples (pixels x dates) in the overlapping period. BELMANIP2.1 includes 445 homogenous sites (Figure 1) representative of the global distribution of the vegetation surface while showing little topography and good level of homogeneity (Weiss et al. 2014). The computed metrics are the 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)), the root mean square error (RMSE), the mean bias, the correlation coefficient (R), the slope and the offset of the least squares linear regression. The distribution of product values per biome type was also analyzed. For comparison purposes, the GEOV2-CGLS were resampled at 5km x 5km. The original resolution of AVHRR products (0.05° for GEOV2-AVHRR and GLASSV4, and 0.08° for GIMMS3g) was kept but at the GEOV2-AVHRR 10-day temporal sampling.
- *Temporal consistency:* We analyzed first the temporal profiles of the satellite datasets over a sample of representative sites of different biome types where ground measurements were also available. The original spatial and temporal resolution of satellite products was used except for GEOV2-CGLS that was resampled at 5km x 5km. The smoothness was also used as a metrics of the temporal consistency. LAI variable results from incremental bio-physical processes. It is therefore expected to show relatively smooth temporal variations except in extreme situations such as flooding, fire or changes in the land-use. High variability in the temporal profiles would indicate a lack of reliability of the derived products. The smoothness of the LAI temporal series was evaluated based on the absolute value of the difference, δLAI, between *LAI*(*t*) product value at date *t* and the mean value between the two closest bracketing dates in a maximum Δt period of 60 days: $\delta LAI = |1/2(LAI(t + \Delta t) + LAI(t -$

$\Delta t)) - LAI(t)|$ (Verger et al. 2011). The smoother the temporal evolution, the smaller the δLAI difference should be.

We also compared the temporal anomalies and trends derived from these long term time series. We analyzed the latitudinal temporal pattern of GEOV2-AVHRR, GIMMS3g and GLASSV4 standardized anomalies over the entire period of each AVHRR time series using Hovmöller diagrams (Hovmöller 1949). The temporal trends of GEOV2-AVHRR, GIMMS3g, and GLASSV4 were analysed based on the non-parametric Mann Kendall trend test for the overlapping period 1982-2011. These tests were conducted for LAI at 0.5° spatial resolution and at the original temporal resolution.

- *Accuracy assessment*: More quantitative assessment was achieved using the available ground-based measurements acquired over 3 km x 3 km DIRECT 2.0 sites in the 1999-2017 period (http://calvalportal.ceos.org/web/olive/site-description; Figure 1). For LAI validation, we compared the satellite products both with ground measurements of true LAI, i.e. accounting for clumping, and effective LAI, i.e. assuming a random distribution of the vegetation elements in the canopy. Each product was interpolated at the date of the ground measurements if two valid dekadal data exist within a maximum period of ±30 days. For comparison purposes, the different satellite products (GEOV2-AVHRR, GEOV2-CGLS, GIMMS3g and GLASSV4) were also validated over the common samples where all products were available.

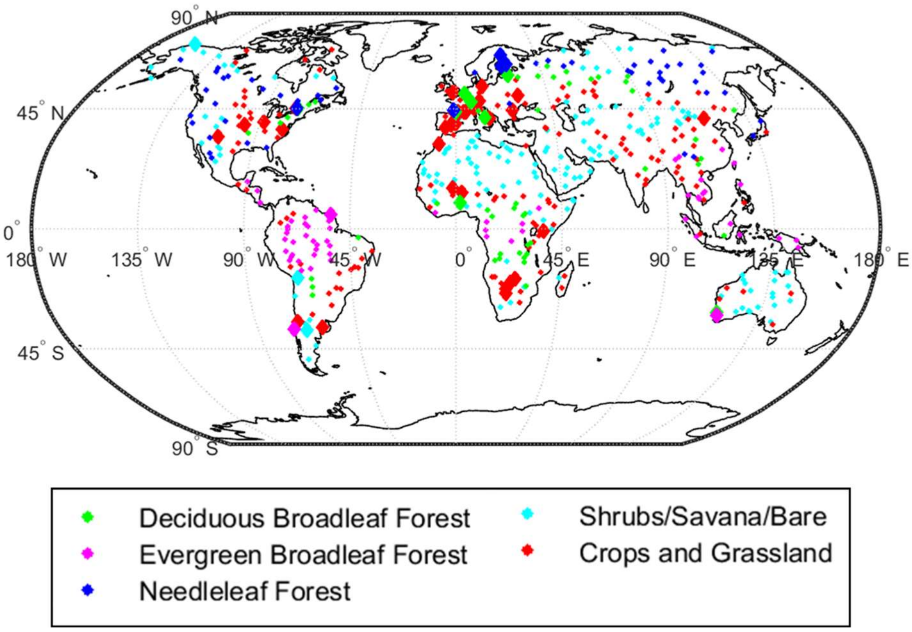

Figure 1. Location of BELMANIP2.1 (small symbols) and DIRECT2.0 (big symbols) sites. The different symbols correspond to the five biome classes as derived from the CCI landcover map (http://maps.elie.ucl.ac.be/CCI/viewer/download/ESACCI-LC-Ph2-PUGv2_2.0.pdf).

3 QUALITY ASSESSMENT OF GEOV2-AVHRR

3.1 SPATIAL CONSISTENCY AND PRODUCT COMPLETENESS

The GEOV2-AVHRR values and associated quality flags show consistent spatio-temporal patterns at the global scale (Figure 2). The seasonal patterns and winter-summer variations in GEOV2-AVHRR biophysical variables reflect the expected regimes of vegetation with low seasonality and very high values in evergreen broadleaf forests in the Amazonian, Congo and South-East Asia Basins; high seasonality and intermediate values over temperate forests and croplands in middle latitudes; and very limited seasonality and low values in the tundra artic areas as well as null seasonality and values close to zero in deserts. The number of valid observations and the length of the composition window also showed the expected seasonal pattern with less number of available observations and long temporal windows in winter time at higher latitudes to accommodate for these sub-optimal conditions during the snowy period. Tropical areas around the equator also show a small number of observations and long compositing windows, in this case, throughout the entire year because of very frequent cloud cover. Conversely, the sparsely vegetated areas and the permanent bare soil areas (e.g. desert) present a large number of observations and, consequently, short compositing windows due to the predominance of sunny days. The RMSE values show a similar seasonality as the associated biophysical variables. The Northern mid-latitudes show as expected a strong contrast between winter and summer seasons, with relatively high RMSE values in summer due to higher LAI values and residual cloud and atmospheric effects. Conversely, in winter, the RMSE decreases with the low LAI values. Note however that the RMSE is not computed over large areas in northern hemisphere in wintertime or in the tropics across the year because at least two valid observations are required. Our compositing method allows filling most of the gaps in the GEOV2-AVHRR time series, even in regions subjected to high occurrence of clouds as for tropical and equatorial regions, and in northern very high latitudes in winter snowy period with very high sun zenith angles where incoming radiation is very low and no valid observations are available in the maximum 120-day compositing window (cf. Figure 2a-b and Figure 3). EBFs, which are mostly located in areas with recurrent cloud cover around the equator, have the highest fraction of gap filled data (Figure 3): 70% of data values are gap filled using the climatology while for the other biome types this is typically lower than 10-15% of cases. This can be explained because either no available observations in the compositing period or because they were noisy and filtered out. Missing data for the period 1981-2018 represents 2% of the land pixels. It corresponds only to some pixels at very high northern latitudes (>70ºN) where GEOCLIM climatology (Verger et al. 2015) is not available (grey areas in Figure 2a-b and Figure 3).

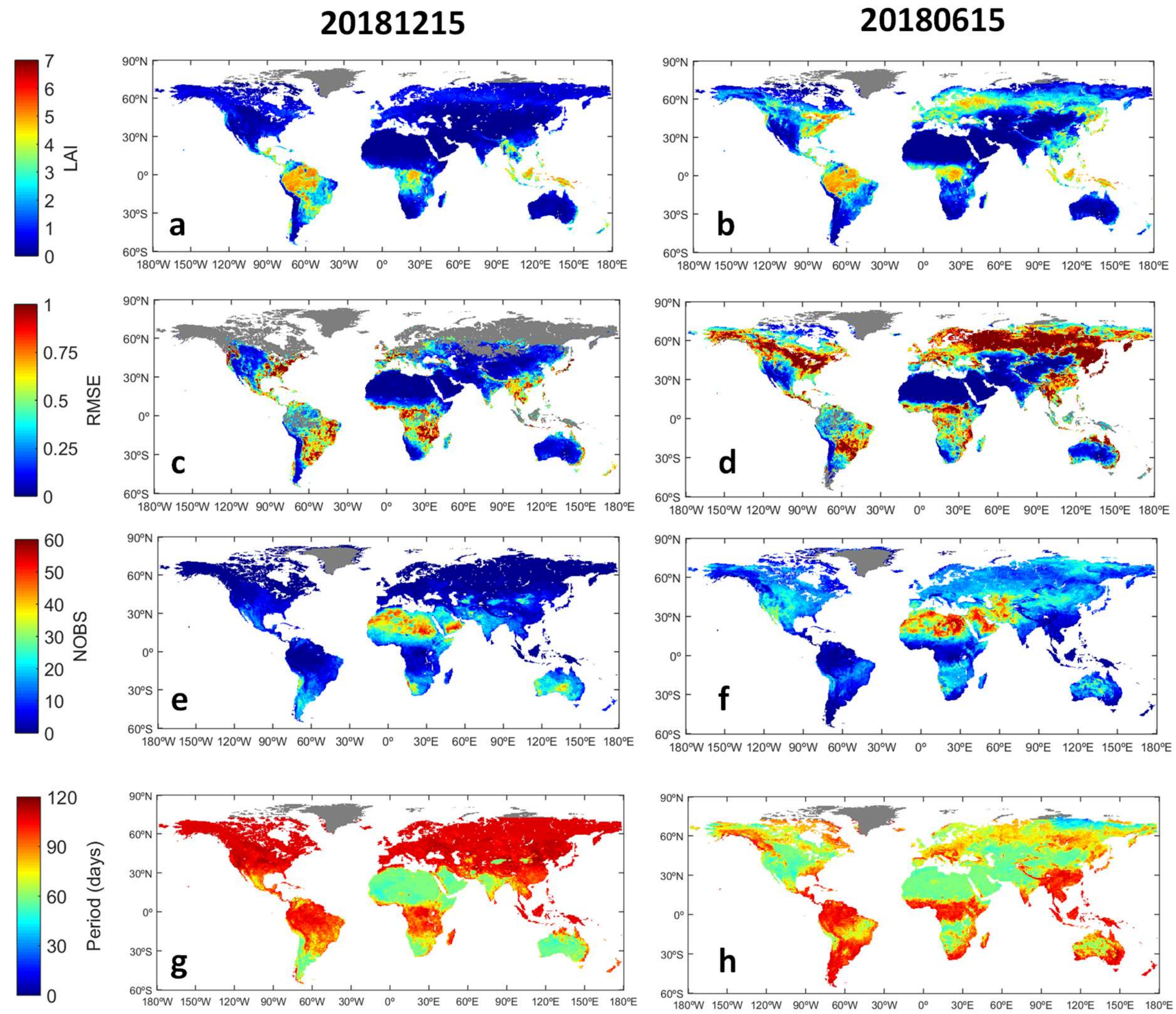

Figure 2. Maps of GEOV2-AVHRR LAI, RMSE LAI, number of valid observations (NOBS) and length of the composition window (period in days) for two dekads centered close to the dates of winter and summer solstice in the northern hemisphere: December 15 (left) and June 15 (right), 2018. Grey color indicates missing data.

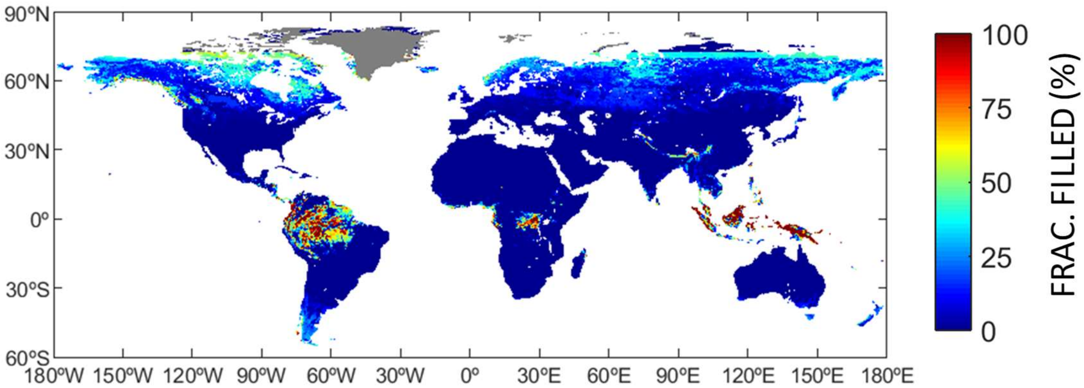

Figure 3. Map of the fraction of filled land pixels for GEOV2-AVHRR, i.e. the fraction of dekads with no valid daily estimates in the compositing window for which the climatology or the linear interpolation is used for filling gaps. The areas in grey correspond to unprocessed pixels (i.e. missing data).

3.2 COMPARISON WITH GEOV2-CGLS

GEOV2-AVHRR and GEOV2-CGLS products show a good overall agreement over the common 1999-2018 period: the GCOS uncertainty requirements (GCOS 2011) are well met with 94% of land pixels are within ±0.5 LAI differences between both products, and 85% of pixels are within ±0.05 FAPAR/FCover differences (Figure 4). The scatterplots of the comparison over the BELMANIP2.1 sites show RMSE of 0.48 LAI, 0.07 FAPAR and 0.08 FCover, with correlation coefficients higher than 0.97, slopes of the linear regressions close to the unity with no offset (Figure 5). However, more detailed inspection of the spatial distribution of the differences reveals lower GEOV2-AVHRR LAI values products in tropical latitudes over EBFs (Figure 4a, Figure 6). The EBFs are mostly located in areas characterized by a high occurrence of clouds which result in high level of noise in the time series due to the remaining atmospheric effects and cloud contamination. For fAPAR and FCover, a negative bias is also observed for some land pixels close to tropical latitudes with less valid observations due to permanent clouds (Figure 9b,c). The LAI is, among the three retrieved variables, the variable the most affected by noise in the data. EBFs are characterized by high levels of LAI and, in these conditions, the models are close to saturation and small variations in the reflectance level may result in high variations in LAI values. EBFs are characterized by lower near infrared reflectance level for a given LAI value mainly due to thicker leaves. To account these specific behaviors of EBFs, GEOV2-CGLS uses specific neural networks for processing EBF or non EBFs pixels while generic neural networks are used in GEOV2-AVHRR with non EBF class distinction.

