vegetation parameters cci

CCI+ Vegetation Parameters

Product Validation and Algorithm Selection Report

Christiaan van der Tol, Simon Blessing, Jorge Sánchez-Zapero, Else Swinnen, Kris Vanhoof, Fernando Camacho

September 2023

vito remote sensing | UNIVERSITY OF TWENTE. ITC | EOLAB | FastOpt | HYGEOS | Imperial College London | University of Antwerp

Distribution list

Author(s)	:	Christiaan van der Tol, Simon Blessing, Fernando Camacho, Jorge Sánchez-Zapero, Kris Vanhoof
Reviewer(s)	:	Else Swinnen
Approver(s)	:	Clement Albergel
Issuing authority	:	VITO

Change record

Release	Date	Pages	Description of change	Editor(s)/Reviewer(s)
V1.0		All	First version	See above

Executive summary

This Product Validation and Algorithm Selection Report presents the methodology and outcome of the algorithm selection for the vegetation CCI product project, between two alternative approaches that were initially selected and developed. The first, TIP, is being used operationally in C3S, but with another method for surface albedo retrieval. It is a two-step algorithm in which OptiAlbedo computes the surface albedo using BRDF kernels factor spectra and TIP is a regression model based on a two-stream radiative transfer model inversion. The second, OptiSAIL, is a numerical inversion of the four-stream PROSAIL radiative transfer model as a more innovative option.

The selection is based on the direct and indirect validation of both algorithms against reference datasets (ground reference measurements and satellite derived) with respect to completeness, temporal stability, uncertainty, precision, and accuracy. The algorithms were also evaluated in terms of implementation aspects of computational demands, processing stability, memory use and data volume, and in terms of qualitative user requirements of ancillary data provision and the prospect for future innovations. The comparison was carried out for the tiles along a transect generated for the first published dataset, CRDP-1, which is based on SPOT-VGT and PROBA-V input (see Figure 1).

Both algorithms provide similar completeness but considering the quality flag OptiSAIL is less restrictive than TIP (using SPOT/VGT input data). Both algorithms provide data that are temporally consistent with reference datasets, but for some land cover types, TIP yields zero LAI values outside the growing season, and OptiSAIL yields some outliers that are not identified as such by the quality flag. In terms of intra-annual precision, OptiSAIL is superior to TIP. Both products show similar inter-annual precision, with median absolute anomalies of around 5%. OptiSAIL outperforms TIP in the comparison with reference ground datasets (DIRECT V2.1, GBOV V3 and AMMA) and reference satellite product (CGLS V2) in all metrics: a better correlation, better accuracy and lower uncertainties for both LAI and fAPAR products, resulting in a larger fraction of samples meeting GCOS goal and threshold requirements. TIP tends to provide lower values than references, mainly for the higher ranges. The LAI of TIP is low even considering that the algorithms provide effective value that requires correction for clumping.

In terms of processing time, OptiSAIL requires a substantially (one order of magnitude) longer time to process that OptiAlbedo+TIP, but the processing of global data for a 21-year period at 300 m resolution is not prohibitively long and can be achieved in the time frame of months of computation time. This can be shortened if additional processing resources are allocated. Memory requirements and stability are comparable between the algorithms. The final data volume for OptiSAIL is substantially larger, due to the larger number of output layers generated (including co-variances between parameters), but the total volume is still acceptable. After repackaging for distribution (removal of non-essential parameters) the OptiSAIL data volume is comparable to that of OptiAlbedo+TIP.

Specific user requirements identified at the start of the project, include the inclusion of the effects of soil and snow cover, the potential for retrieving additional variables of interest: the fraction of FAPAR by chlorophyll (fAPARchl), the possibility to retrieve information from solar induced chlorophyll fluorescence (SIF), and the fAPAR for blue sky. OptiSAIL has the potential to provide these products as a diagnostic variable, in case of SIF after the implementation of additional radiative transfer model code. The possibility exists to extend the product portfolio of Opti-Albedo-TIP as well, but this would be very involving as it requires alternative models such as bi-directional kernels for SIF and modifitaction of TIP.

OptiSAIL was selected due to its overall outperformance of OptiAlbedo-TIP in the validation and the qualitative user requirements, while the additional computational requirements for OptiSAIL are not

critical. Further implementation choices were made including the selection of a 5-day temporal resolution (window), the number of observations per band and per sensor was limited to three, a cloud detection algorithm in OptiSAIL was activated.

Table of Contents

LIST OF ACRONYMS

AMMA	Analyse Multidisciplinaire de la Mousson Africaine
APU	Accuracy, precision, uncertainty
ATBD	Algorithm Theoretical Basis Document
B	Bias
BHR	By-hemispherical reflectance
BRDF	Bidirectional Reflectance Distribution Function
CATCHA	Couplage de l'Atmosphère Tropicale et du Cycle Hydrologique
CEOS (-LPV)	Committee on Earth Observation Satellites (Land Product Validation)
CLM	Community Land Model
CRDP	Climate Research Data Package
CYCLOPES	Project acronym: 'satellite products for change detection and carbon cycle assessment at the regional and global scales'
DIRECT	Database of 3x3 km LAI and fAPAR data for validation
EOS	End of Season
*f*APAR	*f*raction of Absorbed Photosynthetically Active Radiation
GCOS	Global Climate Observing System
LAI	Leaf Area Index
LANDVAL	Land validation sites
MAD	Median absolute difference
MAR	Major axis regression
MD	Median difference
MODIS	Moderate resolution imaging spectrometer
NIR	Near Infra-Red range of the electromagnetic spectrum, here 700—2500 nm
OLS	Ordinary Least Squares
PROBA-V	Project for On-Board Autonomy
PROSPECT	PROperties of leaf SPECTtra
RM	Reference measurement
RMSD	Root mean squared error
RTM	Radiative transfer model
SBA	Sparse vegetated and Bare Areas
SIF	Solar Induced Chlorophyll Fluorescence
SOS	Start of Season
SPOT	Système Pour l'Observation de la Terre
SSD	System Specification Document
STD	Standard deviation
TARTUS	Two-streAm Radiative TransfEr in Snow
TIP	Two-stream Inversion Package
VGT	VEGATATION instrument onboard of SPOT4 & 5
VIS	VISible range of the electromagnetic spectrum, here 400–700 nm
WGCV	Working Group on Calibration and Validation

LIST OF FIGURES

LIST OF TABLES

1 Introduction

1.1 Scope of this document

The purpose of this document is to present the methodology and outcome of the algorithm selection for the vegetation CCI parameters project. Two alternative approaches were initially selected and developed, both based on the inversion of a physically radiative transfer model from solar reflective spectra. The selection is based on the validation of the algorithms against quality (threshold) criteria, an intercomparison to assess the relative differences in the performance against these criteria, and the compliance with user requirements for specific applications and prospects for future development.

In this document, the two alternative algorithms are briefly described with reference to the ATBD (Section 2), followed by the evaluation criteria (Section 3) and methods (Section 4), and the results of the selection process (Section 5). Finally, Section 6 provides a summary and list of risks and mitigation.

1.2 Related documents

Internal documents

Reference ID	Document
ID1	Climate Change Initiative Extension (CCI+) Phase 2 New ECVs: Vegetation Parameters – EXPRO+ - Statement of Work, prepared by ESA Climate Office, Reference ESA-EOP-SC-CA-2021-7, Issue 1.2, date of issue 26/05/2021
VP-CCI_D1.1_URD_V1.1	User Requirement Document: fAPAR and LAI, ESA CCI+ Vegetation Parameters https://climate.esa.int/media/documents/VP-CCI_D1.1_URD_V1.1.pdf
VP-CCI_D2.1_ATBD_V1.3	Algorithm Theoretical Basis Document: fAPAR and LAI, ESA CCI+ Vegetation Parameters http://climate.esa.int/media/documents/VP-CCI_D2.1_ATBD_V1.3.pdf
VP-CCI_D2.4_PVASR_V1.1	Product Validation and Algorithm Selection Report: fAPAR and LAI, ESA CCI+ Vegetation Parameters http://climate.esa.int/media/documents/VP-CCI_D2.4_PVASR_V1.1.pdf
VP-CCI_D4.1_PVIR_V1.2	Product Validation and Intercomparison Report: fAPAR and LAI, ESA CCI+ Vegetation Parameters http://climate.esa.int/media/documents/VP-CCI_D4.1_PVIR_V1.2.pdf
VP-CCI_D4.2_PUG_V1.2	Product User Guide: Lai and fAPAR, ESA CCI+ Vegetation Parameters http://climate.esa.int/media/documents/VP-CCI_D4.2_PUG_V1.2.pdf
VP-CCI_D1.3_PVP_V1.1	Product Validation Plan: fAPAR and LAI, ESA CCI+ Vegetation Parameters http://climate.esa.int/media/documents/VP-CCI_D1.3_PVP_V1.1.pdf
VP-CCI_D3.1_SSD_V1.1	System Specification Document, ESA CCI+ Vegetation Parameters http://climate.esa.int/media/documents/VP-CCI_D3.1_SSD_V1.1.pdf

External documents

Reference ID	Document
JCGM, 2014	JCGM, 2014. International Vocabulary of Metrology–Basic and General Concepts and Associated Terms, Chemistry International — Newsmagazine for IUPAC. Walter de Gruyter GmbH. https://doi.org/10.1515/ci.2008.30.6.21

GCOS-200, 2016	GCOS-200 (2016). The Global Observing System for Climate: Implementation Needs. WMO, Geneva, Switzerland https://library.wmo.int/opac/doc_num.php?explnum_id=3417

1.3 General definitions

Leaf Area Index (LAI) is defined as the total one-sided area of all leaves in the canopy within a defined region, and is a non-dimensional quantity, although units of [m2/m2] are often quoted, as a reminder of its meaning [GCOS-200, 2016]. The selected algorithm in the CCI-Vegetation Parameters project uses a 1-D radiative transfer model, and LAI is uncorrected for potential effects of crown clumping. Its value can be considered as an effective LAI, notably the LAI-parameter of a turbid-medium model of the canopy that would let the model have similar optical properties as the true 3-D structured canopy with true LAI [Pinty et al, 2006]. Additional information about the geometrical structure may be required for this correction to obtain true LAI [Nilson, 1971], which involves the estimation of the clumping index, CI, defined as the ratio between the true and effective LAI [see Fang, 2021 for a review of methods to estimate CI].

Fraction of Absorbed Photosynthetically Active Radiation (fAPAR) is defined as the fraction of Photosynthetically Active Radiation (PAR; solar radiation reaching the surface in the 400-700 nm spectral region) that is absorbed by a vegetation canopy [GCOS-200, 2016]. In contrast to LAI, fAPAR is not only vegetation but also illumination dependent. In the CCI-Vegetation Parameters project we refer to fAPAR as the white-sky value (i.e. assuming that all the incoming radiation is in the form of isotropic diffuse radiation). Total fAPAR is used and no differentiation is made between live leaves, dead foliage and wood.

Fraction of Absorbed Photosynthetically Active Radiation by Chlorophyll (fAPARcab) is defined as the fraction of Photosynthetically Active Radiation (PAR; solar radiation reaching the surface in the 400-700 nm spectral region) that is absorbed by Chlorophyll A + B molecules in the vegetation canopy.

Chlorophyll-A+B leaf pigment concentration is the amount of Chlorophyll A and B molecules per unit leaf area, typically measured in ug.cm-2.

Uncertainty is a measure to describe the statistically expected distribution of the deviation from the true value. Here, it is given as the physical value, which corresponds to the sigma-parameter of a gaussian distribution.

Correlation of uncertainties describes how uncertainties depend on each other. It is important information for error propagation. If, for instance, two measurements X and Y have highly correlated uncertainties, their difference X-Y will have a lower uncertainty than the uncorrelated case. Here, correlation of uncertainty is computed from the posterior variance-covariance matrix.

Surface albedo describes some of the reflectance properties of the surface. Here, we produce bi-hemispheric reflectance (BHR) for diffuse illumination with a reference spectrum for spectral broadband intervals VIS (400—700 nm), NIR (700—2500 nm), and SW (700—2500 nm), as well as directional-hemispherical reflectances (DHR) for the same spectral broadbands, computed for local solar noon.

Accuracy is the degree of the “closeness of the agreement between the result of a measurement and a true value of the measurand” [JCGM, 2014]. Commonly, accuracy represents systematic errors and

often is computed as the statistical mean bias, i.e., the difference between the short-term average measured value of a variable and the true value. The short-term average is the average of a sufficient number of successive measurements of the variable under identical conditions, such that the random error is negligible relative to the systematic error. The latter can be introduced by instrument biases or through the choice of remote sensing retrieval schemes [GCOS-200, 2016].

Precision or repeatability is the “closeness of the agreement between the results of successive measurements of the same measurand carried out under the same conditions of measurement” [JCGM, 2014].

Uncertainty is a “parameter, associated with the result of a measurement that characterizes the dispersion of the values that could reasonably be attributed to the measurand” [JCGM, 2014]. Uncertainty includes systematic and random errors.

2 Algorithm description

During the initial phase of the CCI+ Vegetation Parameters project, two physically based algorithmic chains were adapted and applied to the same data. These are the combination of OptiAlbedo with TIP, and OptiSAIL. OptiAlbedo+TIP is a multi-step algorithm with surface albedo as intermediate product, while OptiSAIL is applied directly to TOC reflectances. Both algorithmic chains handle data from multiple sensors and platforms, that is, multiple solar and viewing geometries, and band combinations, as they are observed during a time window (5 days in cycle 1), for a true multi-sensor retrieval. They retrieve all vegetation parameters jointly, such that they are consistent in terms of the respective physical model and the reflectance data.