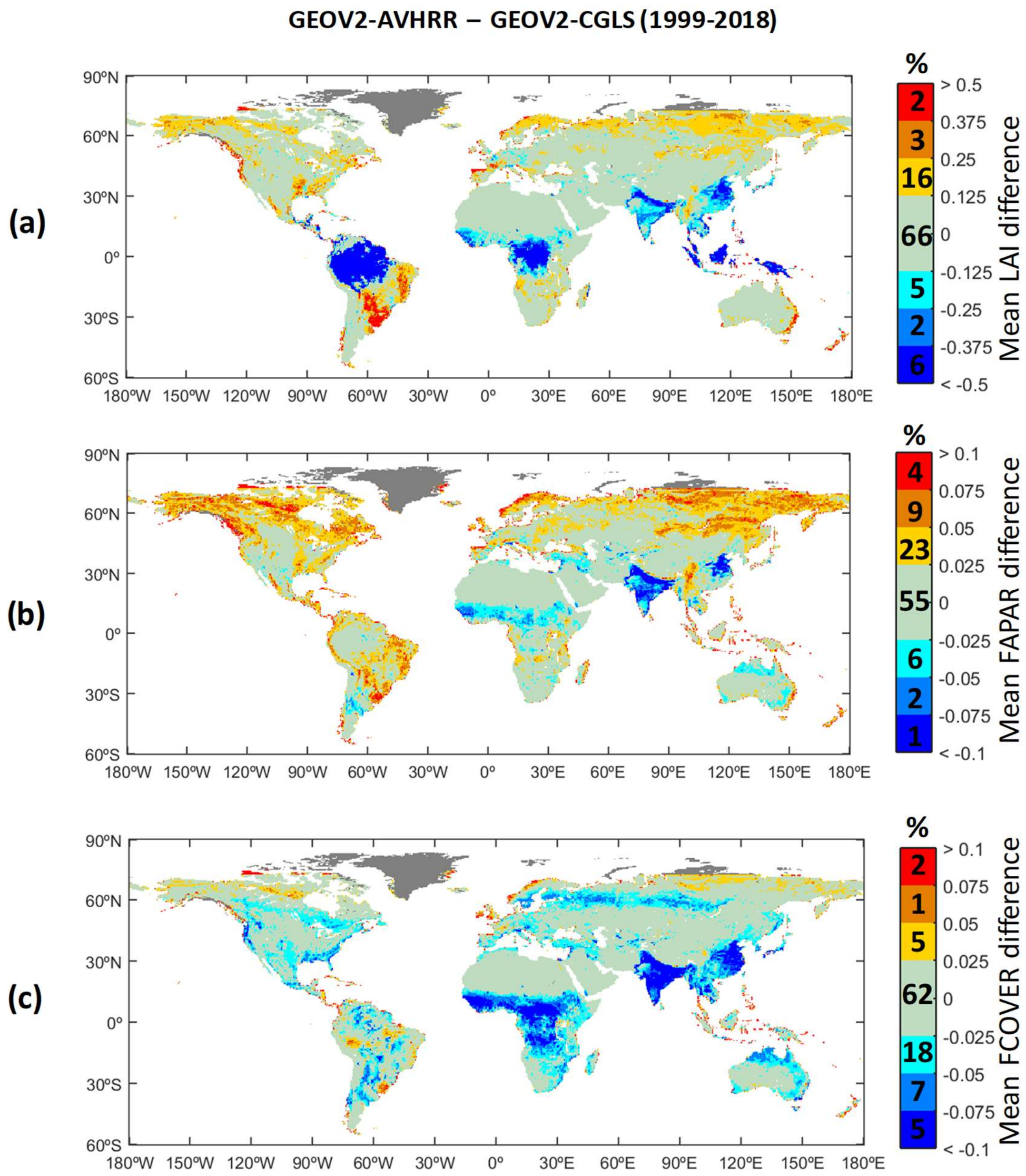

Figure 4. Maps of mean differences between GEOV2-AVHRR and GEOV2-CGLS (a) LAI, (b) FAPAR, and (c) FCover products for the 1999-2018 period. Red (blue) tones indicate higher (lower) GEOV2-AVHRR values than GEOV2-CGLS. The percentage of land pixels for each interval of differences is indicated in the colour bar. The areas in grey correspond to unprocessed pixels (i.e. missing data).

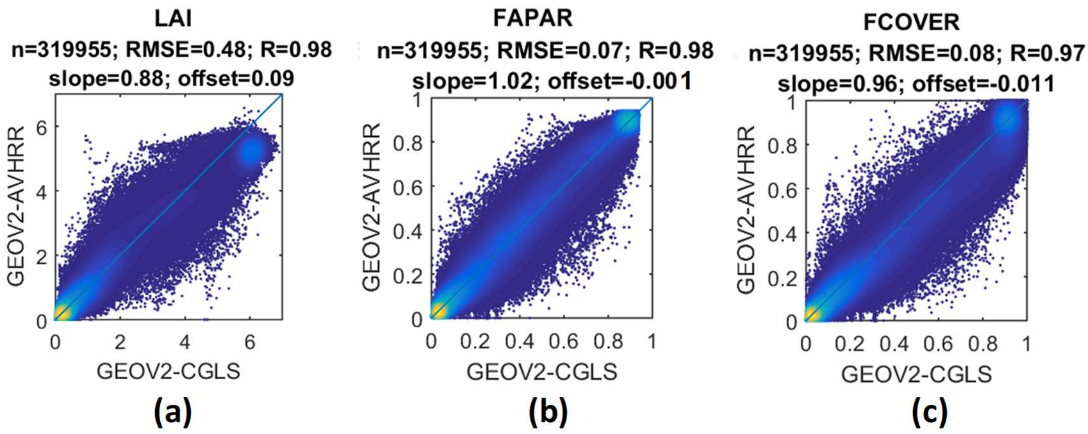

Figure 5. Comparison between GEOV2-AVHRR and GEOV2-CGLS (a) LAI, (b) FAPAR and (c) FCover products over the BELMANIP2 sites for the 1999-2018 period. The statistics of the comparison are indicated: number of samples (n), root mean square error (RMSE), correlation coefficient (R), slope and offset of the least square linear regression.

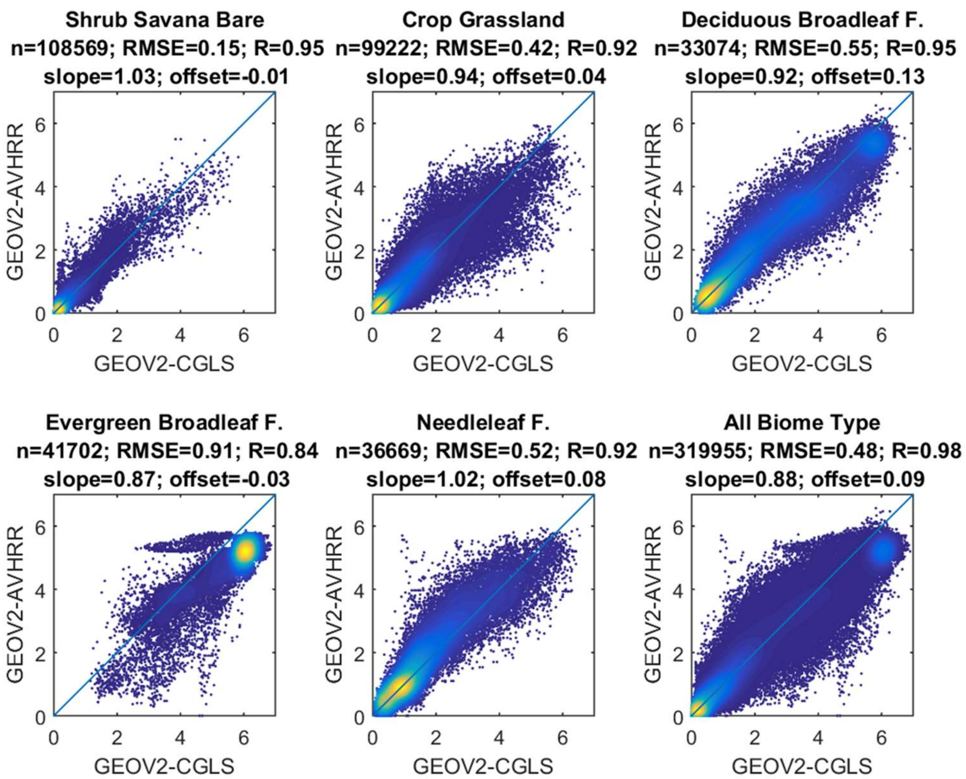

Figure 6. Comparison between GEOV2-AVHRR and GEOV2-CGLS LAI products per biome classes over the BELMANIP2 sites for the 1999-2018 period. The biome classes are derived from the CCI-LC (v1.6.1) landcover (https://www.esa-landcover-cci.org). The statistics of the comparison are indicated: number of samples (n), root mean square error (RMSE), correlation coefficient (R), slope and offset of the least square linear regression.

The LAI differences between GEOV2-AVHRR and GEOV2-CGLS increase with the LAI value due, in particular, to the higher discrepancies for EBFs (Figure 7a). The differences between GEOV2-AVHRR and GEOV2-CGLS LAI are closely linked with the RMSE associated to the GEOV2-AVHRR product (Figure 7b): the higher RMSE values, the higher LAI differences. Then the RMSE quality layer appears as a reliable proxy of the GEOV2-AVHRR product uncertainty. Note that the distribution of RMSE values (Figure 7b, dashed blue line) shows that most values are lower than 0.1. The agreement between GEOV2-AVHRR degrades when the number of available observations in the composition period decreases (NOBS lower than ~5 observations) (Figure 7c) and the period length increases (period higher than ~80 days) (Figure 7d). The distribution of NOBS (Figure 7c, dashed blue line) shows that, in most situations, NOBS is in between 10 to 20, and the period length is lower than 70 days (Figure 7e).

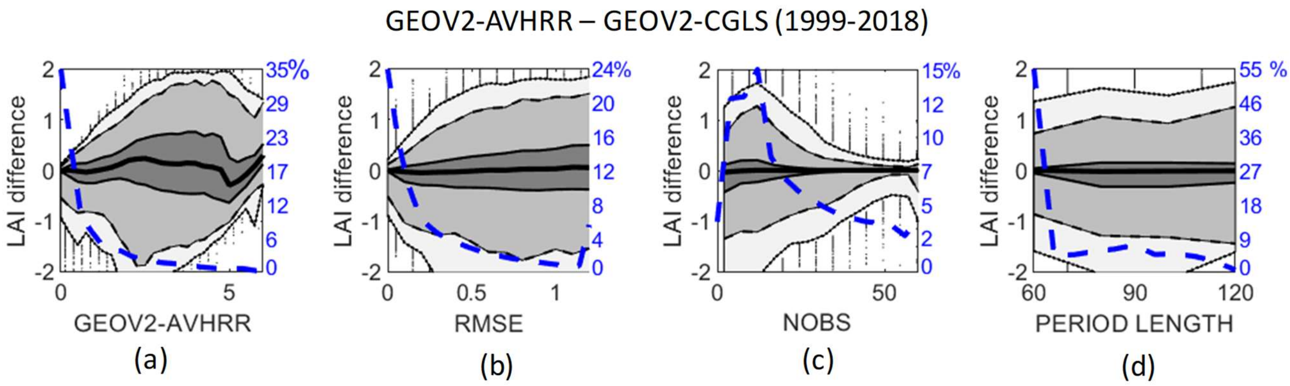

Figure 7. Evaluation of the differences between GEOV2-AVHRR and GEOV2-CGLS LAI products over the BELMANIP2.1 sites for the 1999-2018 period as a function of the GEOV2-AVHRR LAI value (a), the RMSE between the GEOV2-AVHRR product and the daily estimates (b), the number of valid daily estimates in the composition period of GEOV2-AVHRR: NOBS (c), and the length of the composition period. The several gray values correspond to 75% (dark gray), 90% (medium gray) and 95% (light gray) of the population, and the dots to 5th percentile of residual outliers. The bold black solid line corresponds to the median value of the differences. The dashed blue line shows the distribution of values of the variable in the abscissa which frequencies are indicated in the vertical axis on the right.

3.3 Comparison with GIMMS3G and GLASSV4

The comparison of GEOV2-AVHRR with GIMMS3g and GLASSV4 LAI products over the overlapping period (1981-2011) shows 94% and 92% of land pixels, respectively, meet GCOS requirements of ±0.5 LAI differences (Figure 8). The highest differences between LAI products are observed over northern high latitudes where GEOV2-AVHRR shows higher mean LAI values than GIMMS3g and GLASSV4 (red tones in Figure 8a-b), and over Evergreen Broadleaf Forests at latitudes near the Equator where GEOV2-AVHRR shows systematic higher values than GIMMS3g (red tones in Figure 2a) and a good agreement with GLASSV4 (green tones in Figure 2b).

For FAPAR, the comparison of GEOV2-AVHRR with GIMMS3g and GLASSV4 products shows 62% and 65% of pixels, respectively, are within ±0.05 FAPAR requirements (Figure 9). The highest differences with GLASSV4 are again observed over latitudes higher than 50°N where GEOV2-AVHRR shows higher FAPAR values (red tones in Figure 9b). GIMMS3g FAPAR, when available, is systematically higher than the other products for areas with FAPAR<0.2 (blue tones in Figure 9a).

The comparison between GEOV2-AVHRR and GLASSV4 FCover products shows that 56% of land pixels exhibit mean differences within ±0.05 FCover (Figure 10). For northern high latitudes GLASSV4 is systematically higher than GEOV2-AVHRR. Note that in these regions the sign of the FCover differences between GEOV2-AVHRR and GLASSV4 is the opposite as the one observed for FAPAR (c.f. blue tones in Figure 10 corresponds to red tones in Figure 9b) which reveals an internal inconsistency in GLASSV4 dataset.

GEOV2-AVHRR shows a very consistent FAPAR-FCover relationship with FCover<FAPAR/0.94 (Figure 11a). Indeed, when normalized by its maximum value, FAPAR can be approximated by the FIPAR (Fraction of Intercepted Photosynthetically Active Radiation), which is the complementary of the gap fraction at the sun angle corresponding to 10:00 am solar time for a given pixel, and thus, leading to a higher value than FCover defined as the complementary to the gap fraction in the nadir direction. Since instantaneous FAPAR depends on the incident radiation which is a function of the sun position and atmospheric conditions, the GEOV2-AVHRR FCover – FAPAR/0.94 differences gradually increase with the sun zenith angle (Figure 11a). Conversely, GLASSV4 shows more noise and higher FCover – FAPAR/0.94 differences (see distribution of differences in Figure 11b) and unexpected FCover>FAPAR/0.94 particularly for high zenith angle (SZA) values, SZA>70º (Figure 11b). These inconsistencies between GLASSV4 variables are due, in part, to the different retrieval algorithms used for FAPAR (Xiao et al. 2015) and FCover (Jia et al. 2019) as described in Section 2.1. They may also indicate a problem in GLASSV4 FAPAR and/or FCover products for extreme illumination conditions. This may explain the high differences with GEOV2-AVHRR product and the observed biases in GLASSV4 FAPAR (Figure 9b) and FCover (Figure 10) at high latitudes where the SZA values are typically the highest, particularly in winter time.

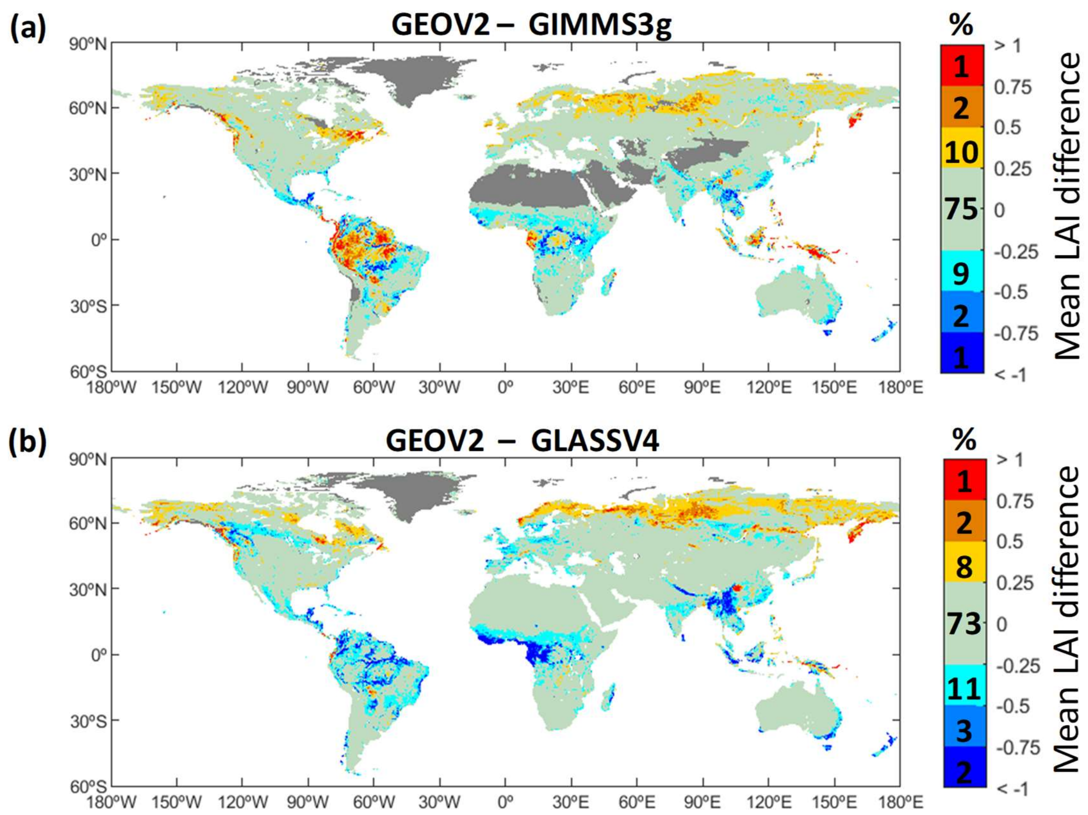

Figure 8. Mean LAI differences between (a) GEOV2-AVHRR and GIMMS3g, and (b) GEOV2-AVHRR and GLASSV4 products for the common period (1981-2011). Red (blue) tones indicate higher (lower) GEOV2-AVHRR values than GIMMS3g (a) and GLASSV4 (b). The percentage of land pixels for each interval of differences is indicated in the colour bar. The areas in grey correspond to unprocessed pixels (i.e. missing data).

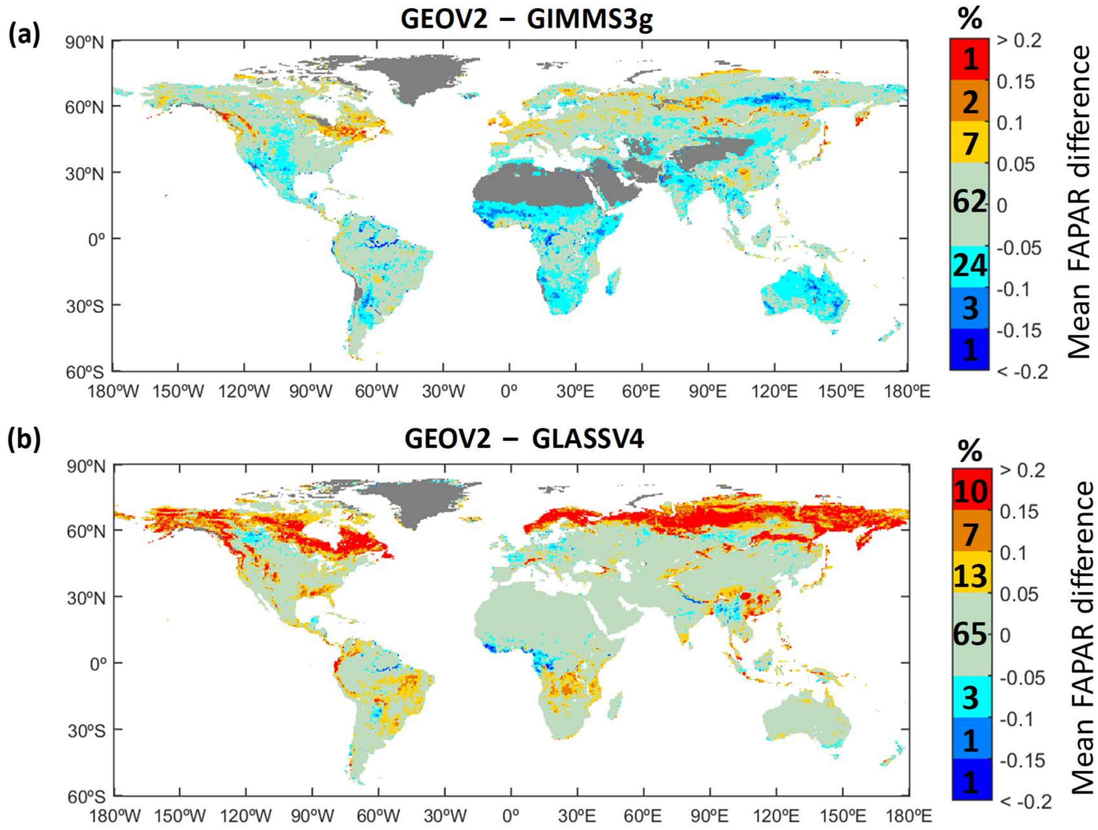

Figure 9. Mean FAPAR differences between (a) GEOV2-AVHRR and GIMMS3g, and (b) GEOV2-AVHRR and GLASSV4 products for the common period (1982-2011). Red (blue) tones indicate higher (lower) GEOV2-AVHRR values than GIMMS3g (a) and GLASSV4 (b). The percentage of land pixels for each interval of differences is indicated in the colour bar. The areas in grey correspond to unprocessed pixels (i.e. missing data).

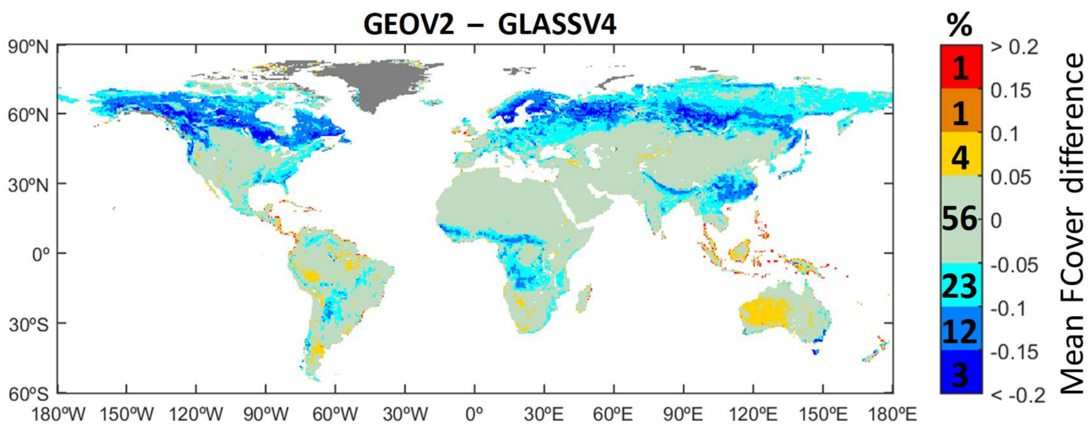

Figure 10. Mean FCover differences between GEOV2-AVHRR and GLASSV4 for the common period (1982-2018). Red (blue) tones indicate higher (lower) GEOV2-AVHRR values than GLASSV4. The percentage of land pixels for each interval of differences is indicated in the colour bar. The areas in grey correspond to unprocessed pixels (i.e. missing data).

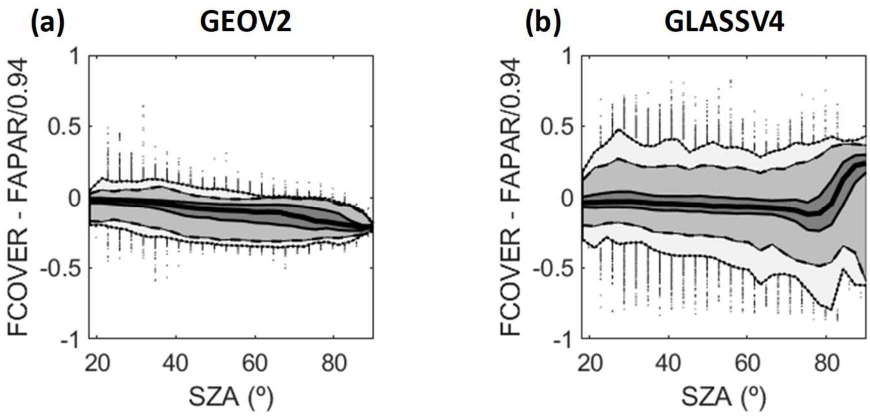

Figure 11. Evaluation of the differences between FCover and FAPAR/0.94 as a function of the sun zenith angle (SZA) at 10:00 am for (a) GEOV2-AVHRR and (b) GLASSV4 products over the LANDVAL sites for the years 1982-2018. The gray levels correspond to 75% (dark gray), 90% (medium gray) and 95% (light gray) of the population, and the dots to 5% percentile of the residual outliers. The bold back solid line corresponds to the median value of the differences.

The scatterplots of the comparison between GEOV2-AVHRR, GIMMS and GLASS LAI products (Figure 12) show RMSE<0.8, correlation coefficient R>0.9, and slope of the linear regression close to the unity with offset close to zero (Table 2). The distributions of LAI values per biomes (Figure 15) are very consistent between GEOV2-AVHRR, GIMMS3g and GLASS except for GIMMS3g Evergreen Broadleaf Forests (EBFs). The observed discrepancies in EBFs are due, in part, to the higher uncertainties in the data (permanent cloud cover, saturation of the signal) but also to the intrinsic limitations and differences of the retrieval algorithms (differences in the training process, filtering, smoothing and gap filling). GEOV2-AVHRR, GLASS and GIMMS3g EBF LAI distributions agree for the location of the peak: maximum frequencies at LAI=5 (Figure 15). GIMMS3g shows a broad LAI frequency distribution for EBFs with lower values than GEOV2-AVHRR and GLASS which are associated to the noise and instabilities observed in the temporal profiles of GIMMS3g (Figure 18). GEOV2-AVHRR and GLASS show sharp frequency distributions of LAI values mainly because of the more efficient filtering of cloud contamination and the algorithm assumptions forcing a very limited seasonality for EBFs. The LAI distributions over needleleaf forest show very low frequency value at LAI=0 for GEOV2-AVHRR as expected while GIMMS3g and GLASS shows higher frequency values which are inconsistently higher than the frequency values at LAI=0 over deciduous broadleaf forest (Figure 15). The statistics of the inter-comparison of LAI products (Table 2) indicates GEOV2-AVHRR generally constitutes an intermediate solution between GIMMS3g and GLASS: the differences of GEOV2-AVHRR with either GLASS (RMSE=0.64) or GIMMS3g (RMSE=0.70) are lower than the differences between GLASS and GIMMS3g (RMSE=0.81).

The scatter plots show very strong agreement for GEOV2-AVHRR as compared to GLASS for all the range of FAPAR values (Figure 13b) and biomes (Figure 16). Note however that the GLASS frequencies at FAPAR=0 are inconsistently higher for needleleaf forests than for deciduous broadleaf forests (Figure 16). GIMMS3g FAPAR is systematically higher than GEOV2-AVHRR

(Figure 13a) and GLASS (Figure 13c) for FAPAR<0.2 across biomes (Figure 16). This positive bias seems to be associated to a problem in the training dataset used to calibrate the GIMMS3g algorithm, in particular the MODIS products used for learning neural networks (Zhu et al. 2013). A similar offset for the low FAPAR values was observed in the validation of MODIS product (Martínez et al. 2013; McCallum et al. 2010). Note that MODIS FAPAR is also used for training GEOV2-AVHRR algorithm but weighted by its uncertainty (Verger et al. 2021-This issue). GEOV2-AVHRR typically provides intermediate FAPAR values between GIMMS3g and GLASS (Table 2): better agreement of GEOV2-AVHRR with either GLASS (RMSE=0.12) or GIMMS (RMSE=0.10) than between GLASS and GIMMS (RMSE=0.13).

The comparison between GEOV2-AVHRR and GLASS shows an overall good agreement: RMSE of 0.09 and >60% of samples within 10% or 0.05 FCover differences (Table 2, Figure 14), but GLASS shows higher intermediate FCover values (Figure 14). The higher differences between GEOV2-AHRR and GLASS FCover are observed for deciduous and needleleaf forests (Figure 17c,e) where GLASS systematically shows higher GLASS values (blue tones in Figure 10 for northern high latitudes). GLASS showed unexpected multimodal FCover distributions for needleleaf forests (Figure 17e).

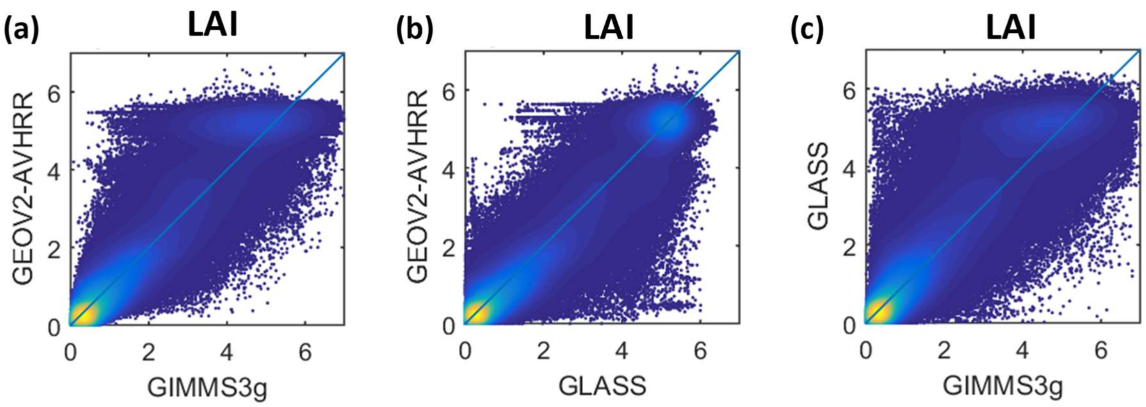

Figure 12. Comparison between (a) GEOV2-AVHRR and GIMMS3g, (b) GEOV2-AVHRR and GLASS, and (c) GLASS and GIMMS3g LAI products over the BELMANIP2 sites for the common samples in the overlapping period 1981-2011. The color represents the point density, from yellow (highest) to blue (lowest density). The statistics of the comparisons are reported in Table 2.

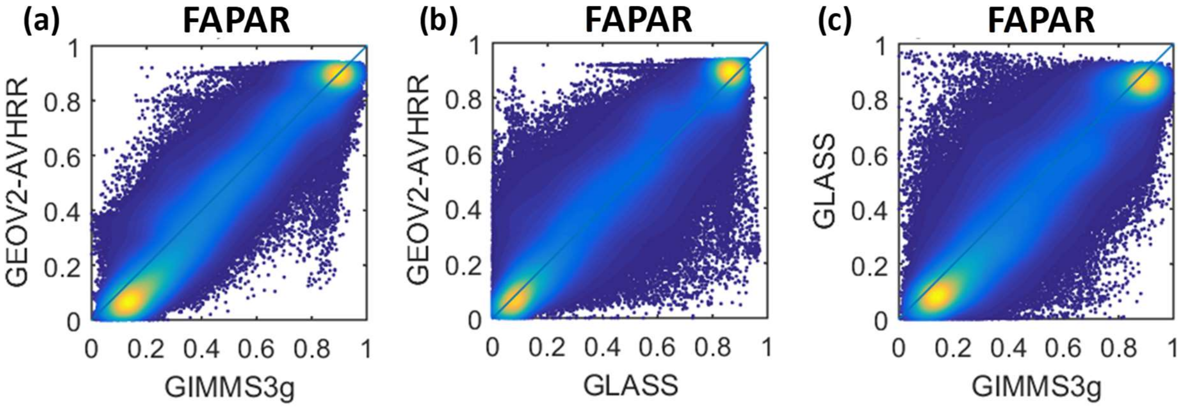

Figure 13. Comparison between (a) GEOV2-AVHRR and GIMMS3g, (b) GEOV2-AVHRR and GLASS, and (c) GLASS and GIMMS3g FAPAR products over the BELMANIP2 sites for the common samples in the overlapping period 1982-2011. The color represents the point density, from yellow (highest) to blue (lowest density). The statistics of the comparisons are reported in Table 2.

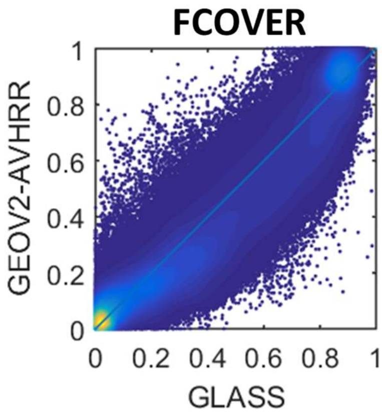

Figure 14. Comparison between (a) GEOV2-AVHRR and GIMMS3g, (b) GEOV2-AVHRR and GLASS, and (c) GLASS and GIMMS3g LAI products over the BELMANIP2 sites for the common samples in the overlapping period 1982-2018. The color represents the point density, from yellow (highest) to blue (lowest density). The statistics of the comparisons are reported in Table 2.