Ideally, and by their definition, an ECV is not model dependent. However, for the retrieval of LAI and fAPAR the use of a model cannot be avoided. To minimize the model dependence, the two selected models are relatively simple and do not require ancillary data that vary in space and time, like land cover type, to avoid potential biases and to facilitate change detection.

As an addition developed during cycle 1, the reflectance data collected over the time window is pre-filtered for bright outliers in the shortest wavelength, to suppress undetected cloud contamination,

Per pixel propagation of TOC reflectance uncertainties (and potentially their correlations) to the final product is also done in both chains. The following subsections give a summary description of both algorithmic chains. Full details can be found in the ATBD [VP-CCI_D2.1_ATBD_V1.2].

2.1 OptiAlbedo + TIP

OptiAlbedo uses a truncated basis for the approximation of the surface BRDF. This basis is derived from a database of synthetic BRDF kernel factor spectra. The retrieved BRDF kernel factors are then converted into spectral broadband albedos (VIS, NIR) as required by TIP, using a regression model. The regression model is adapted per pixel, considering the BRDF kernel factor uncertainties. It is based on the same database of simulated Ross-Li BRDF kernel factor spectra. The regression in OptiAlbedo also determines a parameter describing the snow cover of the soil under the canopy, to provide a consistent snow mask as input for TIP, independent of the sensor-specific and therefore potentially heterogeneous snow flags in the TOC data.

TIP is the Two-stream Inversion Package, based on the two-stream model of Pinty et al (2006). TIP retrieves fAPAR and the turbid medium LAI (effective LAI) from spectral broadband albedos (VIS and NIR) range from inversions of the two-stream model. TIP uses a globally uniform prior for the inversions. By using tabulated inversions in the 2-dimensional observation space (two spectral broadband albedos), TIP is computationally very efficient.

2.2 OptiSAIL

OptiSAIL is an optimisation framework built around the models SAIL4H (Verhoef et al. 2007), PROSPECT-D (Feret et al., 2017), TARTES (Skiles and Painter, 2019), an empirical soil model with a semi-empirical moisture effect, and a cloud contamination model. They directly simulate TOC reflectances for given sets of spectrally invariant parameters (e.g., LAI, leaf pigments etc.) and scene geometries at given bands. In order to retrieve these parameters for observed TOC reflectance data, an inversion is made for each pixel. During cycle-1 of this project, repeatedly cloud-contaminated data was encountered, which was not flagged as such. Therefore, the cloud contamination model of OptiSAIL was activated, which simulates the effect of variable amounts of thin clouds per observation. This significantly reduces the number of outlier retrievals. The inversion in OptiSAIL minimises a cost function with data and prior term. It uses gradient information which is efficiently provided by adjoint code of the models. These adjoint codes are obtained by Automatic Differentiation (AD), which allows for quick adaption of the whole system to changes in the models. OptiSAIL includes an algorithm to model the effects of residual cloud contamination after atmospheric correction; this option has been used.

3 Evaluation criteria

The requirements as formulated in the most recent version of the implementation plan from the Global Climate Observing System (GCOS) programme [GCOS, 2022], which summarize the evolving needs in climate science (Table 1) formed the starting point for selection criteria. The requirements on uncertainty and stability were further specified and used for intercomparison of the two alternative algorithms.
Because the project will not deliver near-real time data, the requirement on data latency does not apply. However, for operationalization, the computational demands (Section **Error! Reference source not found.**) of the algorithm can be major constraint for the timeliness of data provision, and hence, the computational demands were considered in the selection.

Table 1: Requirements documented in GCOS-200 (GCOS, 2022)

	LAI		fAPAR	
	Threshold	**Goal**	**Threshold**	**Goal**
Time series length	20 years	40 years	20 years	40 years
Temporal resolution	10 days	1 day	10 days	1 day
Spatial resolution	250 m	10 m	250 m	10 m
Uncertainty (U)	max(20%,0.1)	max(10%,0.05)	max(10%, 0.005)	max(5%, 0.0025)
Long-term stability	Standard error <6%	Standard error <3%	Standard error <3%	Standard error <1.5%
Timeliness	10 days	1 day	10 days	1 day

The user requirement study [VP-CCI_D1.1_URD] highlighted additional qualitative criteria on LAI and fAPAR that are relevant for the algorithm selection: The functionality of the retrieval algorithm, and preferences for future ancillary vegetation parameter output beyond the product portfolio of LAI and fAPAR.

Hence, three groups of criteria for the algorithm selection have been used, related to the quantitative validation, implementation aspects, and qualitative requirements.

The selected criteria with respect to the uncertainty of LAI and fAPAR and their compliance with uncertainty requirements identified by GCOS include:

1. Completeness;
2. Temporal Consistency;
3. Error evaluation: Accuracy, intra-annual and inter-annual Precision, and Uncertainty;
4. Compliance with GCOS goal and threshold uncertainty requirements (Table 1).

Criteria for evaluating the implementation boundary conditions with respect to processing speed and computational demand include:

1. Processing time/ computational demand;
2. Memory usage;
3. Stability;
4. Data volume;
5. Implementation risks.

Criteria with respect to other (quantitative) user requirements, including the prospect for future expansion of the product portfolio and preferences include:

1. The inclusion of or the ability to correct for the effects of snow;

2. The inclusion of or the ability to correct for the effects of bare soil;
3. The ability to de-couple the effects of seasonally varying chlorophyll content and seasonally varying LAI on FAPAR;
4. To possibility to use the FAPAR and LAI products in synergy with SIF;
5. The possibility to provide a 'blue sky' or 'white sky' FAPAR product besides the black sky product;
6. The inclusion of albedo as an output;
7. The possibility to consider clumping and/or solutions for mixed pixels.

Table 2 summarizes the criteria and corresponding evaluation methods. Because precise weights of the three categories and the individual criteria cannot be given objectively, the final multi-criteria analysis is normative rather than numerical.

Table 2: Summary of all criteria used in the selection matrix

	Evaluation criterion	Indicator	Evaluation method
Validation	Completeness	Percentage of missing values	Comparison of missing values for the two algorithms for all retrievals and best quality retrievals
	Temporal Consistency	Realism of temporal variations	Qualitative inspection of the temporal profiles for both methods. Benchmarking with ground-based and satellite-based references.
	Accuracy	Mean bias, median deviations and slope/offset of regression line	Comparison of bias, median deviations and regression lines against reference data for the two prototypes
	Inter-annual precision	Boxplot of inter-annual anomalies per bins, and median absolute deviation for Cultivated and Evergreen broadleaved forest removed.	Comparison of boxplots of inter-annual anomalies for the two prototypes and their median values.
	Intra-annual precision	Temporal noise assumed to have no serial correlation within a season	Comparison of the smoothness (δ) of the two prototypes, and their median values.
	Uncertainty	Root Mean Square Deviation (RMSD)	Comparison of the RMSD values against reference datasets for the two algorithms.
	Compliance with User Requirements	Percentage of retrievals conform with GCOS goal and threshold uncertainty requirements	Comparison of the percentage of retrievals meeting GCOS uncertainty requirements for both algorithms
Implementation	Processing time / Computational demand	Actual processing time for the processing of half the transect. Analysis done, per processing step, per year, per tile and overall.	Comparison of the processing time between the two algorithms. Evaluation of the processing time wrt the processing resources.
	Memory usage	Actual memory usage for the processing of half the transect. Analysis done, per processing step, per year, per tile and overall.	Comparison of the memory usage between the two algorithms. Evaluation of the memory usage wrt the processing resources.
	Stability	Stability during the processing of the transect	Assessment of reliability of the processing chain (number of hick-ups, etc.)
	Data volume	Actual data volume of the output data set of half the transect. Analysis done, per processing chain, per year, per tile and overall.	Assessment of the required data volume wrt the available resources
	Implementation risks	Dependency of the processing chains on various components (input data, configuration, etc.)	Assessment based on the experience based on CRDP-1.
Qualitative User Requirements	Snow correction	The inclusion of the effect of (partial) snow cover	Evaluation of the way in which snow is represented in the radiative transfer model.
	Accounting for soil	The inclusion varying soil background spectra	Evaluation of the way in which soil reflectance is represented in the radiative transfer model.
	Chlorophyll FAPAR	The possibility to differentiate fAPAR by chlorophyll from total fAPAR.	Comparison with total fAPAR (OptiSAIL), or evaluation of the prospect for including Chlorophyll fAPAR in cycle 2 (TIP).
	Link with SIF	The prospect for including SIF.	Evaluation of compatibility with existing approaches for SIF (literature).
	Blue sky FAPAR	The possibility to provide blue sky fAPAR	Evaluation of the steps required to achieve this output.
	Albedo	The possibility to provide albedo as output	Evaluation of the radiative transfer model.
	Clumping	The possibility to include clumping (or post-processing)	Evaluation of the radiative transfer model wrt approaches in the literature.

4 Evaluation method

4.1 Prototype algorithm validation methodology

The validation methodology, which is described in the validation plan [CCI-VP_D1.3_PVP] was defined to be consistent with the CEOS LPV LAI validation protocol (Fernandes et al., 2014), which is also suitable for fAPAR products. The proposed methodology relies on direct validation and product intercomparison approaches. Both algorithms have been validated directly and indirectly with the same approach, and the metrics compared.

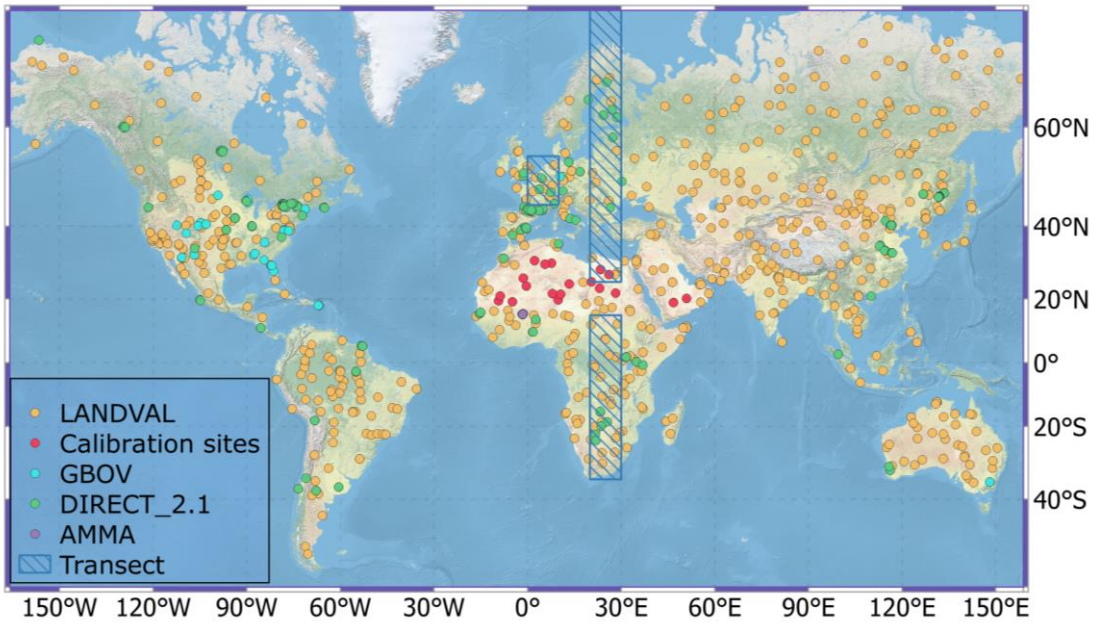

Figure 1: Sampling strategy: A) selected sites from LANDVAL, Calibration Sites, GBOV, DIRECT_2.1 and AMMA. B) Latitudinal Transect (see blue rectangles)

Figure 1 shows the sampling strategy used. It consists of selected sites for direct validation and the transect for indirect validation.

The **direct validation** is carried out against ground data set up-scaled according with the LPV recommendations (Fernandes et al., 2014; Morisette et al., 2006). The confidence in the reference ground-based map derived from empirical transfer functions depends on performances of the transfer functions that should be quantified with appropriate uncertainty metrics. Three different datasets are used for direct validation:

- The **CEOS WGCV LPV DIRECT V2.1** database, hosted at the CEOS cal/val portal (https://calvalportal.ceos.org/lpv-direct-v2.1), compiles LAI and fAPAR averaged values over a 3 km x 3 km area. The ground data was upscaled using high spatial resolution imagery following CEOS WGCV LPV LAI good practices to properly account for the spatial heterogeneity of the site (Garrigues et al., 2008). DIRECT V2.1 database constitutes a major effort of the international community to provide ground reference for the validation of LAI and FAPAR ECVs, with a total of 176 sites around the world (7 main biome types) and 280 LAI values and 128 FAPAR values covering the period from 2000 to 2021.
- **The Ground-Based Observations for Validation (GBOV) V3** (https://land.copernicus.eu/global/gbov), part of the Copernicus Global Land Service, aims at facilitating the use of observations from operational ground-based monitoring networks and

their comparison to Earth Observation products. In case of LAI and FAPAR, the GBOV service performs the implementation and maintenance of a database for the distribution of reference measurements (RMs) and the corresponding Land Products (LPs) (i.e., upscaled maps). Currently, GBOV provides multi-temporal Land Products over 27 sites. The current version (V3) of GBOV LP algorithm takes as input the Reference Measurements (RMs) collected over a given site, in addition to a series of high spatial resolution images. Calibration functions are then derived between RM and Radiative Transfer Model (RTM)-based retrievals, enabling high spatial resolution maps of each RM to be produced.

- **AMMA – Cycle Atmosphérique et Cycle Hydrologique (CATCH) system** has collected a data set composed of LAI, fAPAR and clumping index in the Sahelian rangelands of Gourma region in Mali over the 2005-2017 period. These 1 km x 1 km sites were chosen within large and relatively homogeneous areas to sample the main vegetation types and canopies encountered within the super-site. The variables were derived from the acquisition and the processing of hemispherical photographs taken along 1 km linear sampling (not upscaled). At each sampling date, 100 or 50 hemispherical photographs were acquired at the 1 km herbaceous or 0.5 km forest sites, respectively.