Table 2. Statistics of the intercomparison of GEOV2-AVHRR, GEOV2-CGLS, GIMMS3g and GLASS LAI, FAPAR and FCover variables over the BELMANIP2 sites for the common samples in the overlapping period (1981-2011 for LAI, 1982-2011 for FAPAR, and 1982-2018 for FCover): *n* is number of samples (sites x dates), *%OK GCOS* refers to the 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), mean bias, correlation coefficient (R), slope and offset of the least squares linear regression.

		%OK GCOS	RMSE	Bias	R	slope	offset
LAI (n=376257)	GEOV2-AVHRR – GIMMS	75	0.70	0.04	0.92	1.08	-0.11
	GEOV2-AVHRR – GLASS	77	0.64	-0.10	0.94	1.003	-0.10
	GLASS – GIMMS3g	70	0.81	0.14	0.90	1.09	-0.01
FAPAR (n=370183)	GEOV2-AVHRR – GIMMS	46	0.10	-0.02	0.95	1.10	-0.07
	GEOV2-AVHRR – GLASS	55	0.12	0.01	0.92	1.04	-0.003
	GLASS – GIMMS3g	42	0.13	-0.03	0.91	1.06	-0.06
FCover (n=592736)	GEOV2-AVHRR – GLASS	61	0.09	-0.03	0.96	0.97	-0.01

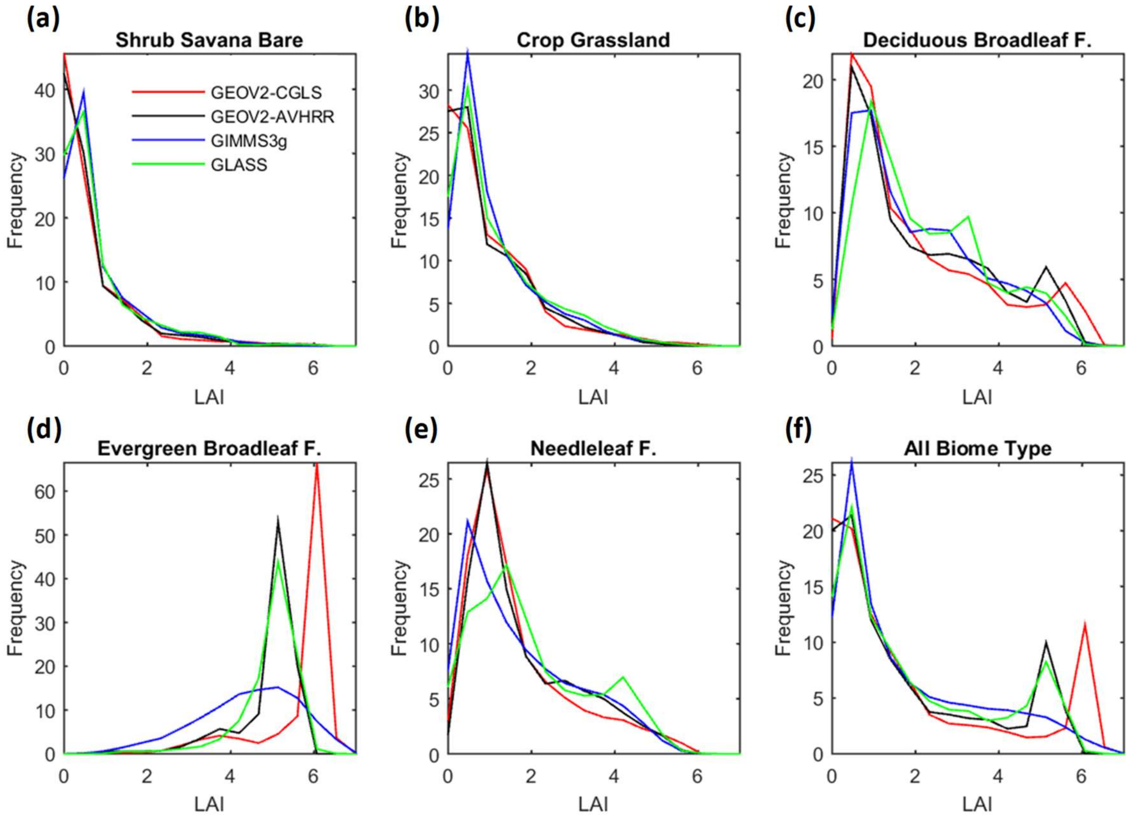

Figure 15. Distribution of GEOV2-AVHRR, GEOV2-CGLS, GIMMS3g and GLASSV4 LAI products per biome type as sampled by the BELMANIP2 sites in the overlapping period 1981-2011. Biome distinction is based on CCI-LC (v1.6.1) landcover (https://www.esa-landcover-cci.org).

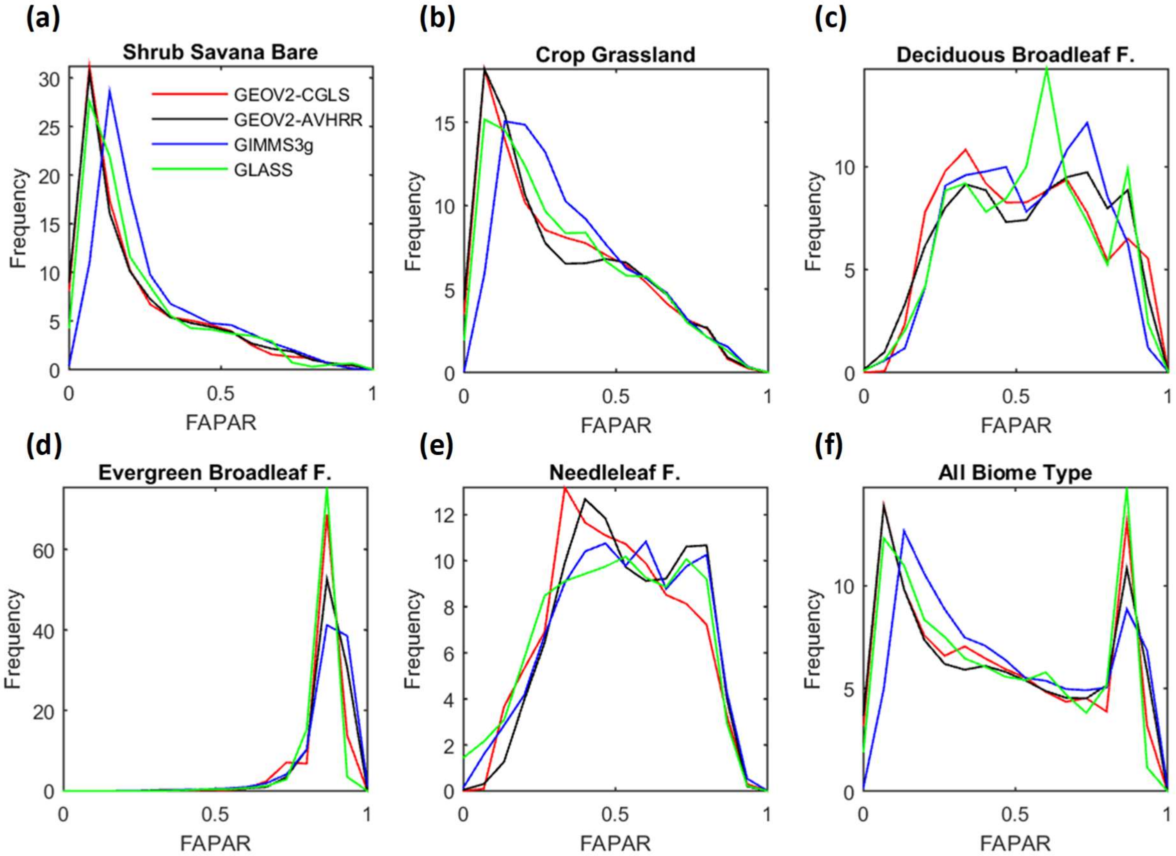

Figure 16. Distribution of GEOV2-AVHRR, GEOV2-CGLS, GIMMS3g and GLASSV4 FAPAR products per biome type as sampled by the BELMANIP2.1 sites in the overlapping period 1982-2011. Biome distinction is based on CCI-LC (v1.6.1) landcover (https://www.esa-landcover-cci.org).

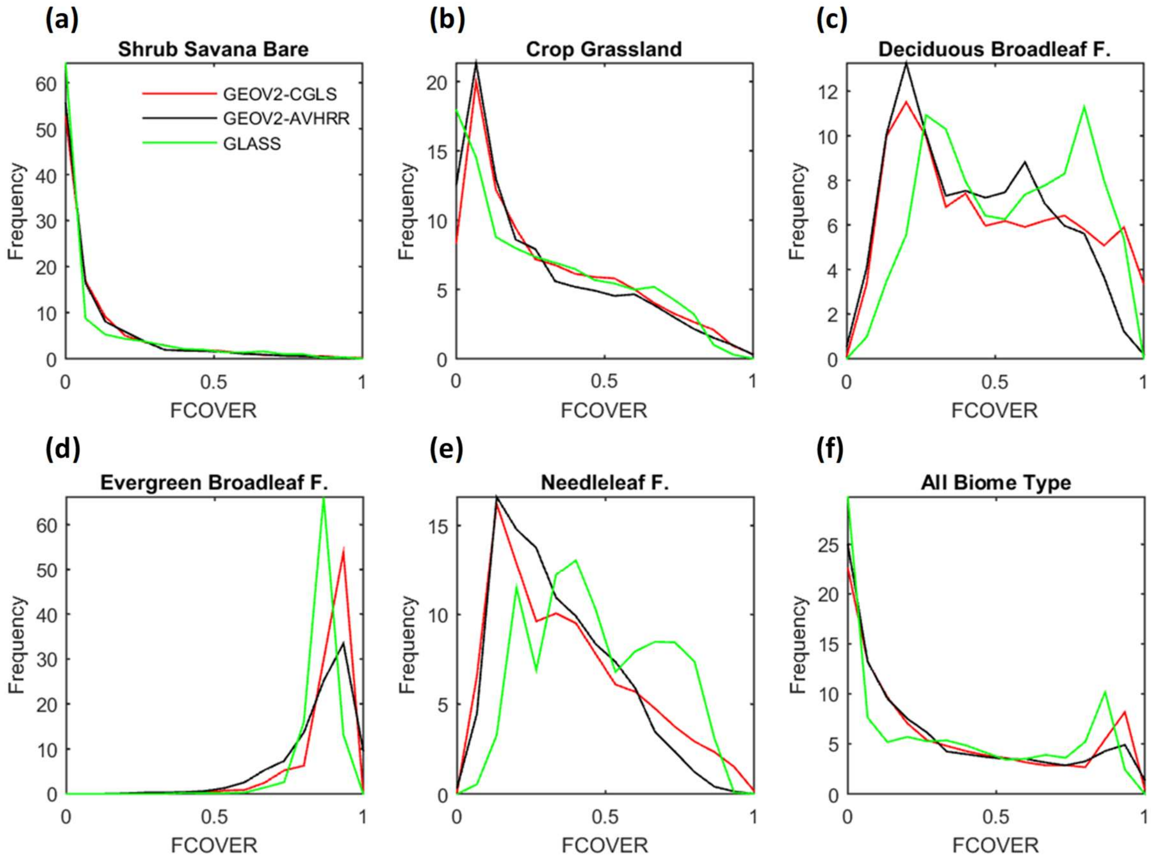

Figure 17. Distribution of GEOV2-AVHRR, GEOV2-CGLS and GLASSV4 FCover products per biome type as sampled by the BELMANIP2.1 sites over the 1982-2018 period. Biome distinction is based on CCI-LC (v1.6.1) landcover (https://www.esa-landcover-cci.org).

3.4 TEMPORAL CONSISTENCY

The analysis is organized per large biome types, selecting few sites showing typical temporal profiles of LAI (Figure 18-Figure 21). For Counami EBF (Figure 18), the effect of residual clouds is very pronounced, creating strongly negatively biased LAI estimates and a shaky temporal profile for GIMMS3g products with abrupt spikes and dips. These inconsistencies are efficiently filtered in GEOV2-AVHRR and GEOV2-CGLS products that for EBFs apply a frequency criterion that rejects values lower than the percentile-80 computed over the entire pixel-based time series [THEIA_ATBD_GEOV2_AVHRR I2.30]. This results in temporal profiles with a very limited seasonality and high level of LAI, FAPAR and FCover (Figure 18) values. GEOV2-CGLS slightly provides higher LAI and FCover values than GEOV2-AVHRR that better agrees with available ground measurements. Similarly, the GLASSV4 products are also characterized by a limited seasonality and high values that agree with GEOV2-AVHRR and ground data (Figure 18).

GEOV2-AVHRR, GEOV2-CGLS, GLASSV4 and GIMMS3g show a similar temporal evolution at Järvselja mixed forest site and also agree with available ground measurements (Figure 19). The highest differences are observed in the winter period: GEOV2-AVHRR and GEOV2-CGLS values highly agree in winter time but they systematically shows higher LAI and FAPAR values than GIMMS3g; and, GLASSV4 shows intermediate LAI and FAPAR values between GEOV2-AVHRR and GIMMS3g and an unexpected inter-annual variability in winter values. In summer time, GIMMS3g product shows an anomalous seasonality with unexpected LAI and FAPAR drops due to unfiltered clouds while GEOV2-AVHRR and GLASSV4 are more robust to the noise in the data and they show more temporally consistent and smooth time evolutions. For FCover, GEOV2-AVHRR and GLASSV4 highly agree in terms of seasonality but GLASSV4 FCover is systematically higher than GEOV2-AVHRR and it also overestimates ground measurements. GLASSV4 FCover show more intra-annual stability in winter FCover values than GLASSV4 LAI and FAPAR with only some spikes for years 1991, 1993 and 1997. Note that the minimum GLASSV4 FCover values in winter time are higher than those of GEOV2-AVHRR FCover and they also are inconsistently higher than GLASSV4 FAPAR/0.94 values for most of years.

For Pshenichne wheat crop site, GEOV2-AVHRR, GEOV2-CGLS and GIMMS3g LAI and FAPAR products agree both in terms of intra-annual and inter-annual temporal evolution as well as in the magnitude of LAI and FAPAR estimates which also agree with available ground measurements (Figure 20). GLASSV4 LAI highly disagrees in terms of phenology as compared to the other products for some particular years and it shows temporal inconsistencies and apparent limitations to correctly reproduce the crop seasonality and the rapid changes in vegetation growth in the periods 1986-1999 and 2007-2010: GLASSV4 LAI shows high LAI values in autumn period after wheat harvest when a minimum in vegetation amount is expected as well as higher LAI values than the other satellite products in the peak of the growing season (Figure 20). This may be due to the rolling approach used to generate continuous LAI time series based on one-year data (Xiao et al. 2016) or to the preprocessing applied to AVHRR reflectances (Tang et al. 2013) in GLASSV4 LAI/FAPAR. The TSCD algorithm (Tang et al. 2013) for filtering residual clouds in AVHRR reflectances assumes land surface reflectance is relatively stable or slowly changing and rapid temporal variations in reflectance are due to cloud contamination. Note the GLASSV4 FAPAR shows the same temporal pattern and limitations since it is derived from GLASSV4 LAI (cf. section 2.1.4). On the contrary, GLASSV4 FCover, which is based on a different retrieval algorithm and preprocessing, well reproduce the vegetation growth and crop cycle and it shows a similar seasonality as GEOV2-AVHRR.

GEOV2-AVHRR, GEOV2-CGLS, GLASSV4 and GIMMS3g LAI and FAPAR products correctly reproduce the phenology at Tessekere savanna site with short growing seasons and a rapid growth of the herbaceous vegetation in the wet-period (Figure 21). Systematic differences are observed for the low LAI and FAPAR values in the dry season: GIMMS3g and to a lesser extend GLASSV4 systematically overestimates low LAI and FAPAR values as compared to GEOV2-AVHRR, GEOV2-CGLS and ground measurements. The temporal trajectories and level of FCover of GEOV2-AVHRR, GEOV2-CGLS and GLASSV4 highly agree.

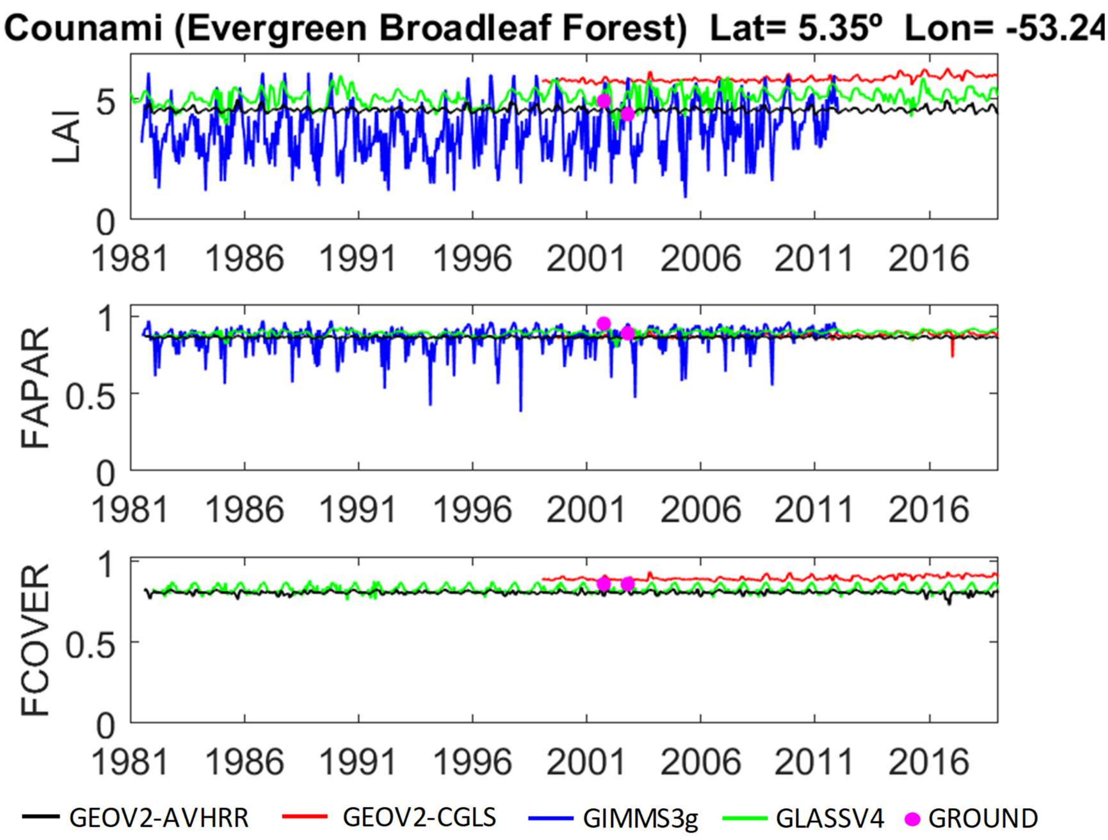

Figure 18. Temporal profiles of GEOV2-AVHRR, GEOV2-CGLS, GIMMS3g and GLASS LAI (top), FAPAR (middle) and FCover (bottom) products over the tropical Counami evergreen broadleaf forest site for the period 1981-2018. The points correspond to the available ground measurements.

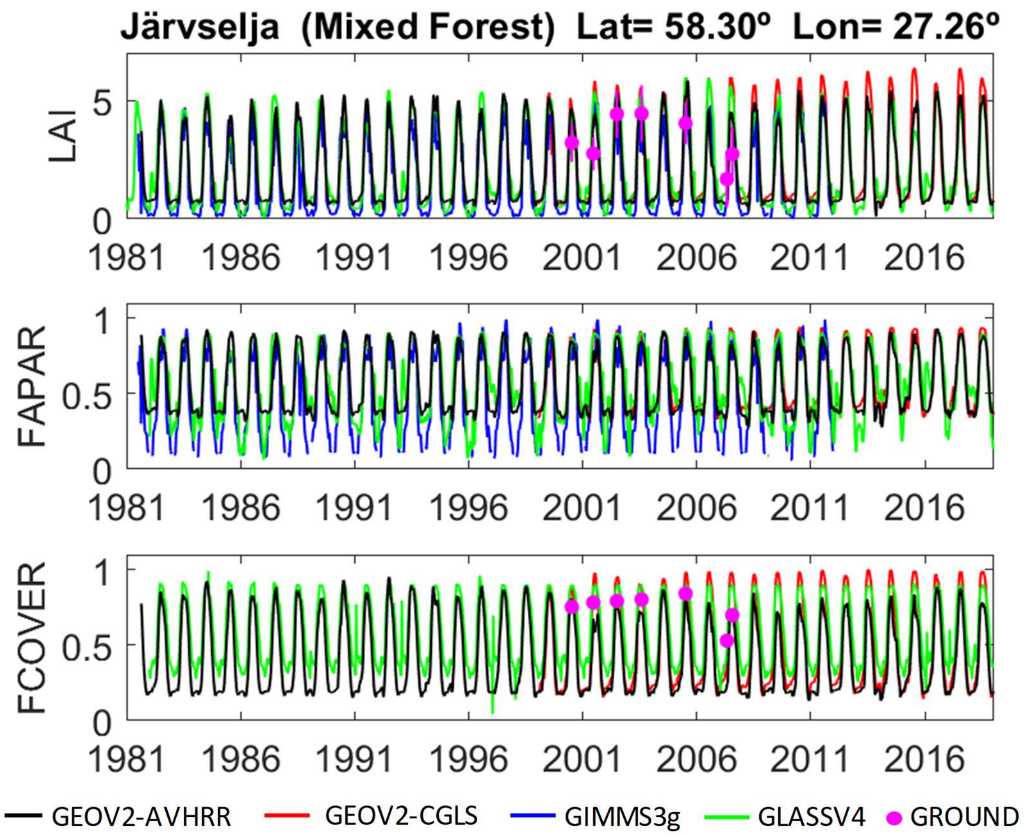

Figure 19. Temporal profiles of GEOV2-AVHRR, GEOV2-CGLS, GIMMS3g and GLASS LAI (top), FAPAR (middle) and FCover (bottom) products over the Järvselja mixed forest site at high northern latitude for the period 1981-2018. The points correspond to the available ground measurements.

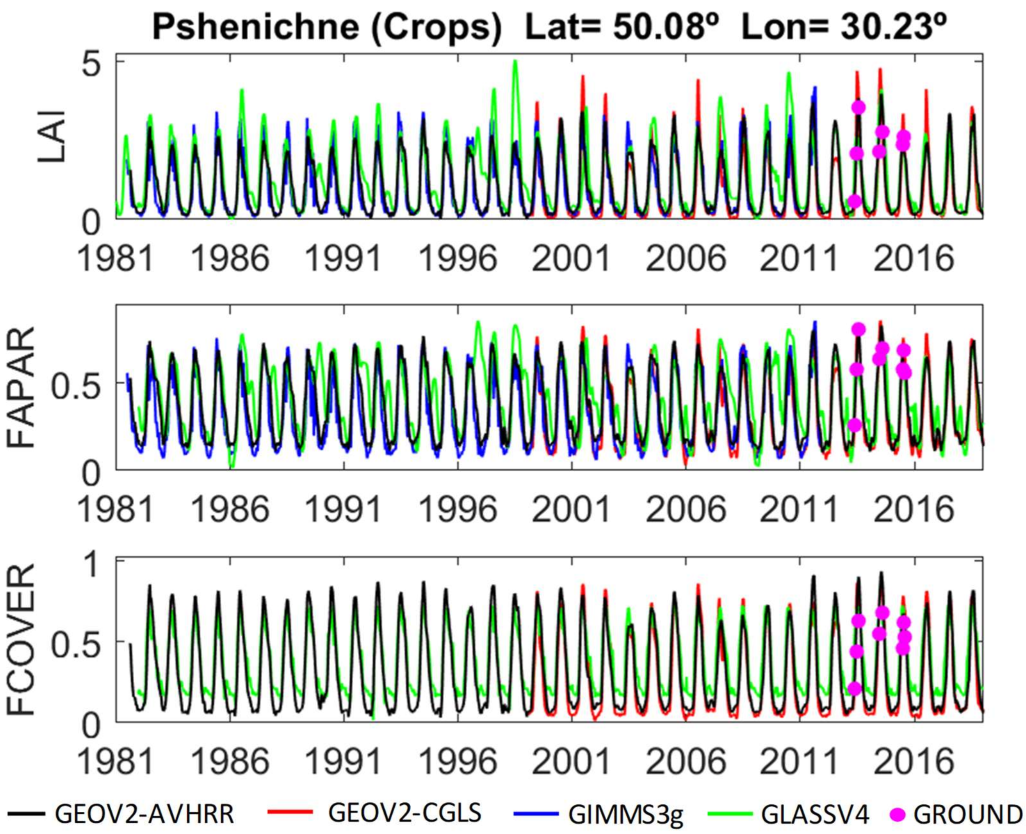

Figure 20. Temporal profiles of GEOV2-AVHRR, GEOV2-CGLS, GIMMS3g and GLASS LAI (top), FAPAR (middle) and FCover (bottom) products over the Pshenichne wheat crop site for the period 1981-2018. The points correspond to the available ground measurements.

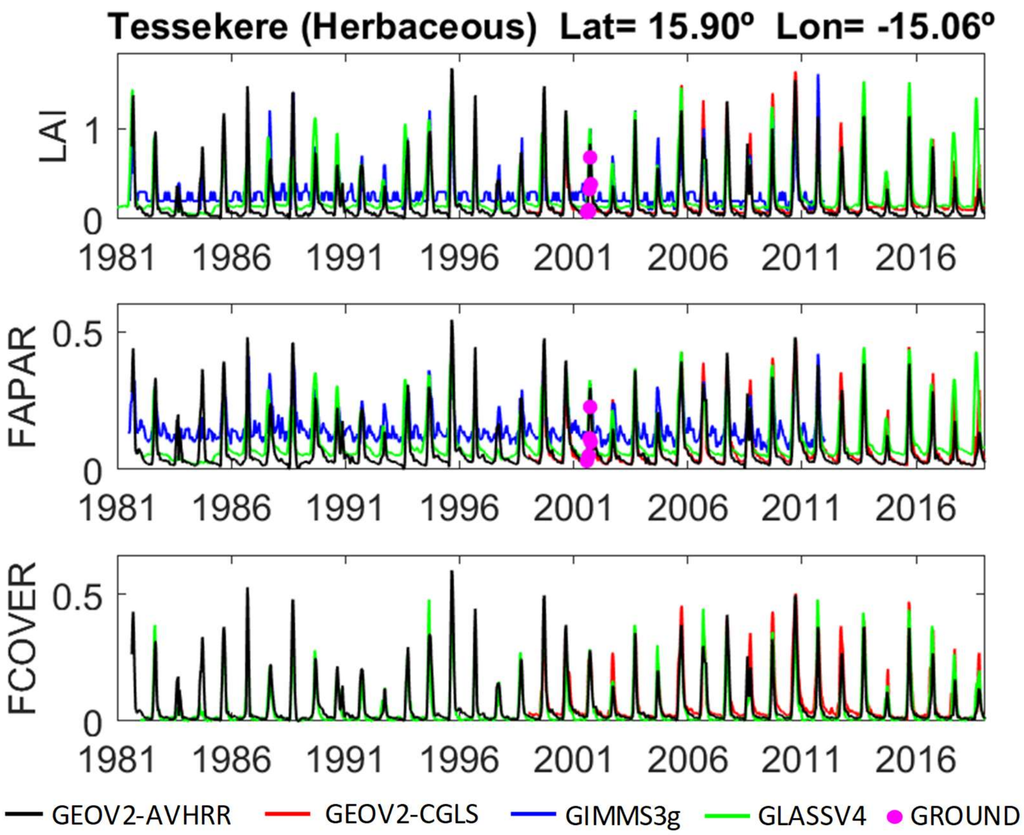

Figure 21. Temporal profiles of GEOV2-AVHRR, GEOV2-CGLS, GIMMS3g and GLASS LAI (top), FAPAR (middle) and FCover (bottom) products over the Tessekere herbaceous savana site for the period 1981-2018. The points correspond to the available ground measurements.

The histogram of δLAI (Figure 22) shows that both GEOV2-AVHRR, GEOV2-CGLS and GLASSV4 products are very smooth with differences lower than 0.1 LAI for most of cases. The GLASSV4 product shows the smoothest temporal evolution at expenses of miss-reproducing rapid variations in the temporal evolution of LAI. On the opposite, GIMMS3g shows the higher frequencies for higher δLAI values indicating a shakier temporal evolution as illustrated in Figure 18. GEOV2-AVHRR and GEOV2-CGLS provide almost identical performances and they show intermediate temporal smoothness as compared to GIMMS3g and GLASSV4.

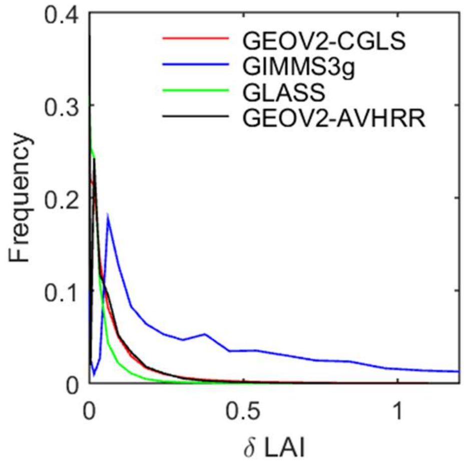

Figure 22. Histogram of the δLAI absolute difference representing temporal smoothness for GEOV2-AVHRR, GEOV2-CGLS, GIMMS3g and GLASSV4 LAI products. Evaluation over the BELMANIP2 sites for the common period 1999-2011.

GEOV2-AVHRR time series show high temporal consistency with no artifacts from the transition between AVHRR sensors except at the beginning of NOAA-11 operating period (start date: 1988/11/08) (Figure 23). In the period from November 1988 to October 1989, the negative anomalies for GEOV2-AVHRR and GLASSV4 can be explained by sub-optimal quality of the AVHRR reflectances in LTDRv4 dataset used as input data of both products. Tian et al. (2015) already observed lower NDVI values in the LTDR AVHRR NDVI products during this period. Negative anomalies are observed in the satellite datasets at tropical latitudes during El Chichon (04/1982–12/1984) and Mount Pinatubo (06/1991–12/1993) volcanic stratospheric aerosol periods. The latitudinal and temporal pattern of the standardized anomalies for the overlapping period 1982-2011 is very consistent between GEOV2-AVHRR, GIMMS3g and GLASSV4 LAI products except the later shows positive anomalies since the year 2000 which are not observed in GEOV2-AVHRR and GIMMS3g. The same positive anomalies since year 2000 are observed in GLASSV4 FAPAR and FCover products (not shown for brevity). These positive anomalies in GLASSV4 LAI can be partially explained by the shift from AVHRR-2 (1981-2000) to AVHRR-3 (2000-2018) sensor. The LTDRv4 AVHRR-3 data from NOAA 16,17,18 and 19, systematically underestimated reflectance values from year 2000 and onward due to a calibration issue (https://landweb.modaps.eosdis.nasa.gov/cgi-bin/ltdr/QA_WWW/displayCase1.cgi?esdt=AVH02C1&caseNum=PM_AVH02C1_17038&caseLocation=cases_data&sensor=NOAA&ver=C005). Conversely, the spectral harmonization applied in GEOV2-AVHRR algorithm [THEIA_ATBD_GEOV2_AVHRR I2.30] effectively corrects the radiometric differences between sensors.