Intercomparisons with similar remote sensing products (i.e., **indirect validation**) can determine whether the products behave similarly in space and time on a global scale and allow identifying differences between products to be investigated in more detail to diagnose product anomalies and devise algorithm refinements. The LAND VALidation (LANDVAL) network of sites (Fuster et al., 2020; Sánchez-Zapero et al., 2020) is used for sampling global conditions in the intercomparison with similar satellite products.

- **CGLS Collection 1km V2** (Verger et al., 2023) products are used for benchmarking. CGLS V2 products are derived from SPOT/VGT and PROBA-V data at global scale from January 1999 to June 2020 at 1/112◦ spatial resolution and 10-day frequency, and in contrast to the data products in this project, they are both gap-filled and smoothed. The products are generated in two steps. The first step is based on neural networks trained on a combination of the existing CYCLOPES and MODIS products (Baret et al., 2013) that generate daily LAI and FAPAR estimates. The second step uses dedicated temporal smoothing and gap filling techniques to provide the final 10-day products and ensure consistency and continuity in the time series.

The following criteria are analysed: product completeness, temporal consistency, error evaluation, which involves Accuracy, Precision and Uncertainty (APU).

Product Completeness: Completeness corresponds to the absence of spatial and temporal gaps in the data. Missing data are mainly due to cloud or snow contamination, poor atmospheric conditions, or technical problems during the acquisition of the images and is generally considered by users as a severe limitation of a given product. It is therefore mandatory to document the completeness of the product (i.e., the distribution in space and time of missing data).

Temporal Consistency: The realism of the temporal variations and the precision of the products were assessed over the 720-site LANDVAL network plus additional sites with availability of ground measurements (i.e., DIRECTV2.1, GBOV, AMMA). The temporal variations of the product under study are qualitatively analysed as compared to reference products and available ground measurements.

Error evaluation: Accuracy, **Precision** and **Uncertainty** (APU) are evaluated by several metrics (Table 3) reporting the goodness of fit between the products and the corresponding reference dataset. **Accuracy** represents systematic errors and often is computed as the statistical mean bias (B). **Precision** represents the dispersion of product retrievals around their expected value and can be estimated by the standard deviation (STD) of the difference between retrieved satellite product and the

corresponding reference estimates. **Uncertainty** includes systematic and random errors and can be estimated by the Root Mean Square Deviation (RMSD). In addition to these metrics, other statistics are useful to evaluate the goodness of fit between two datasets including linear model fits. For this purpose, Major Axis Regression (MAR) is computed instead Ordinary Least Squares (OLS) because it is specifically formulated to handle error in both the x and y variables (Harper, 2014). In case of LAI, CEOS LPV recommends RMSD as the overall performance statistic to evaluate the accuracy, due to limitation in the temporal availability of ground datasets (Fernandes et al., 2014). It should be noted that strong and/or multiple outliers affect the classical metrics described above (i.e., B and STD): in such cases using the median deviation (MD) instead of the mean bias to estimate systematic error and the median absolute deviation (MAD) as a measure of precision is more suitable.

Two additional aspects of the precision have also been evaluated, notably the inter-annual and intra-annual precision (Fernandes et al., 2014).
Intra-annual precision (smoothness) corresponds to temporal noise assumed to have no serial correlation within a season. In this case, the anomaly of a variable from the linear estimate based on its neighbours can be used as an indication of intra-annual precision. It can be characterized (Weiss et al., 2007) as follows: for each triplet of consecutive observations, the absolute value of the difference between the centre P(ti) and the corresponding linear interpolation between the two extremes P(ti-1) and P(ti+1) is computed:

$$\delta = \left| P(t_i) - P(t_{i-1}) - \frac{P(t_{i-1}) - P(t_{i+1})}{t_{i-1} - t_{i+1}} (t_{i-1} - t_i) \right| \qquad \text{Eq. 1}$$

The distribution of the intra-annual precision is analysed, and the median δ value is used as a quantitative indicator of the inter-annual precision (Fernandes et al., 2014; Wang et al., 2019). Hence, the lower median of δ values, the higher the inter-annual precision.

Anomalies of an upper and lower percentile of variable are indicators **of inter-annual precision**, i.e., dispersion of variable values from year to year (Fernandes et al., 2014). It can be assessed providing a boxplot of the absolute anomalies for a given product between consecutive years per bins, and its median value. Note that cultivated sites are not considered in this analysis due to the non-natural variability in this land cover type due to agricultural practices (e.g., crop rotation). In addition, Evergreen Broadleaf Forest sites are not considered in the analysis since they are typically affected by cloud coverage for most of the products, and values are filled in case of products using gap-filling techniques.

Table 3: Validation metrics for product validation

Statistics	Comment
N	Number of samples. Indicative of the power of the validation
B	Mean Bias. Difference between average values of x and y. Indicative of accuracy and offset.
MD	Median deviation between x and y. Best practice reporting the accuracy.
STD	Standard deviation of the pair differences. Indicates precision.
MAD	Median absolute deviation between x and y. Best practice reporting the precision.
RMSD	Root Mean Square Deviation. RMSD is the square root of the average of squared errors between x and y.
MAR	Slope and offset of the Major Axis Regression linear fit. Indicates some possible bias
R	Correlation coefficient. Indicates descriptive power of the linear accuracy test. Pearson coefficient is used.

4.2 Implementation aspects

A set of tiles was used to evaluate the processing time, memory, stability and data volume following the methods summarized in Table *4*.

Table 4: Evaluation methods for the implementation aspects

Criterion	Method
Processing time / Computational demand	Evaluation by processing a set of tiles on the CCI Vegetation cluster processing environment and analysing the tracked performance metrics.
Memory usage	Assessment the memory used per individual component of both processing algorithms, as recorded during the processing of the full time series for 11 tiles.
Stability	Evaluation of processing behaviour.
Data volume	Assessment of the data volume after processing.
Implementation risks	Assessment of the risks.

4.3 Qualitative user requirements

The user requirement assessment, based on an online survey and interviews with experienced users of LAI and fAPAR products, resulted in priorities and key features of a product of added value. The inclusion of snow and soil were identified as important, since they influence the quality of the products in at the start of season (SOS) and end of season (EOS) in temperature, boreal and arctic climate. Furthermore, consistency among the three products fAPAR, LAI and surface albedo and with other data sets was considered important. The users identified additional vegetation data products of interest, including fAPAR in different illumination conditions (for example blue sky fAPAR), fAPAR differentiated by pigments, and approaches to use solar induced chlorophyll fluorescence (SIF). Finally, most users prefer an LAI product that accounts for clumping (true rather than effective LAI), and a product that is as little model-specific as possible.

The methodology included an evaluation to what degree these elements are already implemented in the algorithms, and an evaluation of the feasibility to implement this in a future development of the algorithm. This was evaluated by considering algorithms that have been published in the scientific literature.

5 Results of the assessment

5.1 Uncertainty of LAI and fAPAR: Prototype algorithm validation

5.1.1 Product completeness

Figure 2 shows the temporal evolution of missing values for TIP and OptiSAIL prototype algorithms for two different periods: 2004-2005 (based on SPOT/VGT input data) and 2019 (based on PROBA-V input data) over LANDVAL sites. The maps of the percentage of missing values over LANDVAL sites are displayed in Figure 3.

The findings are:

- Both algorithms show similar distribution of missing values over LANDVAL sites, with higher fraction of missing data in wintertime (northern hemisphere) and equatorial areas, as expected due the higher presence of clouds and snow (in case on northern regions).
- When quality flags are used to remove pixels with suboptimal quality (see dashed lines in Figure 2), TIP provides slightly higher fraction of missing values than OptiSAIL for SPOT-VGT and similar values for PROBA-V.

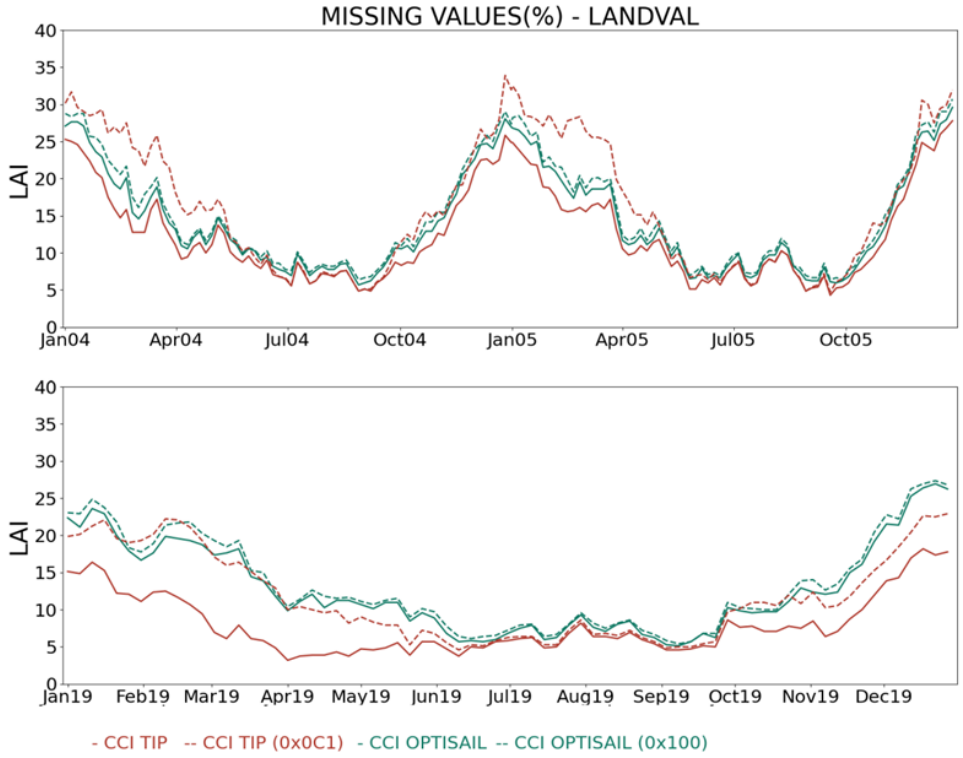

Figure 2: Temporal variation of the percentage of missing values (computed over LANDVAL sites) for TIP (red) and OptiSAIL (green) during 2004-2005 (top, based on SPOT-VGT) and 2019 (bottom, based on PROBA-V). The computation of gaps was performed considering all pixels (continuous lines) and filtering using quality flags (dashed lines).

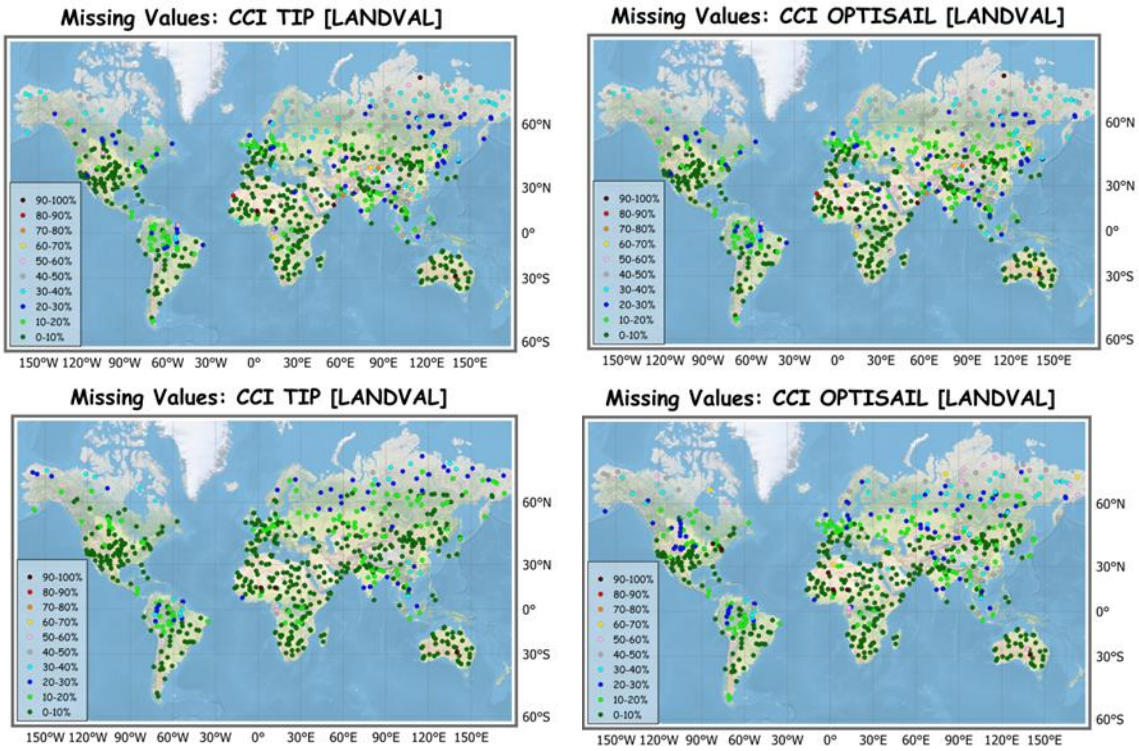

Figure 3: Maps of missing values (computed over LANDVAL sites) for TIP (left side) and OptiSAIL (right) during 2004-2005 (top, based on SPOT/VGT) and 2019 (bottom, based on PROBA-V).

5.1.2 Temporal consistency

The realism of the temporal variations of TIP and OptiSAIL was evaluated over sites with availability of multi-temporal ground observations over DIRECT V2.1 (Figure 4), GBOV V3 (Figure 5) and AMMA (Figure 6) sites. CGLS V2 products were also included in the comparison for benchmarking. The spatial support for the intercomparison is 3km^2 (i.e., averaged values of 3x3 SPOT/VGT or PROBA-V pixels) in case of DIRECT V2.1 and GBOV V3. However, for AMMA sites, the spatial support is 1km^2 (1 pixel) as ground data is provided over homogeneous transects of around 1km. Only best quality pixels according to quality flags are displayed.

The findings are:

- For DIRECT V2.1 sites (Figure 4), both TIP and OptiSAIL algorithms are temporally consistent with ground observations. OptiSAIL typically reaches higher values for LAI and fAPAR for crops and forests, which seems to be more consistent in magnitude with ground data (and CGLS V2).
- For GBOV V3 sites (Figure 5), OptiSAIL provides better temporal agreement than TIP with ground data over forest sites, reaching higher values during the leaf-on season which is more consistent with ground observations (and CGLS V2). Both OptiSAIL and TIP display consistent temporal trajectories for croplands, grasslands and shrublands but OptiSAIL shows several outliers which are not properly identified by quality flags.
- Similar results are observed over AMMA grassland sites (Figure 6), where both TIP and OptiSAIL algorithms are temporally consistent with multi-temporal ground observations.
- In all cases, TIP algorithm provides noisier temporal trajectories than OptiSAIL, mainly during the leaf-on season over forests and crops.
- TIP provides in some cases unrealistic values of zero during winter season over areas where the references provided non-zero values (e.g., AMMA#4, #5, #10, #11 or #12 grassland sites in Figure 6).

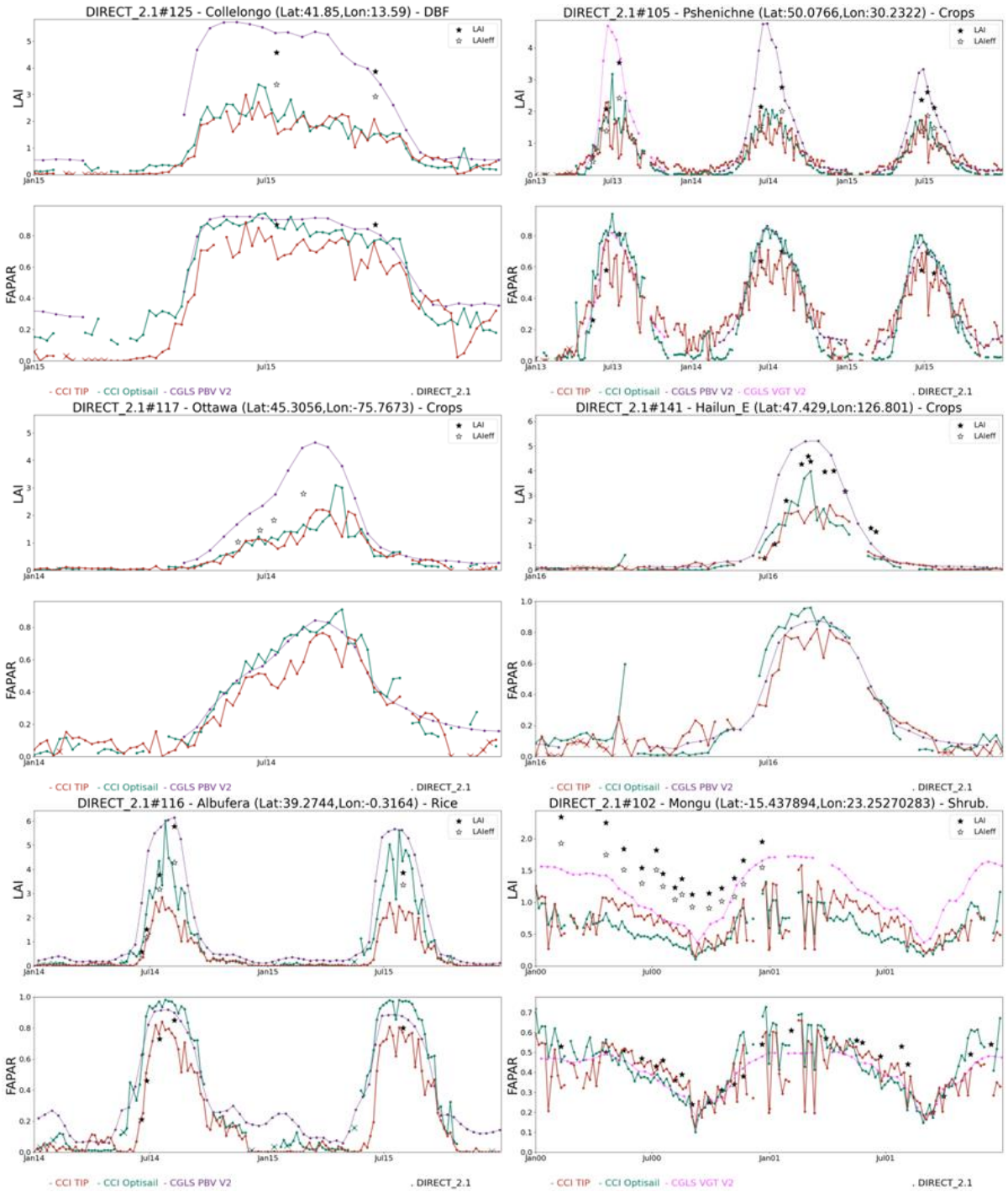

Figure 4: Temporal profiles over a selection of DIRECT V2.1 sites with availability of multi-temporal ground observations of TIP (red), OptiSAIL (green) and CGLS V2 (purple) products. Crosses in TIP and OptiSAIL represent pixels which are discarded by quality flags.

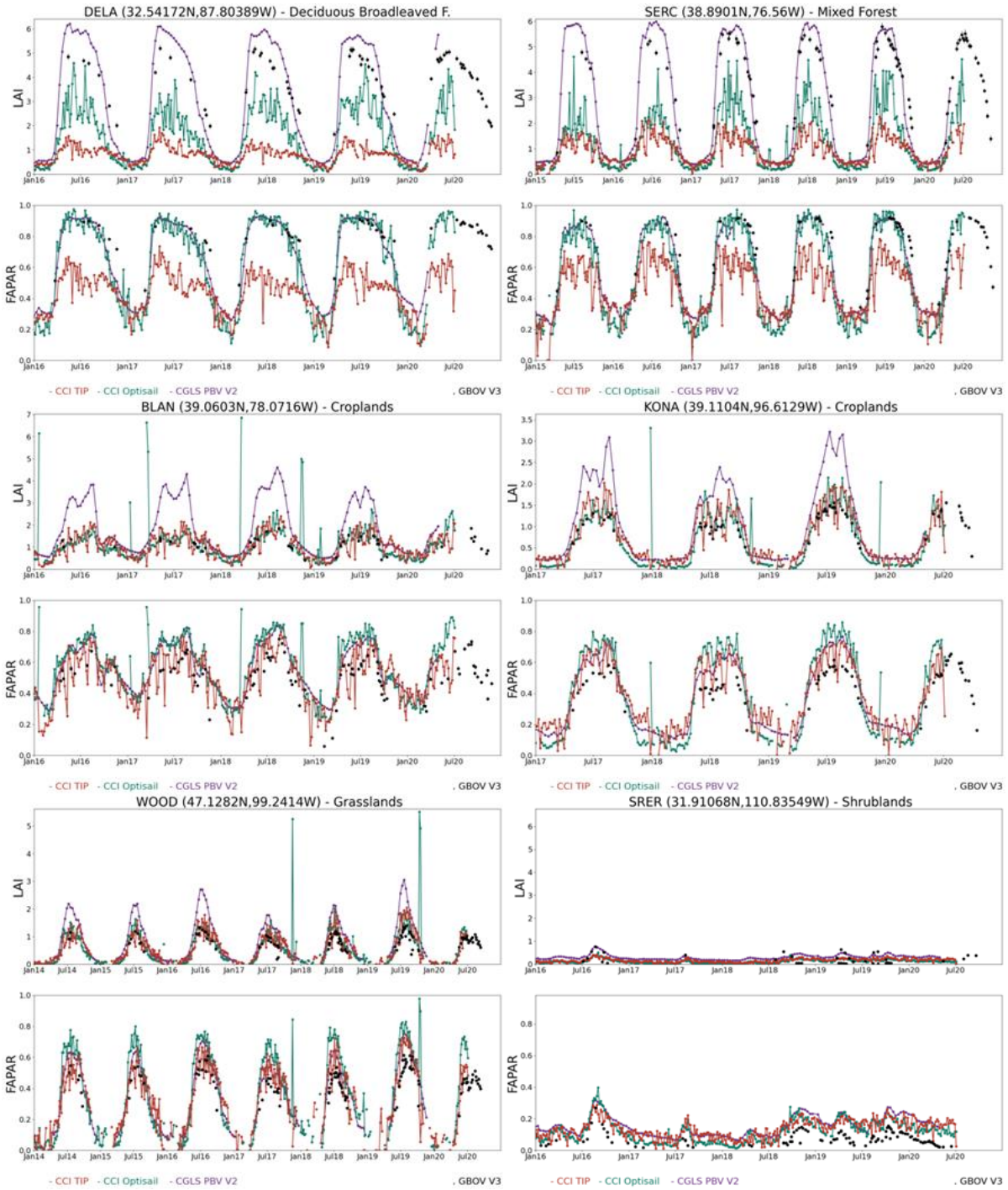

Figure 5: Temporal profiles over a selection of GBOV V3 sites with availability of multi-temporal ground observations of TIP (red), OptiSAIL (green) and CGLS V2 (purple) products.

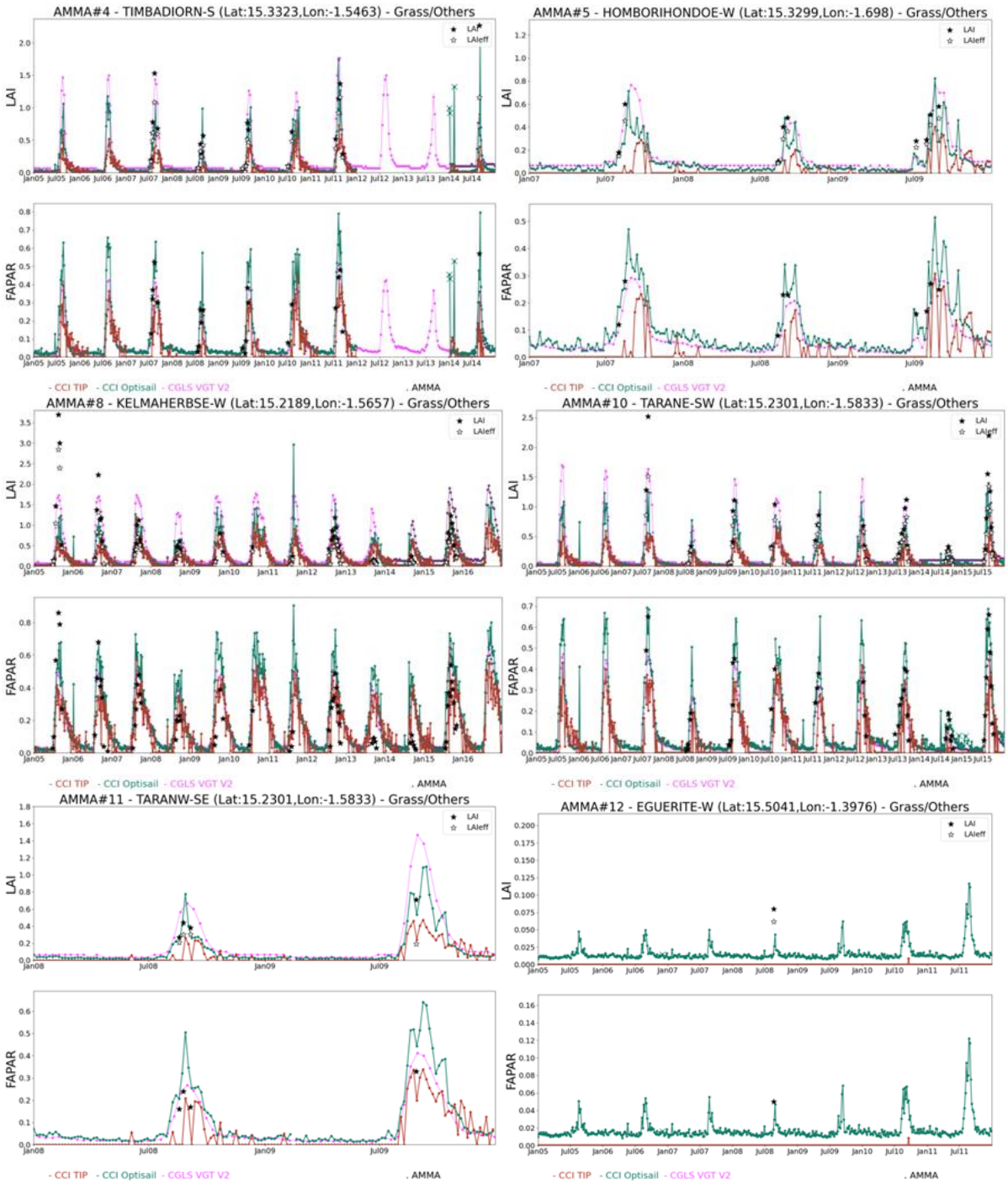

Figure 6: Temporal profiles over a selection of AMMA sites with availability of multi-temporal ground observations of TIP (red), OptiSAIL (green) and CGLS V2 (purple) products. Crosses in TIP and OptiSAIL represent pixels which are discarded by quality flags.

Figure 7 and Figure 8 show scatterplots between TIP and OptiSAIL LAI and fAPAR products versus DIRECT V2.1 effective LAI and fAPAR ground-based reference maps. Concomitant 'best quality' samples between both satellite products under study are used, and comparison was performed at 3km^2 (i.e., average values of 3x3 pixels).