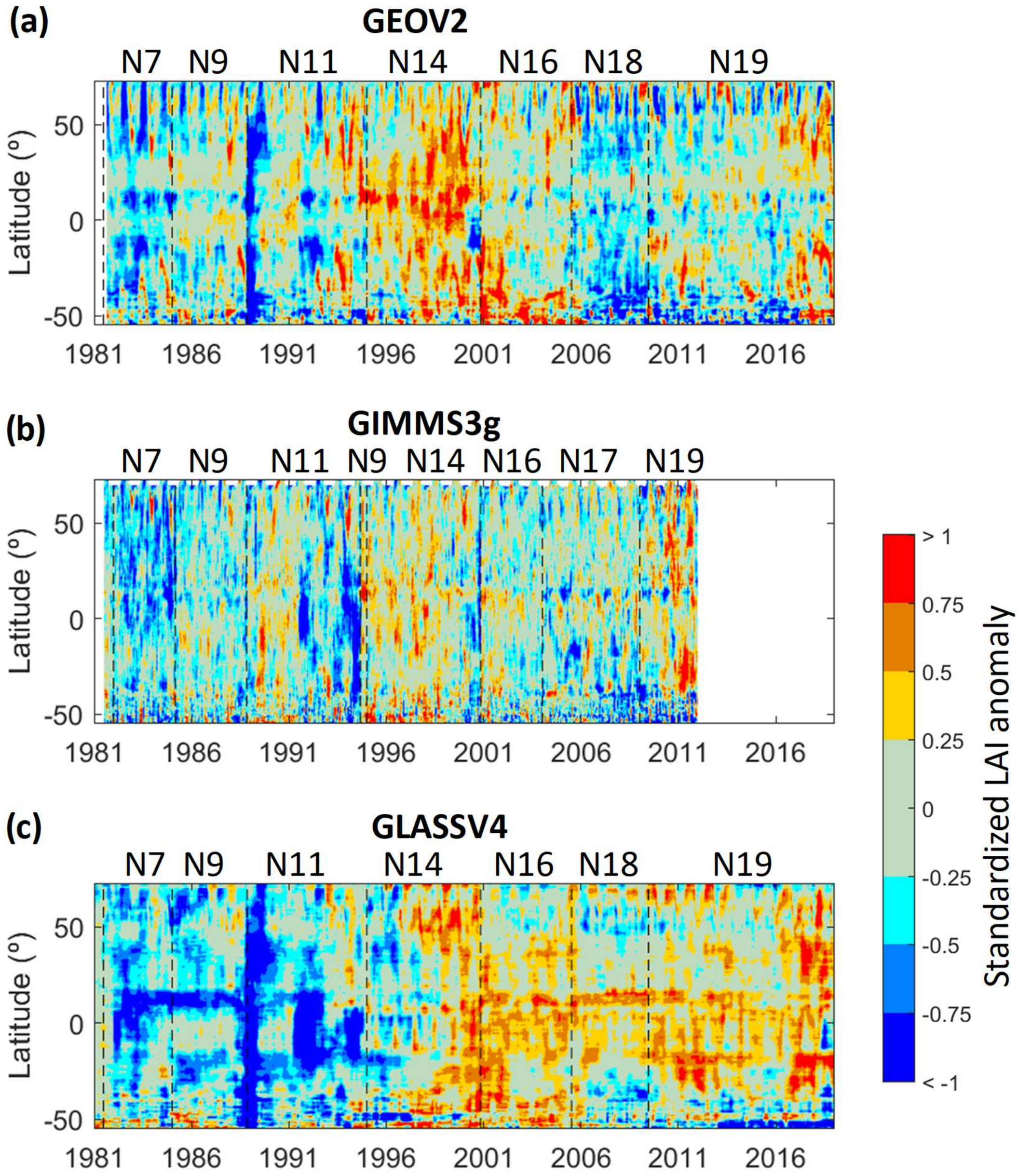

Figure 23. Hovmöller diagram (Hovmöller 1949) of the standardized anomalies for (a) GEOV2-AVHRR, (b) GIMMS3g and (c) GLASSV4 LAI products. Time is plotted along the abscissa and latitude along the ordinate. The black dashed lines indicate the date of change of sensors as described in Table 1.

GEOV2-AVHRR, GIMMS3g and GLASSV4 show consistent spatial patterns for the trends in LAI for the 1981-2011 common period (Figure 24). GEOV2-AVHRR agrees in the sense of trends with either GIMMS3g or GLASSV4 in 86% of cases, with both GIMMS3g and GLASSV4 in 55% of cases and it shows a mismatch only in 14% of land pixels (Figure 25). Greening in GEOV2-AVHRR LAI is observed in 79% of land pixels and browning in 21%. However, higher discrepancies exist in the magnitude and the significance of the observed trends (Figure 24). The main differences are for latitudes near the Equator and in the South Hemisphere where GLASSV4 product shows significant greening trends while the other products show non-significant greening or even browning trends over the Amazon or the Congo Basin.

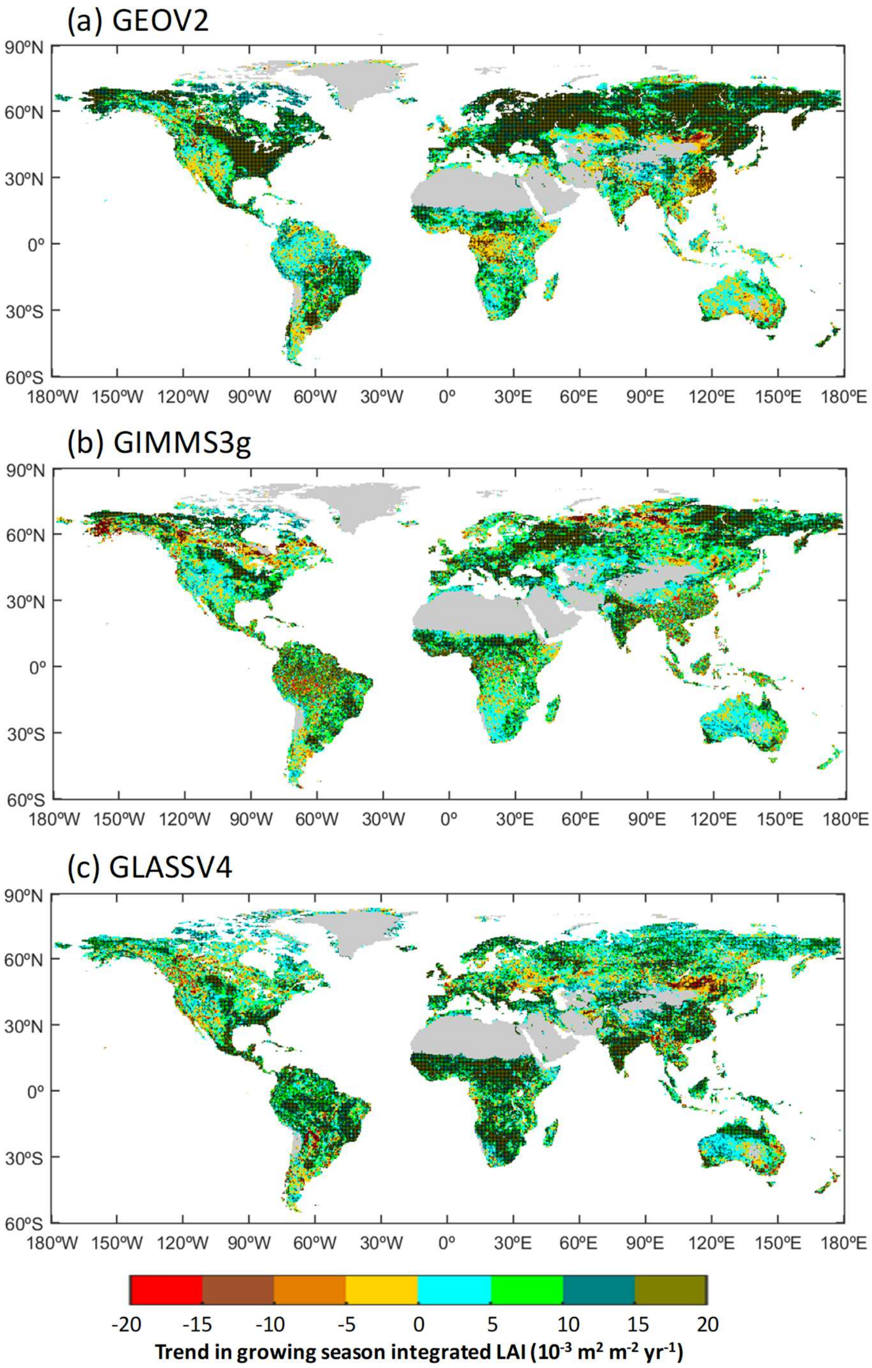

Figure 24. Spatial pattern of trends in the growing season integrated LAI derived from (a) GEOV2-AVHRR, (b) GIMMS3g and (c) GLASSV4 products for 1982-2011. Back dots highlight trends that are statistically significant (Mann–Kendall test; $p<0.05$).

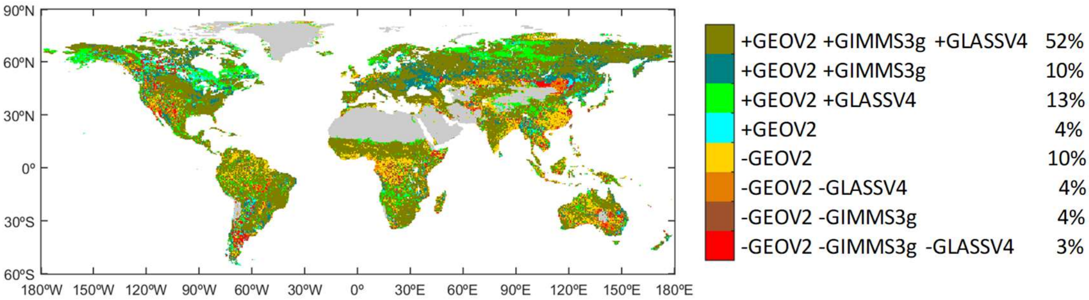

Figure 25. Spatial agreement of increase (green tones) or decrease (red tones) in growing season integrated LAI derived from GEOV2-AVHRR, GLASSV4 and GIMMS3g for 1982-2011. Cyan and yellow indicate, respectively, the regions with increase and decrease in GEOV2-AVHRR but not agreement with either GLASSV4 nor GIMMS3g. The percentage of agreement is indicated on the right of the legend.

3.5 Comparison with ground data

Confrontation with ground measurements shows similar performances for GEOV2-AVHRR, GEOV2-CGLS, GLASSV4 and GIMMS3g satellite datasets although GEOV2-AVHRR products perform the best in terms of the overall error: RMSE of 0.81 LAI, 0.10 FAPAR and 0.13 FCover (Figure 26, Table 3). GEOV2-AVHRR FCover better agree with ground data than GLASSV4 FCover for deciduous broadleaf forests and needleleaf forests (Figure 26) confirming that the differences between both products for these biomes (Figure 17) are mostly due to an overestimation of GLASSV4 to the actual FCover.

This validation is limited by the low number of available ground-based measurements that were mostly achieved in non-problematic conditions in northern mid latitudes close to the maximum peak of vegetation (Fuster et al. 2020; Garrigues et al. 2008). The validation is also affected by the uncertainty associated to reference maps which is expected to be around 1 LAI units for forest (Fernandes et al. 2003) or around 0.5 LAI for croplands (Martínez et al. 2009) and up to 0.1 for FAPAR (Gobron et al. 2008). These uncertainties are even larger for FCover reference maps mainly due to the very limited footprint of ground measurements since only 0-10º range of view of Digital Hemispherical Photography imagery is used for FCover ground sampling as compared to 0-60º range used for LAI/FAPAR (Fuster et al. 2020). Further, placing the camera below the canopy may modify its architecture an introduces artificial gaps in very dense and homogeneous canopies that results in an underestimation of actual FCover as measured from upwards DHP. This may partially explain the positive bias observed for both GEOV2-AVHRR and GLASSV4 products for crops and grassland sites (Figure 26) but also the overestimation reported for other products including CGLS Collection 1km (Camacho et al. 2013) and Collection 300m (Fuster et al. 2020) FCover products.

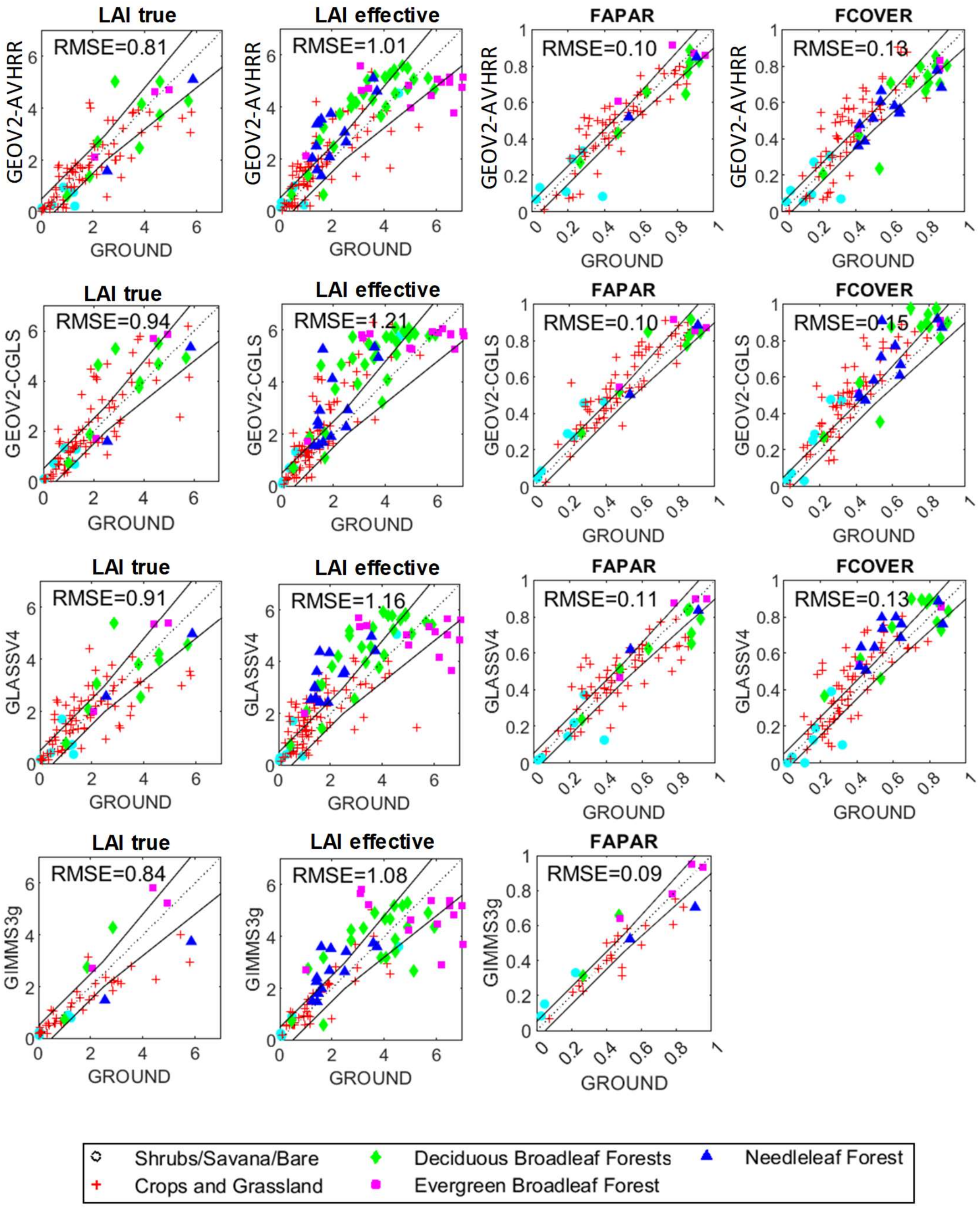

Figure 26. Comparison of GEOV2-AVHRR, GEOV2-CGLS, GLASSV4 and GIMMS3g products with ground measurements of true LAI, effective LAI, FAPAR and FCover in the 1999-2017 period. The different symbols correspond to the five biome classes as derived from the CCI landcover map (http://maps.elie.ucl.ac.be/CCI/viewer/download/ESACCI-LC-Ph2-PUGv2_2.0.pdf). The dotted line corresponds to the 1:1 line. The solid lines represent the GCOS accuracy criteria: max(20%, 0.5) LAI and max(10%, 0.05) FAPAR (FCover) (GCOS 2011). The root mean square error (RMSE) is indicated and other statistics are provided in Table 3.

Table 3. Statistics of the comparison of GEOV2-AVHRR, GEOV2-CGLS, GLASSV4 and GIMMS3g with ground measurements of true LAI, effective LAI, FAPAR and FCover over the DIRECT2.0 sites for all available samples in the 1999-2017 period and (*) for the common samples in the 1999-2011 period: 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 least squares linear regression. The best agreement with ground data is highlighted.