The findings for LAI are:

- OptiSAIL shows better agreement than TIP in terms of correlation (R of 0.82 vs 0.79), overall uncertainty (RMSD of 0.9 vs 1.1) and accuracy (mean bias of -0.5 vs -0.7). Despite of the fact that both satellite products are compared with ground-based effective LAI maps, negative biases and slopes lower than 1 are found. However, OptiSAIL reaches higher values (slope of 0.69) than TIP (slope of 0.52), which is more realistic.
- OptiSAIL provides higher percentage of cases within goal (17%) and threshold (28%) GCOS uncertainty requirements than TIP (10% and 22%).

The findings for fAPAR are:

- Similar performances are found for TIP and OPTISAL. TIP provides slightly better overall agreement than OptiSAIL (RMSD = 0.14 vs 0.15) but slightly worse correlation (R of 0.80 vs 0.82). TIP provides mean negative bias, mainly observed for higher values (slope = 0.84). OptiSAIL shows systematic positive bias of 0.07 compared with DIRECT V2.1 but with linear relationship (slope of 1).
- OptiSAIL provides higher number of cases than TIP within GCOS goal requirement (7% vs 10%) and lower number for threshold level (28% vs 22%).

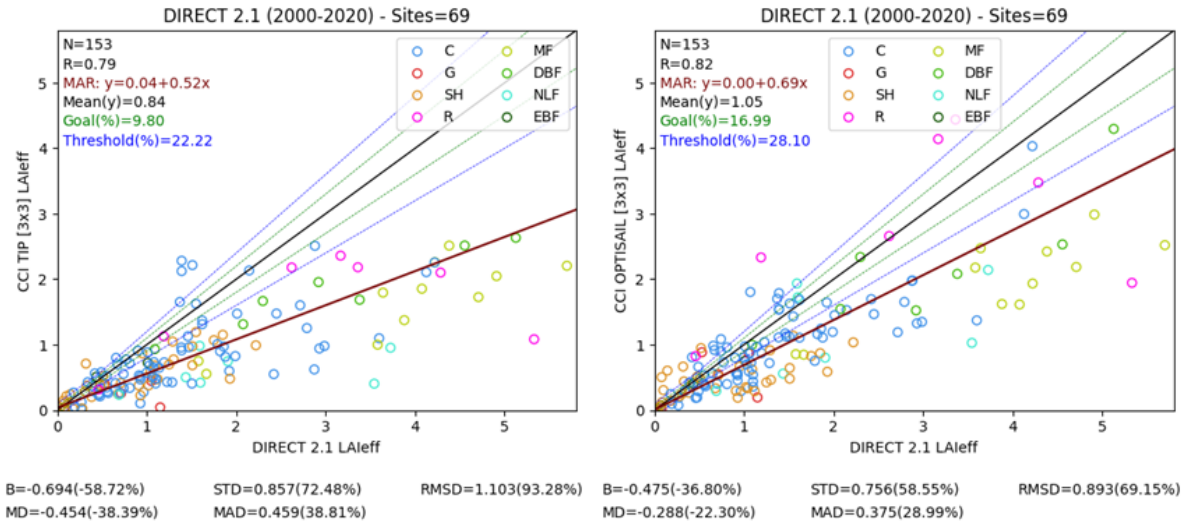

Figure 7: Scatterplots between TIP (left) and OptiSAIL (right) LAI products versus DIRECT V2.1 effective LAI ground-based maps. 'C' stands for cultivated, 'G' for grasslands, 'SH' for shrublands, 'R' for rice, 'MF' for mixed forests, 'DBF' for deciduous broadleaved forests, 'NLF' for needle-leaf forests and 'EBF' for evergreen broadleaved forests. Green and blue lines stand for goal and threshold levels, respectively.

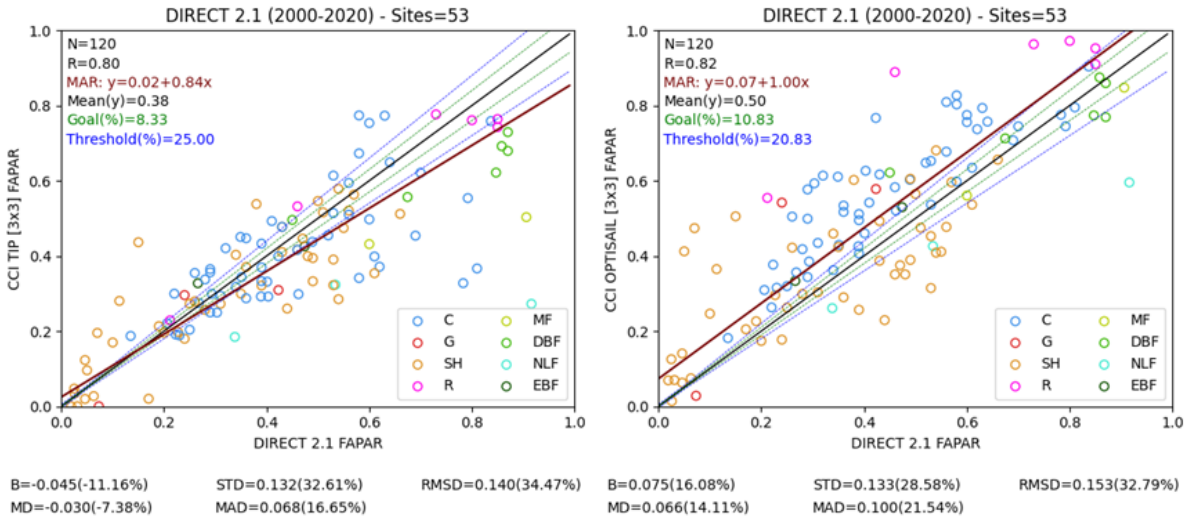

Figure 8: As in Figure 7 for fAPAR.

Figure 9 and Figure 10 show the scatterplots between TIP and OptiSAIL LAI and fAPAR products versus GBOV V3 LAI and fAPAR ground-based reference maps. Concomitant 'best quality' samples between both satellite products under study are used, and comparison was also performed at 3km^2 (i.e., average values of 3x3 pixels). The results for LAI are presented for forest and for non-forest sites separately due to the different level of clumping. Therefore, larger discrepancies due to the different definition (effective LAI in case of TIP and OptiSAIL and actual LAI in case of GBOV V3) are expected in those sites with higher clumping (i.e., over forests).

The findings for LAI are:

- For forest sites (Figure 9, top side) OptiSAIL shows better agreement than TIP in terms of correlation (R of 0.89 vs 0.77) and shows lower uncertainties (RMSD of 1.56 vs 2.16), reaching higher values (slope of 0.82 vs 0.39). Considering that the clumping index in forest is typically about 0.6-0.7 (Chen et al., 2005), the LAIeff values of OptiSAIL are more realistic than those of TIP.
- For non-forest sites (Figure 8, bottom side) OptiSAIL shows slightly better overall agreement than TIP for all validation metrics: accuracy (mean bias of 0.11 vs 0.13), precision (STD of 0.27 vs 0.31) and uncertainty (RMSD of 0.29 vs 0.33). OptiSAIL provides significantly high percentage of cases within goal (25%) and threshold (48%) GCOS uncertainty requirements than TIP (17% and 33%).

The findings for fAPAR are:

- OptiSAIL (RMSD = 0.14) provides better overall uncertainty than TIP (RMSD = 0.19), and better correlation (R of 0.82 vs 0.74).
- TIP provides systematic positive bias for lower fAPAR ranges (fAPAR < 0.3) and the opposite trend for higher ranges (fAPAR >0.6). OptiSAIL shows slight positive mean bias of 0.02, which is mainly observed for non-forest sites.
- Both satellite products provide higher values than GBOV V3 for non-forest sites, which is the same tendency to that found for other satellite products produced in CGLS, LSA SAF or MODIS services [VP-CCI_D4.1_PVIR] and could be partly due to GBOV V3 uncertainties.
- As for LAI, OptiSAIL provides significantly high percentage of cases within goal (18%) and threshold (33%) GCOS uncertainty requirements than TIP (4% and 9%).

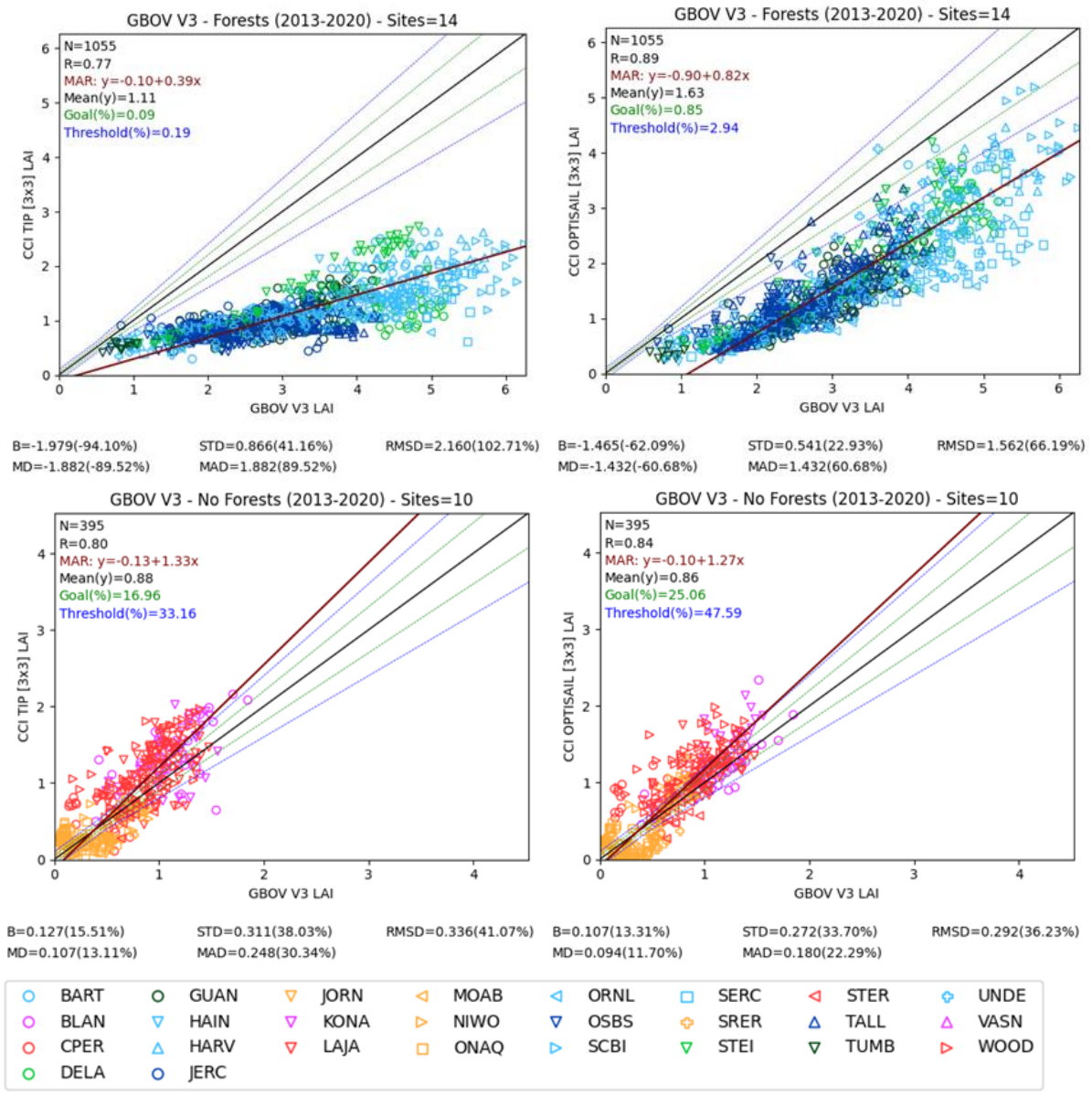

Figure 9: Scatterplots between TIP (left) and OptiSAIL (right) LAI products versus GBOV V3 LAI ground-based maps. Forest sites are presented at the top (dark and light green represent EBF and DBF, dark and light blue represent NLF and mixed forests) and non-forest sites at the bottom side (purple, red and orange represent croplands, grasslands and shrublands). Green and blue lines stand for goal and threshold levels, respectively.

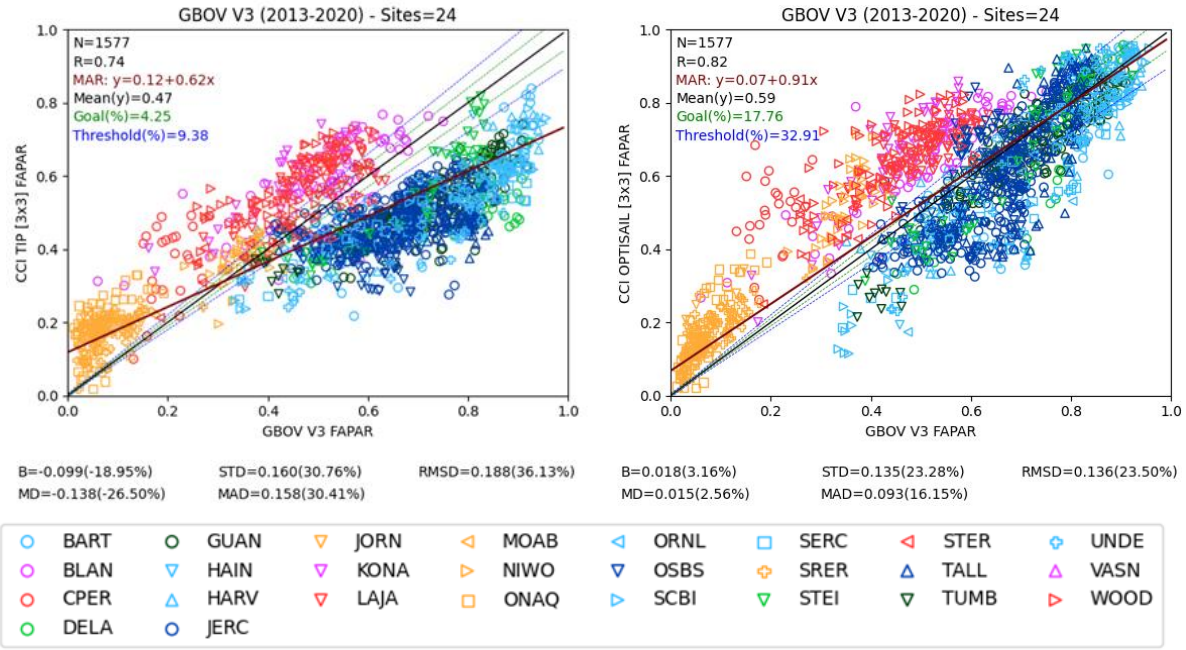

Figure 10: Scatterplots between TIP (left) and OptiSAIL (right) fAPAR products versus GBOV V3 fAPAR ground-based maps. Dark and light green represent EBF and DBF, dark and light blue represent NLF and mixed forests, and purple, red and orange stand for croplands, grasslands and shrublands. Green and blue lines stand for goal and threshold levels, respectively.

Figure 11 and Figure 12 show the scatterplots between TIP and OptiSAIL LAI and fAPAR products versus AMMA LAIeff and fAPAR ground data. Concomitant 'best quality' samples between both satellite products under study are used, and comparison was performed at 1km^2 of spatial support as AMMA ground measurements are provided over transects of around 1km.

The findings for LAI are (Figure 11):

- OptiSAIL shows better agreement than TIP in terms overall uncertainty (RMSD of 0.3 vs 0.4), accuracy (mean bias close to zero vs -0.2) and linear relationship. Despite to the fact that both satellite products are compared with ground-based effective LAI maps, TIP provides large negative biases for the higher LAI values (slope of 0.52).
- OptiSAIL provides high percentage of cases within goal (29%) and target (46%) GCOS uncertainty requirements than TIP (16% and 31%).

The findings for fAPAR are (Figure 12):

- OptiSAIL shows, again, better agreement than TIP in the comparison with AMMA ground measurements in terms of overall uncertainty (RMSD of 0.15 vs 0.17) and accuracy (positive mean bias of 0.5 in case of OptiSAIL and negative mean bias of -0.1 in case of TIP). The positive bias of OptiSAIL is mainly observed for higher fAPAR ranges (slope of 1.23).
- OptiSAIL also provides high percentage of cases within goal (6%) and target (13%) GCOS uncertainty requirements than TIP (3% and 7%).

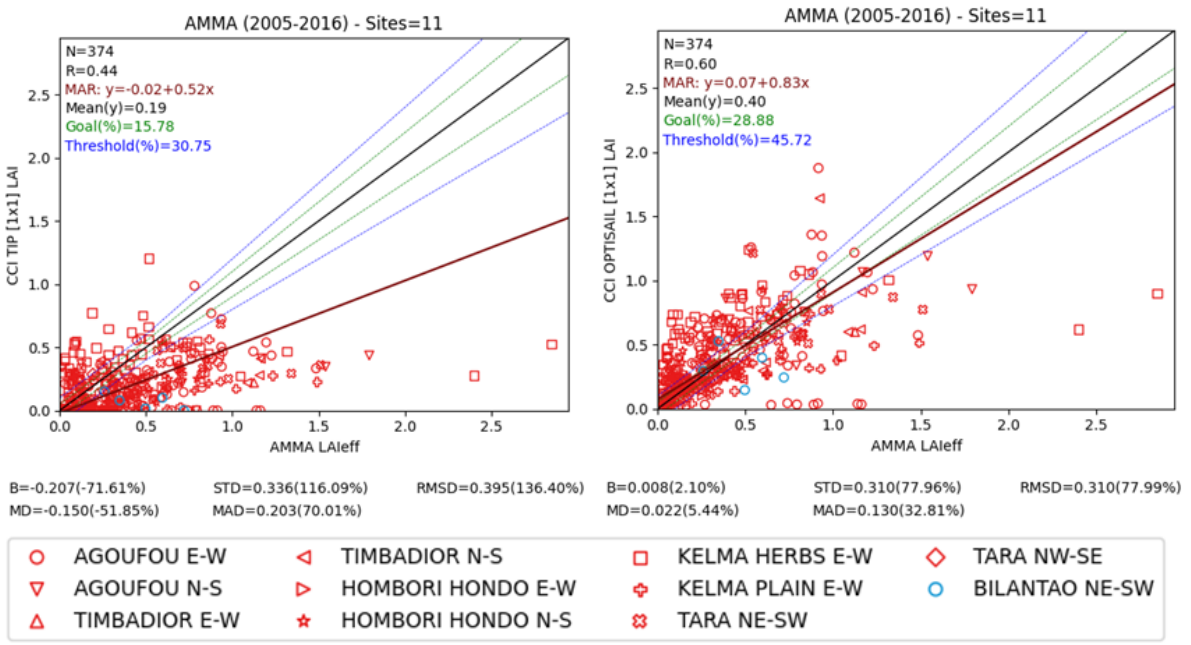

Figure 11: Scatter-plots between TIP (left) and OptiSAIL (right) LAI products versus AMMA LAI ground data. Green and blue lines stand for goal and threshold levels, respectively.

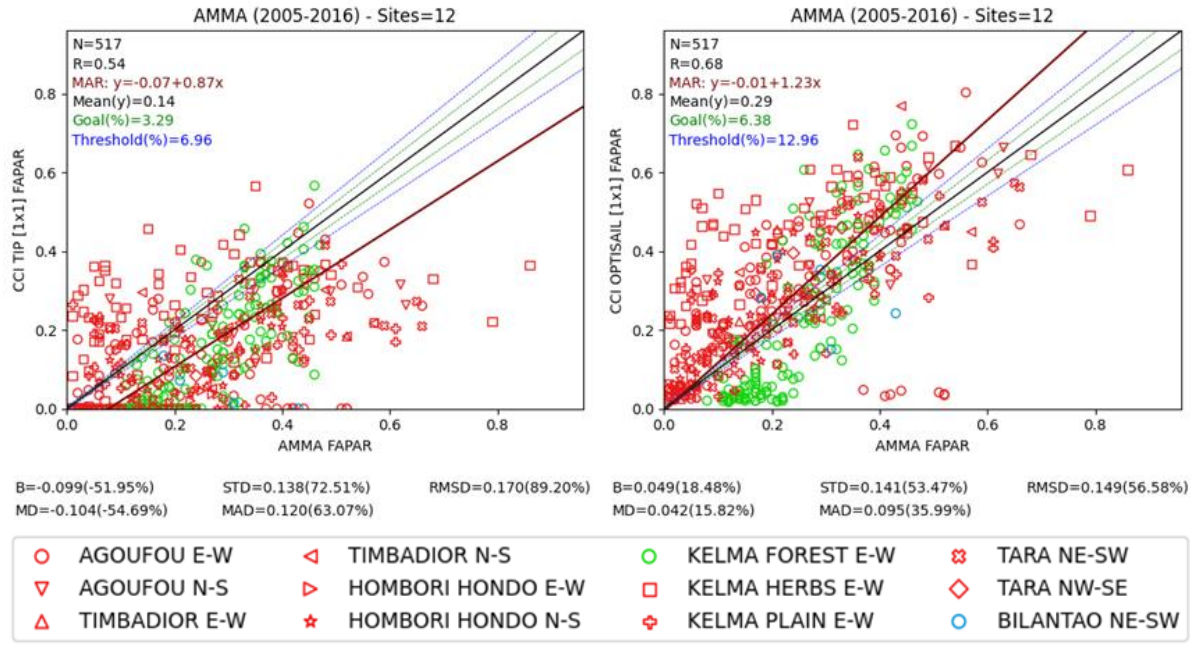

Figure 12: As in Figure 11 for fAPAR.

5.1.3 Error evaluation

The overall consistency between pair of satellite products (TIP, OptiSAIL and CGLS V2) is evaluated over best quality retrievals of LANDVAL sites during 2004-2005 (SPOT/VGT) and 2019 (PROBA-V). Figure 13 and Figure 14 show the scatterplots between pair of products for LAI and fAPAR respectively.

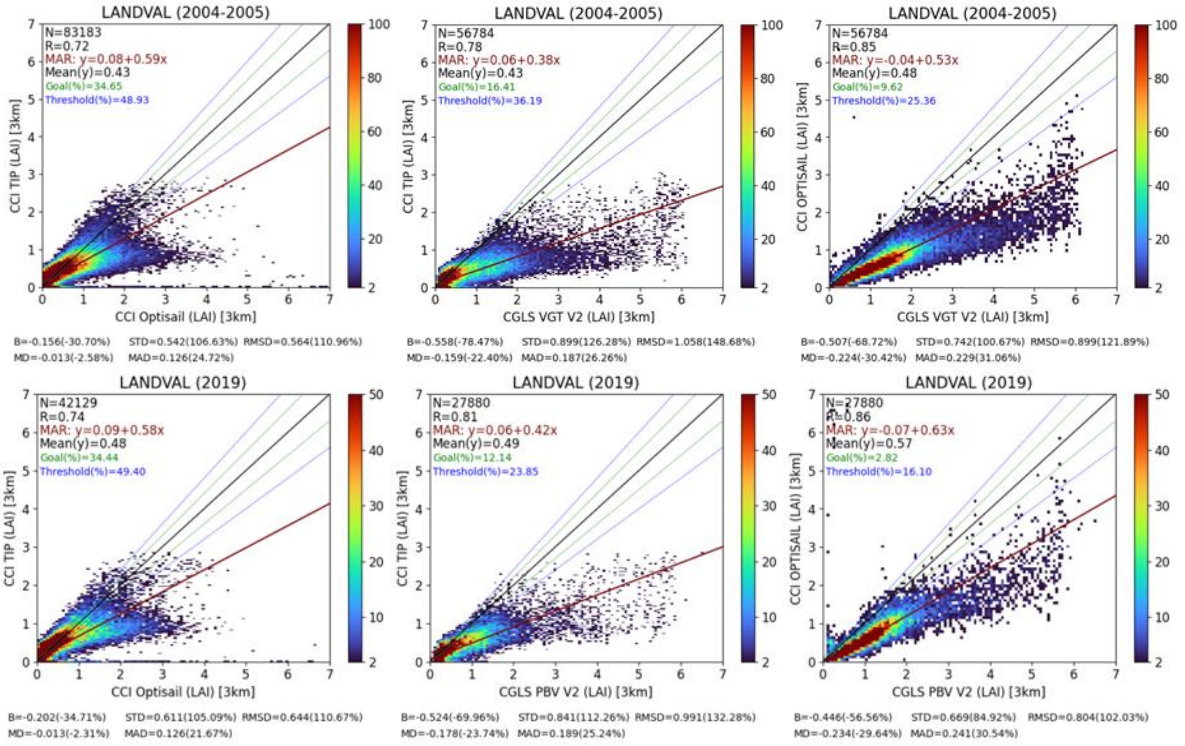

Figure 13: Scatter-plots between pair of satellite LAI products (colorbar represents density of points). Computation over best quality retrievals over LANDVAL sites for 2004-2005 (Top) and 2019 (bottom). From left to right: TIP versus OptiSAIL, TIP vs CGLS V2 and OptiSAIL vs CGLS V2. Green and blue lines stand for goal and threshold levels, respectively.

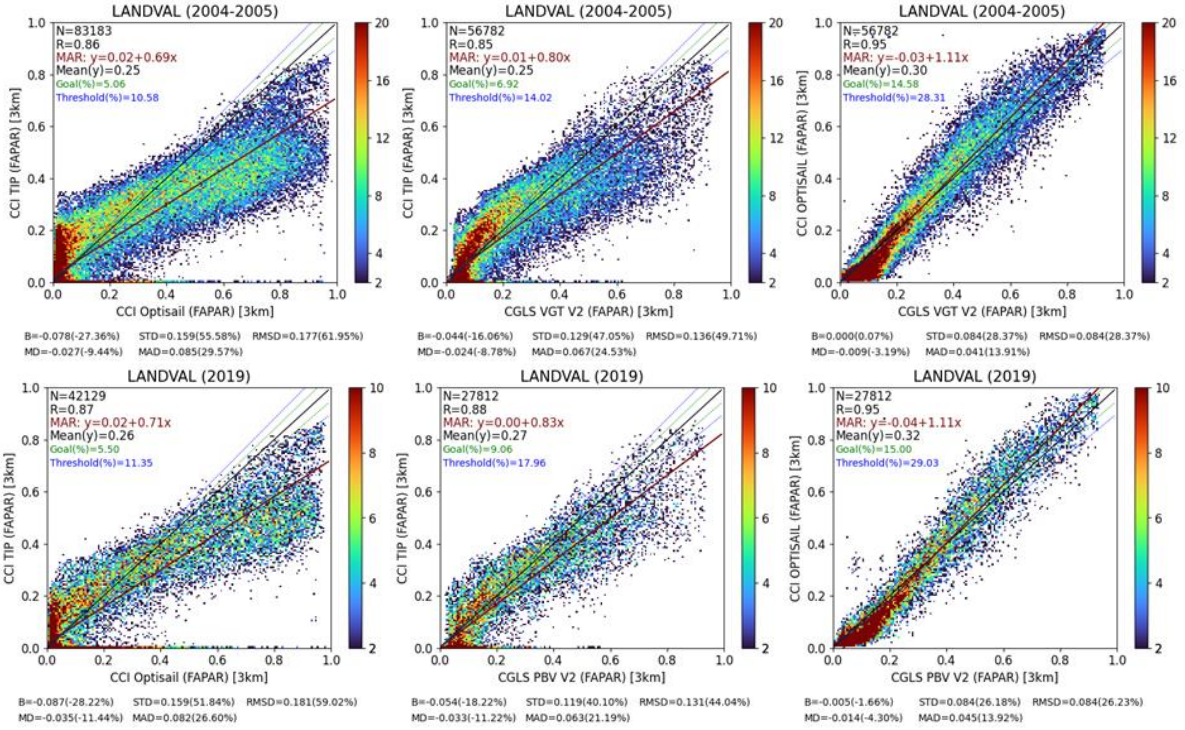

Figure 14: As in Figure 13 for fAPAR products.