		Num. samples	%OK GCOS	RMSE	Bias	R	slope	offset
LAI true	**GEOV2-AVHRR**	**95**	**63**	**0.81**	**-0.10**	**0.85**	**0.90**	**0.09**
	GEOV2-CGLS	95	61	0.94	0.24	0.85	1.17	-0.10
	GLASSV4	95	59	0.91	0.04	0.80	0.87	0.30
	GIMMS3g	49	59	0.84	-0.18	0.86	0.81	0.16
	GEOV2-AVHRR*	**49**	**69**	**0.74**	**-0.14**	**0.89**	**0.93**	**-0.03**
	GEOV2-CGLS*	49	67	0.91	0.004	0.86	1.09	-0.16
	GLASSV4*	49	76	0.84	0.04	0.86	0.97	0.09
LAI effective	**GEOV2-AVHRR**	**147**	**47**	**1.01**	**0.31**	**0.84**	**0.94**	**0.44**
	GEOV2-CGLS	147	50	1.21	0.72	0.87	1.16	0.35
	GLASSV4	147	47	1.16	0.51	0.82	0.97	0.59
	GIMMS3g	90	57	1.08	0.08	0.83	0.83	0.52
	GEOV2-AVHRR*	**90**	**51**	**1.05**	**0.25**	**0.85**	**0.92**	**0.45**
	GEOV2-CGLS*	90	57	1.14	0.59	0.89	1.11	0.29
	GLASSV4*	90	51	1.17	0.52	0.85	0.96	0.62
FAPAR	**GEOV2-AVHRR**	**79**	**48**	**0.10**	**0.001**	**0.91**	**0.97**	**0.01**
	GEOV2-CGLS							
	GLASSV4	79	47	0.11	-0.01	0.87	0.89	0.04
	GIMMS3g	31	52	0.09	0.001	0.93	0.89	0.05
	GEOV2-AVHRR *	31	58	0.09	-0.001	0.94	1.01	-0.01
	GEOV2-CGLS*	31	61	0.08	0.02	0.96	0.99	0.03
	GLASSV4*	**31**	**61**	**0.07**	**-0.01**	**0.96**	**0.98**	**-0.004**
FCover	**GEOV2-AVHRR**	**97**	**35**	**0.13**	**0.03**	**0.85**	**1.07**	**0.001**
	GEOV2-CGLS*	97	27	0.15	0.11	0.91	1.14	0.05
	GLASSV4	97	31	0.13	0.05	0.88	1.13	-0.01

4 CONCLUSION

GEOV2-AVHRR products capitalize on the efforts undertaken to pre-process the AVHRR temporal series, resulting in the Long Term Data Record data, and the recent development of improved processing of biophysical variables resulting in the Copernicus Global Land Service (CGLS) Version 2 of Collection 1km LAI, FAPAR and FCover products (GEOV2-CGLS). The GEOV2-AVHRR algorithm was designed to ensure (i) GCOS requirements and (ii) high consistency with the biophysical products developed in the recent years, and particularly with GEOV2-CGLS products derived from VEGETATION and PROBA-V sensors. The comparison between GEOV2-AVHRR and GEOV2-CGLS demonstrates that this objective was achieved: 94% of land pixels meet GCOS requirements for LAI with differences within ±0.5 LAI and 85% of pixels are within ±0.05 FAPAR/FCover requirements. The comparison with such similar long-term products derived from AVHRR data shows GEOV2-AVHRR products meet GCOS and user requirements in >90% of land pixels for LAI, >60% of pixel for FAPAR and >50% of pixels for FCover with differences with GIMMS3g and GLASSV4 products within ±0.5 LAI, ±0.05 FAPAR and ±0.05 FCover. GEOV2-AVHRR constitutes an intermediate solution between GIMMS3g and GLASS V4 in terms of the magnitude of products: GIMMS3g and GLASS V4 agree better with GEOV2-AVHRR (~0.6 LAI and ~0.10 FAPAR) than between them (~0.8 LAI and ~0.13 FAPAR); temporal smoothness: GEOV2-AVHRR improves temporal consistency reducing the noise observed in GIMMS3g and the over-smoothing in GLASS V4; and, spatio-temporal continuity: 2% of missing data for GEOV2-AVHRR at the global scale as compared to 18% of GIMMS3g and no missing data for GLASS V4. GEOV2-AVHRR also improves in terms of internal consistency between LAI, FAPAR and FCover variables. The spatio-temporal pattern of temporal anomalies of GEOV2-AVHRR time series is very consistent with GIMMS3g and across NOAA AVHRR sensors. The temporal trends in GEOV2-AVHRR LAI agrees with GIMMS3g and/or GLASSV4 in 85% of land pixels with 80% of greening and 20% of browning as evaluated over the common period (1982-2011). GEOV2-AVHRR products show the best accuracy against ground measurements: 0.81 LAI, 0.10 FAPAR and 0.13 FCover.

The NOAA-19 orbit degradation and the significant data loss in the recent years highly impacts the quality of GEOV2-AVHRR products. Thus it is recommended not to extend LTDR-based GEOV2-AVHRR time series after end of 2018. In the future, other input datasets will be explored to extend and reprocess GEOV2-AVHRR 1981-2018 time series. In particular, the GEOV2-AVHRR production could be switched from the AVHRR instrument onboard NOAA-19 to the same instrument from MetOp-B mission.

The user should use the products with due attention to the quality flags and the associated uncertainties, in particular for areas with long periods of missing data. The GEOV2-AVHRR long-term data records are expected to contribute to monitor the global change and to play a key role in Earth science modelling.

5 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., Leroy, M., Roujean, J.L., Knorr, W., Lambin, E., & Linderman, M. (2003). CYCLOPES User Requirement Document. In. Avignon: INRA-CSE

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

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

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

Fernandes, R., Burton, C., Leblanc, S., & Latifovic, R. (2003). Landsat-5 TM and Landsat-7 ETM+ based accuracy assessment of leaf area index products for Canada derived from SPOT-4 VEGETATION data. *Canadian Journal of Remote Sensing., 20*, 241-258

Fernandes, R., Plummer, S., Nightingale, J., Baret, F., Camacho, F., Fang, H., Garrigues, S., Gobron, N., Lang, M.L., R., LeBlanc, S., Meroni, M., Martinez, B., Nilson, T., Pinty, B., Pisek, J., Sonnentag, O., Verger, A., Welles, J., Weiss, M., & Widlowski, J.L. (2014). Global Leaf Area Index Product Validation Good Practices. Version 2.0. . In M.R. G. Schaepman-Strub, & J. Nickeson (Ed.), *Best Practice for Satellite-Derived Land Product Validation: Land Product Validation Subgroup (WGCV/CEOS)* (pp. 1-76)

Fuster, B., Sánchez-Zapero, J., Camacho, F., García-Santos, V., Verger, A., Lacaze, R., Weiss, M., Baret, F., & Smets, B. (2020). Quality Assessment of PROBA-V LAI, fAPAR and fCOVER Collection 300 m Products of Copernicus Global Land Service. *Remote Sensing, 12*, 1017

Garrigues, S., Lacaze, R., Baret, F., Morisette, J.T., Weiss, M., Nickeson, J.E., Fernandes, R., Plummer, S., Shabanov, N.V., Myneni, R.B., Knyazikhin, Y., & Yang, W. (2008). Validation and intercomparison of global Leaf Area Index products derived from remote sensing data. *Journal of Geophysical Research: Biogeosciences, 113*, G02028

GCOS (2010). GCOS-13. : Implementation plan for the global observing system for climate in support of the UNFCCC (2010 update). In, *GCOS-138* (p. 186): WMO

GCOS (2011). Global Climate Observing System - Systematic Observation Requirements for Satellite-Based Products for Climate, 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, *https://www.wmo.int/pages/prog/gcos/Publications/gcos-154.pdf* (p. 138). Geneva, Switzerland: World Meteorological Organization

Gobron, N., Pinty, B., Aussedat, O., Taberner, M., Faber, O., Mélin, F., Lavergne, T., Robustelli, M., & Snoeij, P. (2008). Uncertainty estimates for the FAPAR operational products derived from MERIS — Impact of top-of-atmosphere radiance uncertainties and validation with field data. *Remote Sensing of Environment, 112*, 1871-1883

He, L., Chen, J.M., Pisek, J., Schaaf, C.B., & Strahler, A.H. (2012). Global clumping index map derived from the MODIS BRDF product. *Remote Sensing of Environment, 119*, 118-130

Hovmöller, E. (1949). The Trough-and-Ridge Diagram. *Tellus, 1*, 62-66

Jia, K., Liang, S., Liu, S., Li, Y., Xiao, Z., Yao, Y., Jiang, B., Zhao, X., Wang, X., Xu, S., & Cui, J. (2015). Global Land Surface Fractional Vegetation Cover Estimation Using General Regression Neural Networks From MODIS Surface Reflectance. *IEEE Transactions on Geoscience and Remote Sensing, 53*, 4787-4796

Jia, K., Yang, L., Liang, S., Xiao, Z., Zhao, X., Yao, Y., Zhang, X., Jiang, B., & Liu, D. (2019). Long-Term Global Land Surface Satellite (GLASS) Fractional Vegetation Cover Product Derived From MODIS and AVHRR Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 12*, 508-518

Liang, S., Zhao, X., Liu, S., Yuan, W., Cheng, X., Xiao, Z., Zhang, X., Liu, Q., Cheng, J., Tang, H., Qu, Y., Bo, Y., Qu, Y., Ren, H., Yu, K., & Townshend, J. (2013). A long-term Global LAnd Surface Satellite (GLASS) data-set for environmental studies. *International Journal of Digital Earth, 6*, 5-33

Martínez, B., Camacho, F., Verger, A., García-Haro, F.J., & Gilabert, M.A. (2013). Intercomparison and quality assessment of MERIS, MODIS and SEVIRI FAPAR products over the Iberian Peninsula. *International Journal of Applied Earth Observation and Geoinformation, 21*, 463-476

Martínez, B., García-Haro, F.J., & Camacho-de Coca, F. (2009). Derivation of high-resolution leaf area index maps in support of validation activities: Application to the cropland Barrax site. *Agricultural and Forest Meteorology, 149*, 130-145

McCallum, I., Wagner, W., Schmullius, C., Shvidenko, A., Obersteiner, M., Fritz, S., & Nilsson, S. (2010). Comparison of four global FAPAR datasets over Northern Eurasia for the year 2000. *Remote Sensing of Environment, 114*, 941

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

Pisek, J., Chen, J.M., Lacaze, R., Sonnentag, O., & Alikas, K. (2010). Expanding global mapping of the foliage clumping index with multi-angular POLDER three measurements: Evaluation and topographic compensation. *ISPRS Journal of Photogrammetry and Remote Sensing, 65*, 341-346

Tang, H., Yu, K., Hagolle, O., Jiang, K., Geng, X., & Zhao, Y. (2013). A cloud detection method based on a time series of MODIS surface reflectance images. *International Journal of Digital Earth, 6*, 157-171

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. (2014). Near real time vegetation monitoring at global scale. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7*, 3473-3481

Verger, A., Baret, F., Weiss, M., Filella, I., & Peñuelas, J. (2015). GEOCLIM: A global climatology of LAI, FAPAR, and FCOVER from VEGETATION observations for 1999–2010. *Remote Sensing of Environment, 166*, 126-137

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., Sánchez-Zapero, J., Weiss, M., Descals, A., Camacho, F., Lacaze, R., & Baret , F. (2023). Copernicus Global Land leaf area index, fraction of absorbed PAR and vegetation cover at 1km resolution: consistency and improvements of version 2. *Remote Sensing, Under review*

Weiss, M., Baret , F., & Verger, A. (2014). BELMANIP2: Enhancement of the CEOS-BELMANIP ensemble of sites used for the validation of land products from medium resolution sensors. In J. Sobrino (Ed.), *Recent Advances in Quantitative Remote Sensing (RAQRS IV)*. Valencia, Spain

Xiao, Z., Liang, S., Sun, R., Wang, J., & Jiang, B. (2015). Estimating the fraction of absorbed photosynthetically active radiation from the MODIS data based GLASS leaf area index product. *Remote Sensing of Environment, 171*, 105-117

Xiao, Z., Liang, S., Wang, J., Xiang, Y., Zhao, X., & Song, J. (2016). Long-Time-Series Global Land Surface Satellite Leaf Area Index Product Derived From MODIS and AVHRR Surface Reflectance. *IEEE Transactions on Geoscience and Remote Sensing, 54*, 5301-5318

Yuan, H., Dai, Y., Xiao, Z., Ji, D., & Shangguan, W. (2011). Reprocessing the MODIS Leaf Area Index products for land surface and climate modelling. *Remote Sensing of Environment, 115*, 1171-1187

Zhu, Z., Bi, J., Pan, Y., Ganguly, S., Anav, A., Xu, L., Samanta, A., Piao, S., Nemani, R., & Myneni, R. (2013). Global Data Sets of Vegetation Leaf Area Index (LAI)3g and Fraction of Photosynthetically

Active Radiation (FPAR)3g Derived from Global Inventory Modeling and Mapping Studies (GIMMS) Normalized Difference Vegetation Index (NDVI3g) for the Period 1981 to 2011. *Remote Sensing, 5*, 927-948

6 ANNEX: ASSESSMENT OF EXTENDED TIME SERIES FOR 2019-2021

We evaluate here the extended GEOV2-AVHRR 1981-2018 time series for the years 2019-2021 (Figure 27) using the available LTDR V4 AVHRR NOAA-19 data for these years. However, NOAA-19's orbit has been degrading in recent years (known issue of AVHRR data as reported in https://landweb.modaps.eosdis.nasa.gov). Specifically, the equator crossing time has been getting later and later, which increases the angle of the sun throughout the daytime side of the orbit. This impacts processing and results in a significant loss of data from each orbit because the BRDF correction algorithm fails at high sun angles. The impact of the degradation can be observed globally and has significantly increased over the year 2019.

The NOAA-19 orbit degradation also impacts GEOV2-AVHRR products (Figure 27). The number of valid observations for generating GEOV2-AVHRR have been reducing in recent year globally (Figure 28). AVHRR data is missing systematically in the south hemisphere for all dates over the years 2020 and 2021 (and partially in 2019) and globally in winter time. The quality of observations has also been degrading resulting in higher uncertainty in GEOV2-AVHRR estimates, as the increase in RMSE values for years 2019-2021 reflects (Figure 29). Note that in most of situations the RMSE was not computed for years 2019-2021 because not enough valid observations (less than 2 observations in a 120-day compositing window) (Figure 29). The analysis of anomalies of GEOV2-AVHRR time series shows more clearly the impact of NOAA-19 orbit degradation on products' quality. After 1st January 2019, GEOV2-AVHRR LAI shows negative bias in the south hemisphere where AVHRR data is systematically missing and LAI shows positive anomalies in the north hemisphere (Figure 30). Note that GEOV2-AVHRR products for the years 2019-2021 mostly results from the climatology gap filling because practically no valid observations are available globally (Figure 28). The observed anomalies are due to the loss of data and do not represent actual temporal changes in vegetation amount. Then GEOV2-AVHRR products for the years 2019-2021 are not suitable for most of application and particularly for the analysis of temporal changes. The analysis of anomalies of NOBS (Figure 30) showed clearly the significant data loss due to orbit degradation in recent years. This also results in positive anomalies in RMSE-LAI (Figure 30) as commented. Then even if it is technically possible to generate GEOV2-AVHRR products for the period 2019-2021 from LTDR AVHRR NOAA-19 data, it is recommended not to extend LTDR-based GEOV2-AVHRR time series after end of 2018. In this sense, GLASS V4 products, which were generated from the same LTDR dataset, were processed only until end of 2018. In the future, other input datasets will be explored to extend and reprocess GEOV2-AVHRR time series. In particular, the LTDR V5 production was switched from the AVHRR instrument onboard NOAA-19 to the same instrument from MetOp-B on September 2020. The GEOV2-AVHRR could be reprocessed and extended using all the AVHRR records from the beginning of the MetOp-B mission to the present, and updated annually.