The findings for LAI are (Figure 13):

- Large discrepancies between TIP and OptiSAIL (RMSD around 0.6) are found with OptiSAIL typically providing higher values (mean biases of around 0.2). It should be noted the high density of unrealistic zero values in case of TIP.
- Large differences with CGLS V2 due to the different LAI definitions (true LAI values in case of CGLS and effective LAI in case of TIP/OptiSAIL). OptiSAIL shows better agreement in the comparison with CGLS for all validation metrics (correlation, linear regression, bias, scattering, RMSD).
- Very similar results and validation metrics are found when the comparisons are performed for SPOT/VGT (Figure 13, top) or for PROBA-V (Figure 13, bottom) input data.

The findings for fAPAR are (Figure 14):

- Large differences between TIP and OptiSAIL with an RMSD value of 0.18. TIP provides higher values than OptiSAIL for the lower ranges and the opposite trend is found for the higher ranges. Again, TIP shows high density of unrealistic zero values.
- The comparison with CGLS V2 shows that OptiSAIL shows better agreement (RMSD = 0.08) than TIP (RMSD = 0.13). TIP shows negative bias mainly observed for higher fAPAR ranges whereas OptiSAIL shows almost no mean bias compared with CGLS V2.
- Very similar results and validation metrics are found when the comparisons are performed for SPOT/VGT (Figure 14, top) or PROBA-V (Figure 14, bottom) input data.

The histograms of LAI and fAPAR product values are analysed per biome type for TIP, OptiSAIL and CGLS V2 products, and presented in Figure 15 and Figure 16 respectively.

The findings for LAI are (Figure 15):

- TIP and OptiSAIL provide typically lower values than CGLS V2 for forest sites, as expected due to clumping index mainly impacts in dense canopies. OptiSAIL provides high frequency of LAI values towards higher ranges than TIP, mainly for EBF and NLF, which seems to be more realistic considering the clumping values reported by Chen et al. (2005) for these biomes.
- Similar distributions of LAI retrievals are found between both CCI proposed algorithms (TIP and OptiSAIL) and CGLS V2 for non-forest sites, except for cultivated.

The findings for fAPAR are (Figure 16):

- OptiSAIL and CGLS V2 provide very similar distribution of retrievals for forest cases and cultivated whereas TIP is typically shifted towards lower values.
- Some discrepancies in the distribution of values are found between the three products for herbaceous and shrublands whereas very similar distributions are found for sparse vegetated and bare areas (SBA).

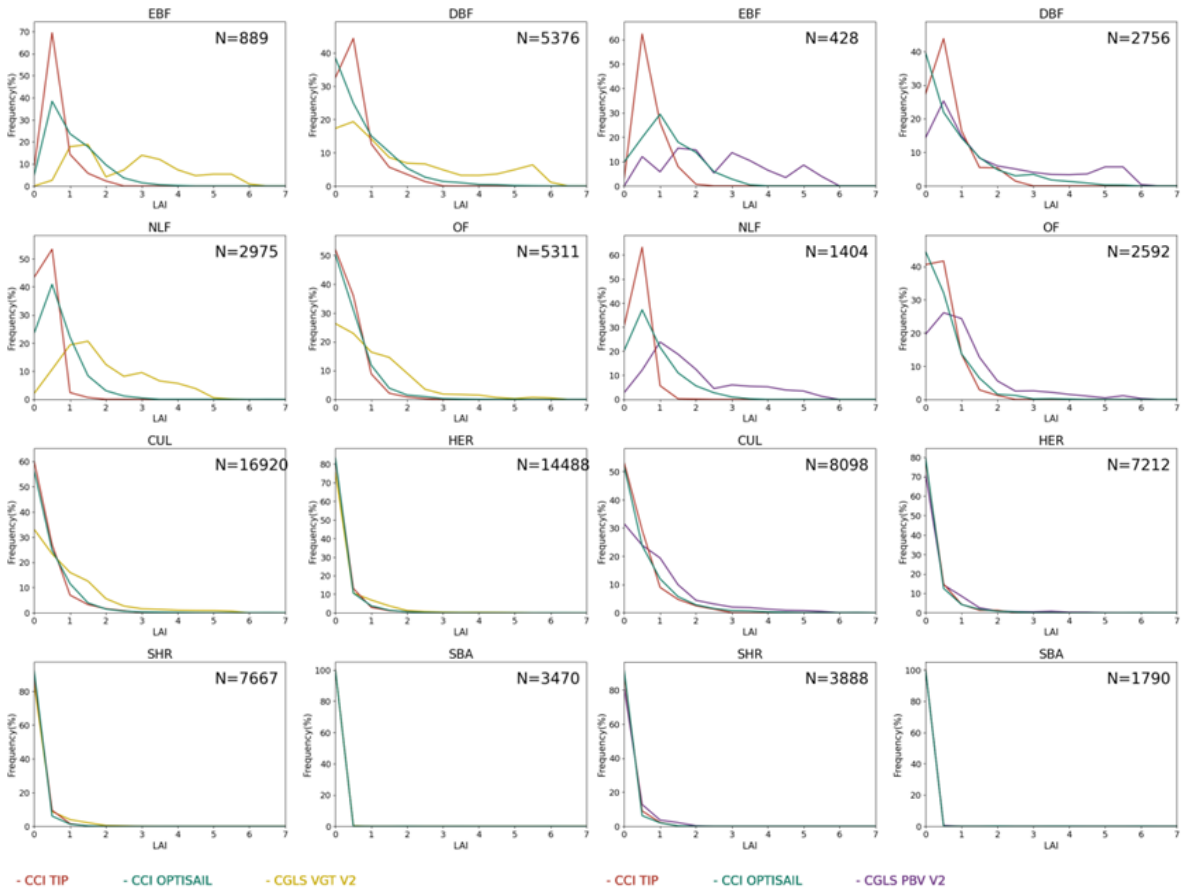

Figure 15: Distribution of LAI values for TIP, OptiSAIL and CGLS V2 products per main biome type for two different periods depending of input data: 2004-2005 SPOT/VGT (left) and 2019 PROBA-V (right).

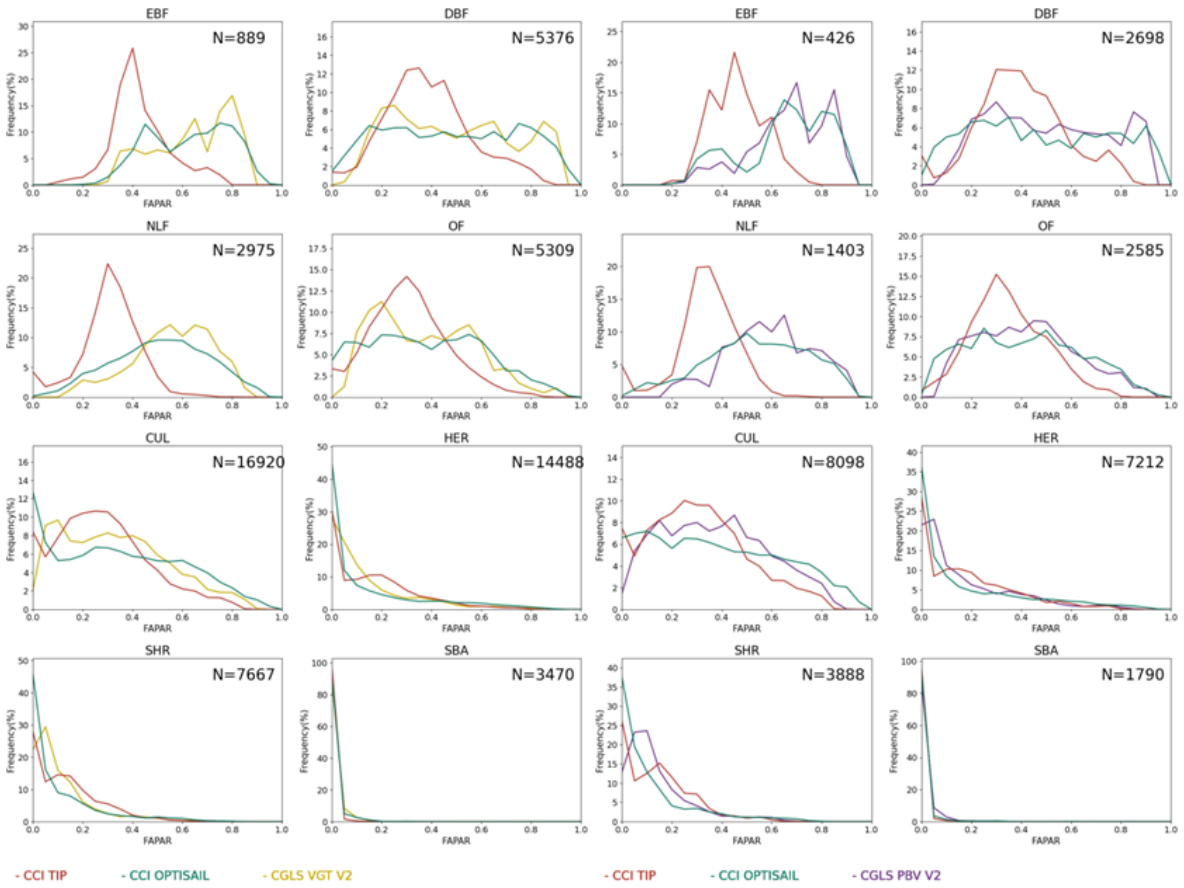

Figure 16: As in Figure 15 for fAPAR products.

5.1.4 Intra-annual Precision

The histograms of the intra-annual precision (δ, the so-called smoothness) for TIP and OptiSAIL are presented in Figure 17 and Figure 18 for two different periods: 2004-2005 (SPOT-VGT) and 2019 (PROBA-V). The computation is performed over LANDVAL sites and median δ values are provided as indicative of the intra-annual precision of the products. The main conclusions are:

- Both algorithms show similar distributions of δ values, but OptiSAIL provides, in overall, lower δ values than TIP, which is translated in high stability at short time scale and confirming the main findings of the visual inspections of temporal trajectories.
- Median δ values are lower in case of OptiSAIL even when it provides higher values than TIP.

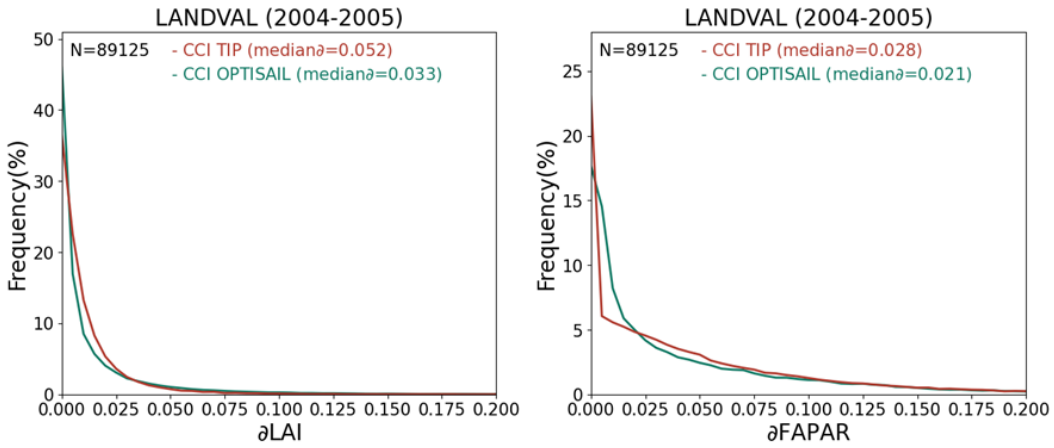

Figure 17: Histograms of delta function (smoothness) for TIP (red) and OptiSAIL (green) LAI (left) and fAPAR (right) over LANDVAL sites during the 2004-2005 period. Computation for all LANDVAL sites.

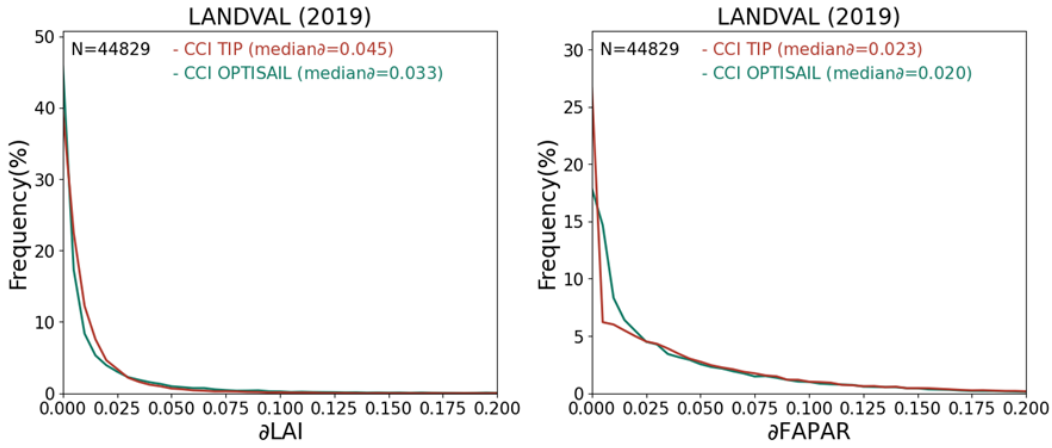

Figure 18 Histograms of delta function (smoothness) for TIP (red) and OptiSAIL (green) LAI (left) and fAPAR (right) over LANDVAL sites during the year 2019. Computation for all LANDVAL sites.