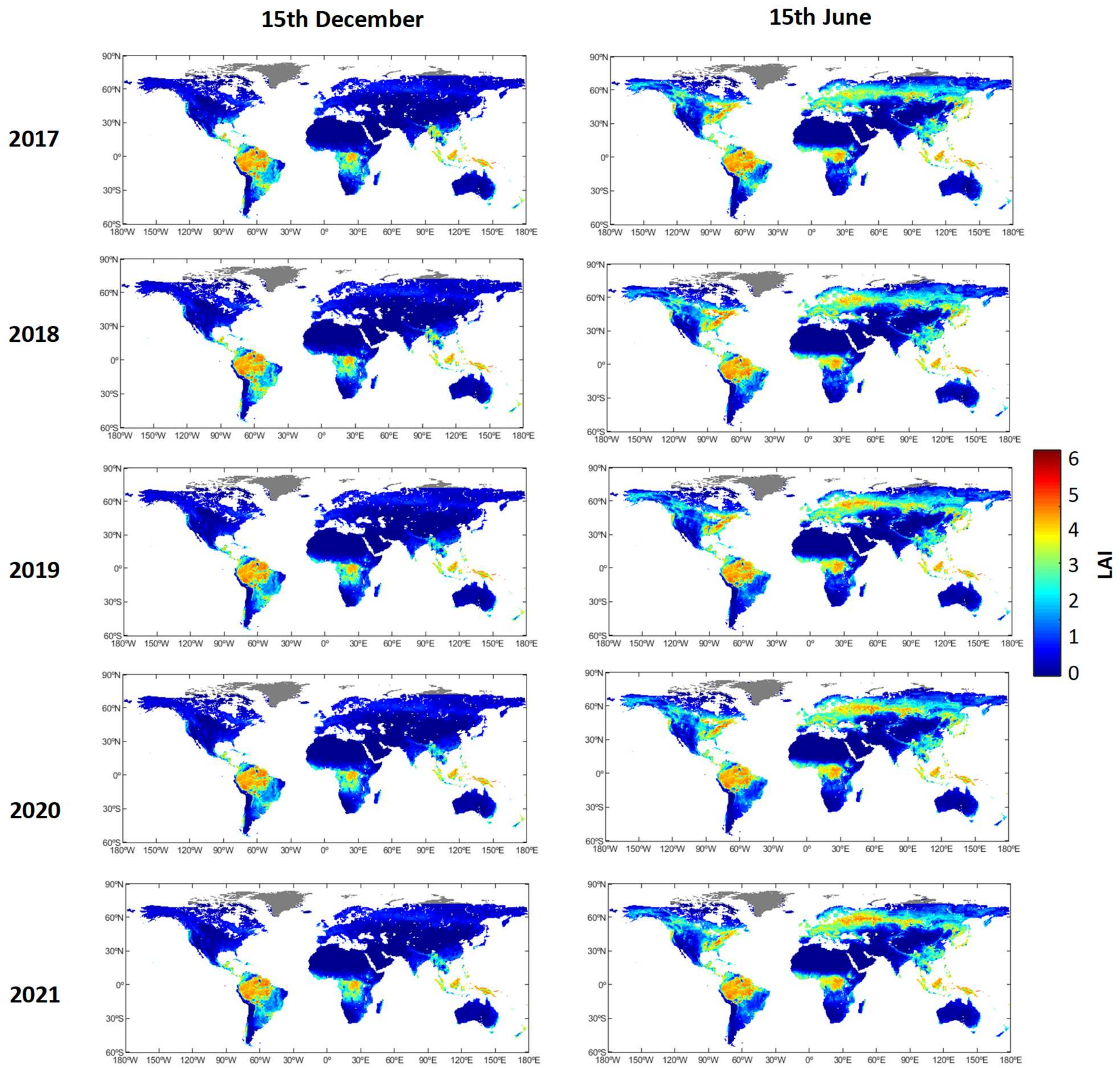

Figure 27. Map of GEOV2-AVHRR LAI for winter and summer solstice dates over the years 2017-2021. The areas in grey correspond to pixels with no data.

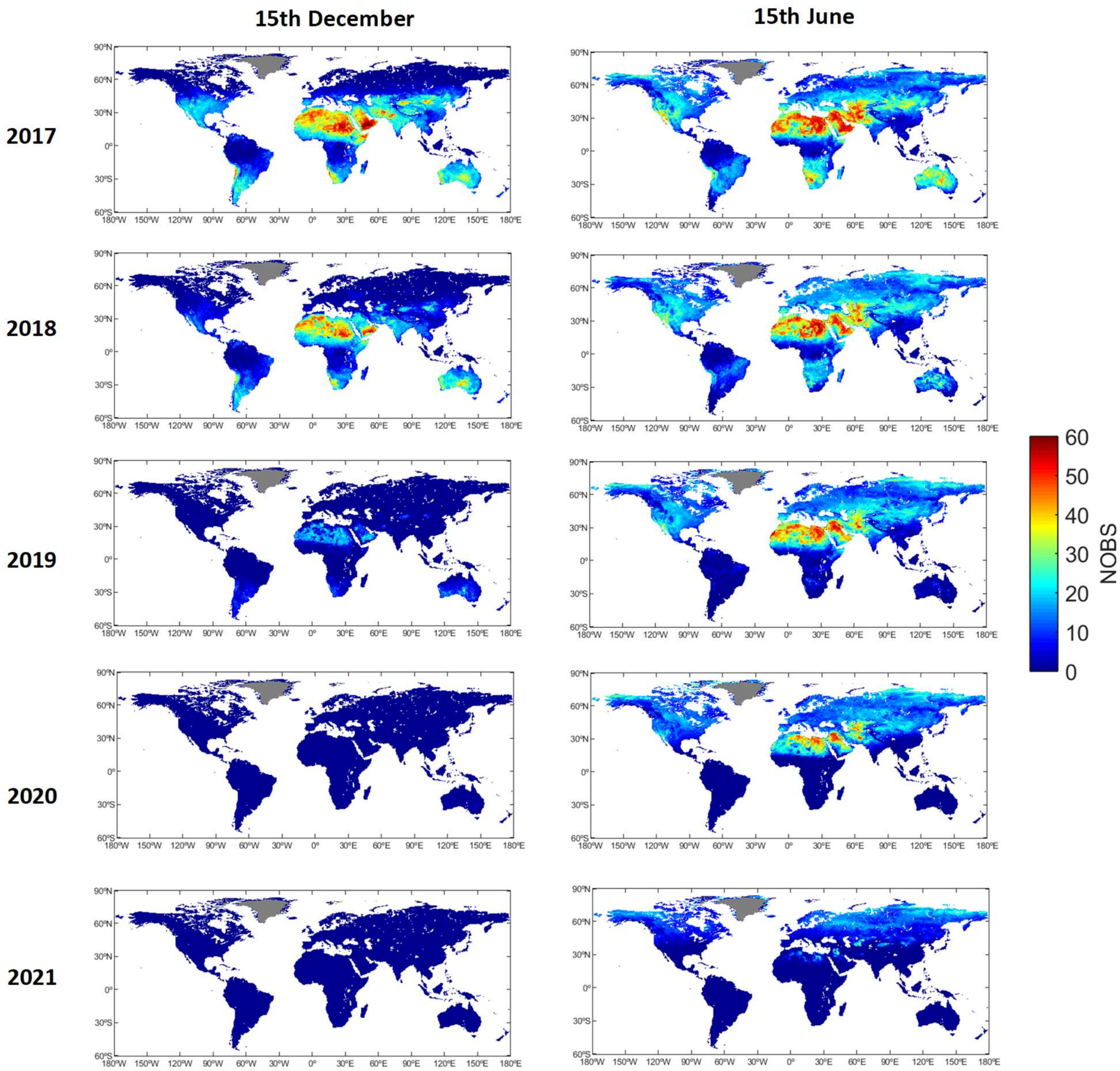

Figure 28. Map of the number of valid observations (NOBS) in the compositing window for winter and summer solstice dates over the years 2017-2021. The areas in grey correspond to pixels with no data.

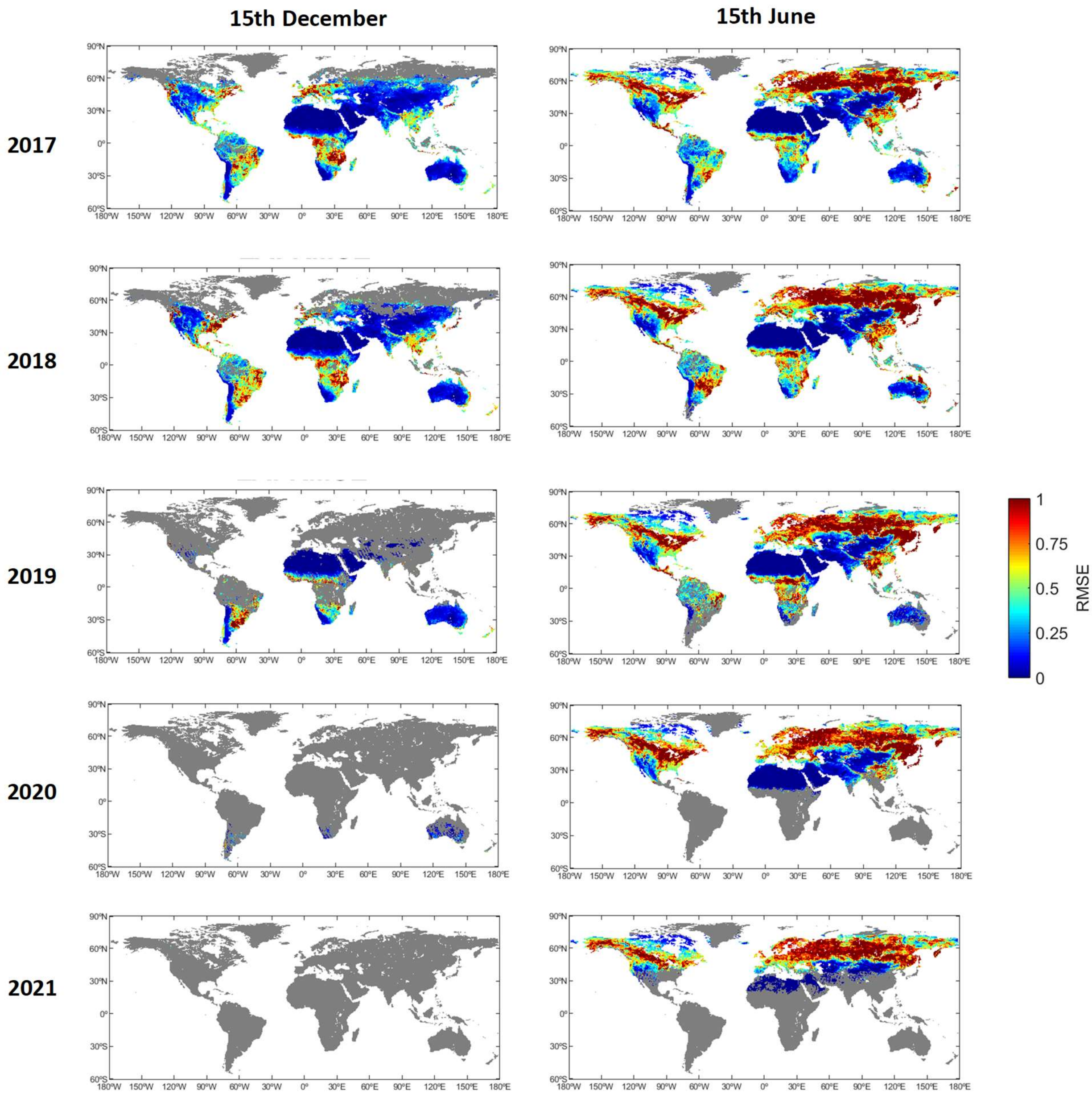

Figure 29. Map of the RME LAI for winter and summer solstice dates over the years 2017-2021. The areas in grey correspond to pixels with no data.

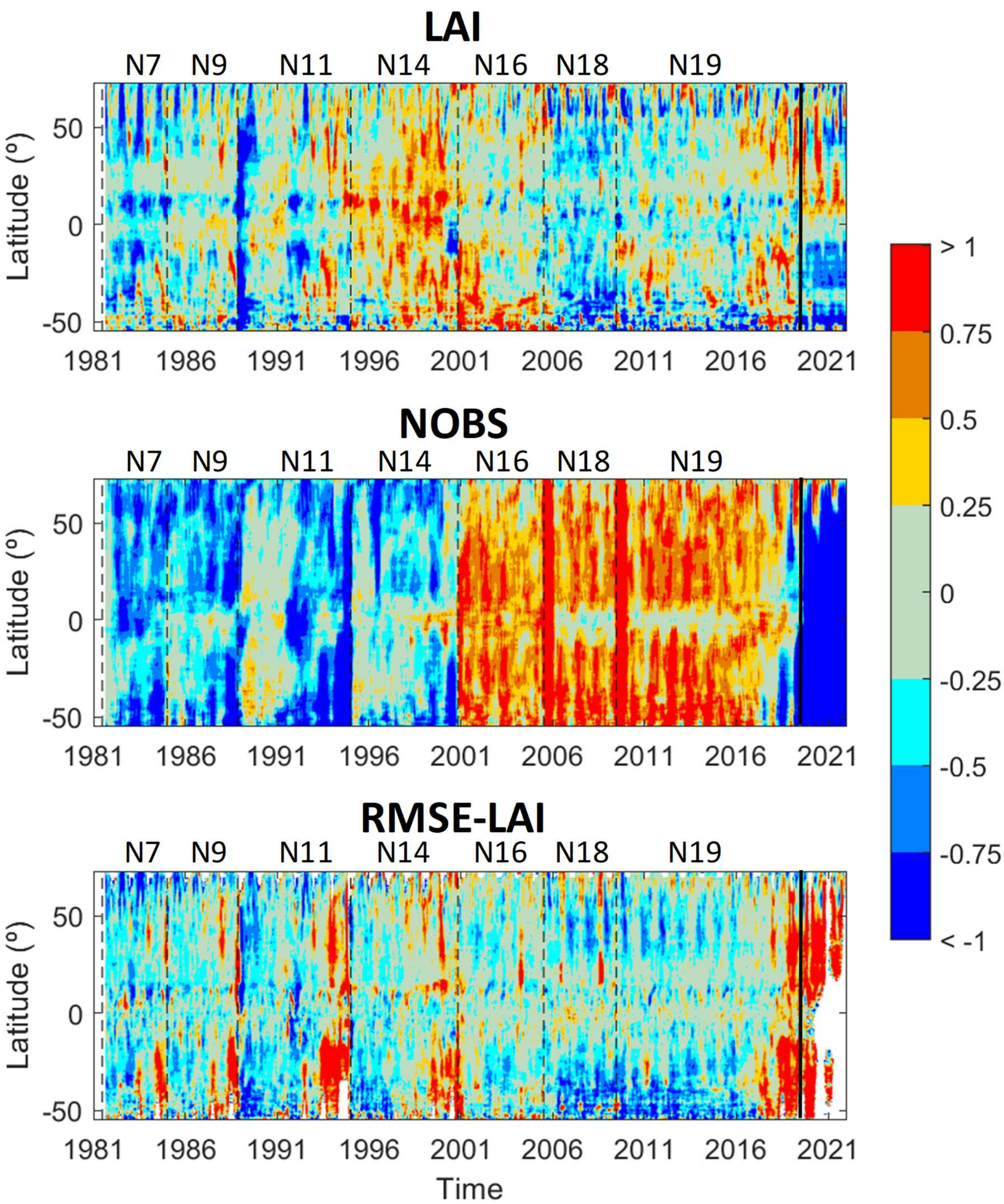

Figure 30. Hovmöller diagram of GEOV2-AVHRR LAI, NOBS and RMSE-LAI. Time is plotted along the abscissa and latitude along the ordinate. The dashed back lines indicate the date of change of NOAA (N) satellites. The continuous black line indicates the 31 December 2018 for which the GEOV2-AVHRR processing is recommended to be interrupted because of the NOAA-19 orbital degradation and its impact on products' quality.