5.1.5 Inter-annual Precision

To investigate the inter-annual precision of TIP and OptiSAIL, box-plots per bin value of absolute inter-annual anomalies (year 2005 versus 2004) of the products under study were computed using the upper 95th and lower 5th percentiles over LANDVAL sites. Figure 19 and Figure 20 show the inter-annual precision for LAI and fAPAR products. The median of the absolute anomaly is proposed as overall indicator of inter-annual precision.

The main conclusions are:

- As expected, OptiSAIL shows slightly higher anomalies than TIP as it provides higher LAI and fAPAR values.
- In relative terms, both products provide very similar values, with inter-annual precision of around 5% for LAI and fAPAR.

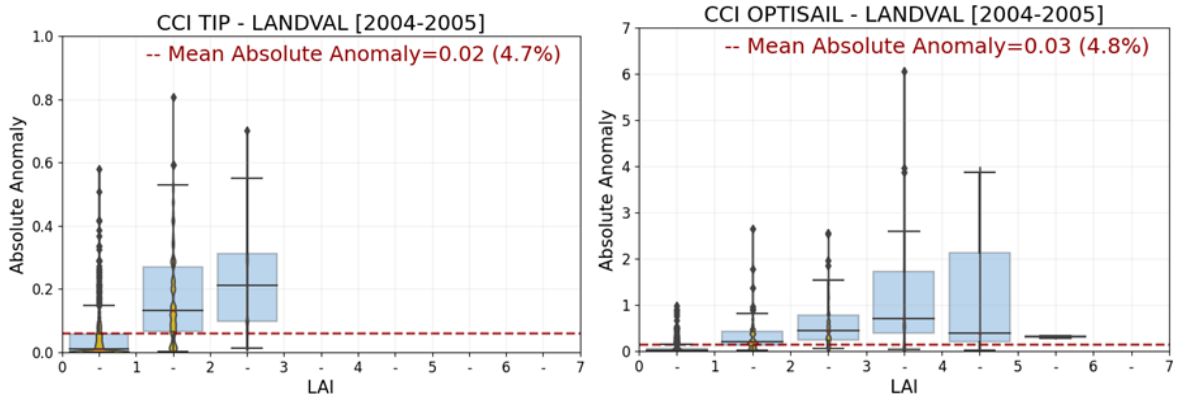

Figure 19: Box-plots of inter-annual absolute anomalies of TIP and OptiSAIL (year 2005 vs year 2004) per bin LAI value. Black bars in each box indicate median values and the dashed red line corresponds to the median absolute anomaly including all LAI ranges.

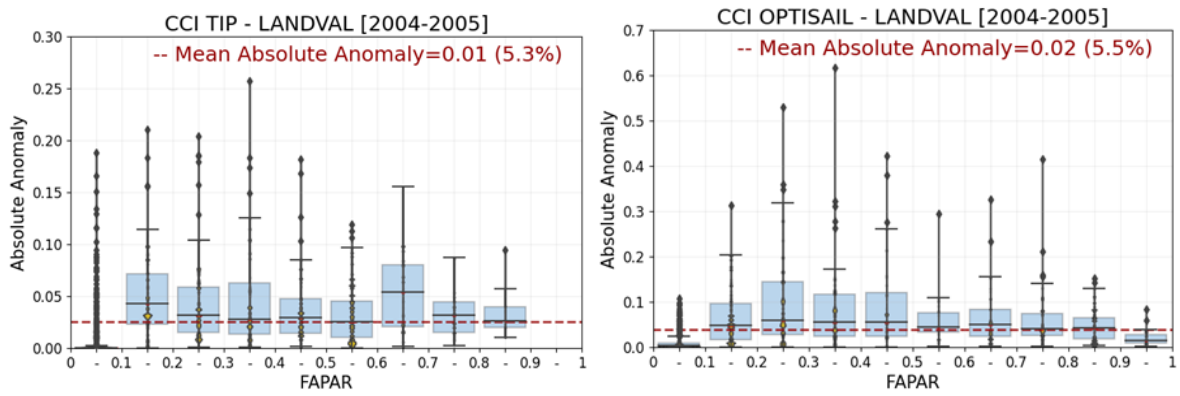

Figure 20: As in Figure 19 for fAPAR.

5.2 Implementation aspects

5.2.1 Processing time or computational demand

The computational demands of OptiAlbedo+TIP and OptiSAIL were evaluated by processing a set of tiles on the CCI Vegetation cluster processing environment and analysing the tracked performance metrics. This made it possible to draw conclusions on the behaviour of the algorithm implementations in a real-world scenario rather than on isolated runs.

The CCI Vegetation processing environment is part of a larger Hadoop/Spark cluster [VP-CCI_D3.1_SSD], with a dedicated slice of resources reserved for the project. The resources currently available are:

- 678 virtual CPU cores, across multiple processing nodes, with different hardware capabilities
- 939 GiB of total virtual memory, so about 1.38 GiB per virtual CPU core

As generated by the CCI Vegetation processing chain, a 1-year long timeseries with a 5-day interval contains 73 dates (74 for leap years). For both the OptiAlbedo+TIP and the OptiSAIL algorithm, all 73 dates for a tile can be processed in parallel. With the 678 available CPU cores, in principle 9 tiles can be processed simultaneously. However, at the start of a Hadoop/Spark process, an upper memory limit has to be specified for distributed jobs, which is then reserved per job. If a job exceeds this memory limit, it is pre-emptively terminated by the scheduling software. The upper memory limit is chosen sufficiently high to prevent this. Because the memory limit per job is also used by the scheduling software to check that the total memory reserved by all jobs does not exceed the total available virtual memory of the cluster, the value of the memory limit determines which part of the available virtual CPU cores are used. A (too) high means value results in not all available virtual CPU cores being used, because the number of parallel jobs will be limited.

It is notable that OptiSAIL has a clear advantage over OptiAlbedo+TIP in this respect, because it can use multiple CPU cores per job, without requiring more memory. In the current setup, OptiSAIL is configured to use 4 CPU cores. The memory available per virtual CPU core in the processing environment is, although reasonably large, often not sufficient when working with the large raster data of tiles, so that more memory needs to be reserved. Because OptiSAIL uses multiple CPU cores per job, it can make better use of the available CPU cores, without exceeding the total available virtual cluster memory. As an additional benefit, this reduces the total simultaneous I/O calls to the shared storage in the cluster, because fewer parallel jobs are reading and writing data (that are using multiple CPU cores for processing).

Table 5 shows the time required to process one timeseries date for one tile, per individual component of the OptiAlbedo+TIP and OptiSAIL algorithms, as recorded during processing of 21 yearly timeseries (2000 to 2020), for 11 tiles. OptiSAIL clearly takes a substantially longer computation time, even considering that OptiSAIL uses 4 CPU cores, while TIP and OptiAlbedo use only 1.

Table 5: Processing time for one timeseries date of one tile, per individual processing chain component.

Algorithm	Job	Minimum [minutes]	Maximum [minutes]	Mean [minutes]	Median [minutes]
Optialbedo+TIP	TIP	0.02	6.08	1.28	1.18
Optialbedo+TIP	OptiAlbedo	0.01	11.56	3.89	3.55
OptiSAIL	OptiSAIL	0.12	219.29	48.01	46.47

The OptiAlbedo and TIP components are always run sequentially, so we can simply combine their times, as shown in Table 6.

Table 6: Processing time for one date of one tile, per processing chain.

Algorithm	Minimum [minutes]	Maximum [minutes]	Mean [minutes]	Median [minutes]
OptiAlbedo+TIP	0.03	17.64	5.17	4.73
OptiSAIL	0.12	219.29	48.01	46.47

Table 7 shows the mean processing time for one date of one tile, over all tiles and per year. For OptiSAIL, years that include multiple sensors take a bit longer, but not dramatically. There is no actual overlap in VGT1 and VGT2 data, but OptiAlbedo+TIP and OptiSAIL use a 10-days window for input data selection, and both VGT1 and VGT2 can occur together in that window in 2003. However, this has no noticeable impact on processing time. For the period 2013-10-15 to 2014-06-02 the Proba-V and VGT2 data overlap.

Table 7: Mean processing time for one date of one tile, over all tiles, per processing chain and per year.

	Mean Time [minutes]	
Job	**OptiAlbedo+TIP**	**OptiSAIL**
Year		
2000	5.37	46.53
2001	5.76	46.82
2002	5.80	45.13
2003	6.14	44.58
2004	4.29	41.84
2005	4.64	44.09
2006	5.39	48.60
2007	5.53	49.67
2008	6.61	49.06
2009	5.82	50.46
2010	5.90	48.74
2011	5.36	47.10
2012	5.35	46.92
2013	**5.33**	**57.94**

2014	**5.17**	**64.59**
2015	4.01	45.37
2016	4.11	45.59
2017	4.23	44.41
2018	4.30	46.34
2019	4.30	46.47

Table 8 shows some variation among the timing results for the individual tiles, probably depending on the number of pixels actually processed.

Table 8: Mean processing time for one date of one tile, over all years, per processing chain and per tile.

	Mean Time [minutes]	
Job	**OptiAlbedo+TIP**	**OptiSAIL**
Tile		
X18Y02	5.87	42.23
X20Y00	3.70	18.76
X20Y01	5.52	33.83
X20Y02	6.60	53.04
X20Y03	5.22	38.13
X20Y04	4.61	43.49
X20Y06	4.95	62.73
X20Y07	4.90	57.00
X20Y08	5.22	58.12
X20Y09	5.33	64.37
X20Y10	4.97	56.36

However, due to the limited availability of CPU cores and memory in the processing cluster environment, potentially not all timeseries dates/jobs for a tile will be scheduled at the same time (they are all submitted for processing at the same time but may have to wait for resources). So, although in principle all jobs run in parallel, the average processing time for an individual job is not necessarily a good indicator of how long it typically takes to process a tile. As an illustration, see the OptiSAIL processing timeline graphs in

Figure 21, where the blue lines represent the running jobs for a tile.

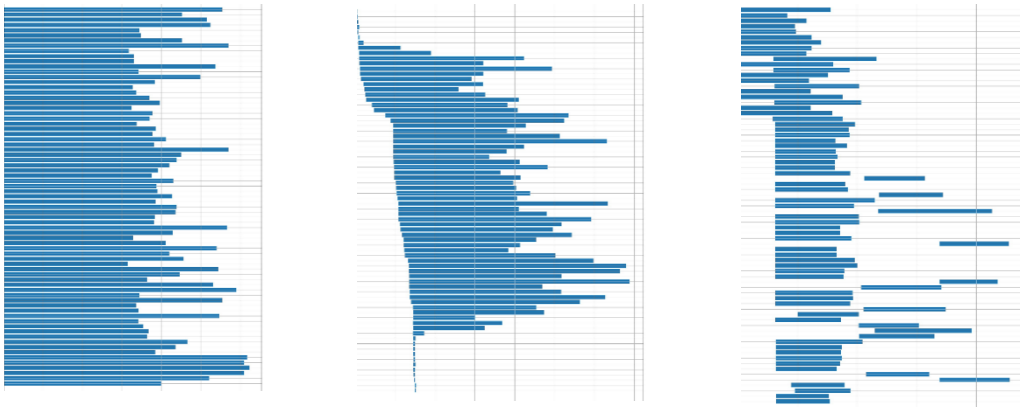

Figure 21: Examples of distribution of running jobs on the cluster. The blue lines represent the running jobs for a tile.

This can be taken into account by checking the actual time it takes to process a complete time series of 1 year for 1 tile, from the start of the first job to end of the last job (where each job calculates one date in the timeseries), as shown in Table 9.

Table 9: Actual processing time to complete processing of 1 tile for the full time series.

Algorithm	Minimum [minutes]	Maximum [minutes]	Mean [minutes]	Median [minutes]
OptiAlbedo+TIP	4.49	15.09	8.87	8.90
OptiSAIL	33.31	219.29	92.16	88.90

The time required to process the full 21 years timeseries (2000 to 2020) for 11 tiles (half the transect) took (from the start of the first job to the end of the last job) is shown in Table 10.

Table 10: Processing time required to process half of the transect (11 tiles).

Algorithm	Total Time [hours]	Total Time [days]
OptiAlbedo+TIP	5.84	0.24
OptiSAIL	88.80	3.70

By extrapolating this, an estimate can be made of the time required to process all 397 non-water tiles in the global CCI Vegetation tile grid (Table 11).

Table 11: Extrapolation of the processing time for global processing.

Algorithm	Estimated Time [days]	Estimated Time [weeks]	Estimated Time [months]
OptiAlbedo+TIP	8.66	1.24	0.29
OptiSAIL	133.54	19.08	4.45

Given the current processing environment, processing the full 21-years timeseries globally can be achieved in 4.5 months with OptiSAIL and 1.5 weeks with OptiAlbedo+TIP. Obviously processing for both algorithms scales very well, such that adding more cluster resources should have a more or less

linear impact on the total processing time. The computation time is given for the 1 km resolution data set. For the 300 m data set, these figures should be multiplied by 9.

5.2.2 Memory usage

Table 12 shows the memory used for processing one date of one tile, per individual component of the OptiAlbedo+TIP and OptiSAIL algorithms, as recorded during processing of 21 yearly timeseries (2000 to 2020), for 11 tiles. Note that the results are obtained for the processing settings used to generate the datasets, but in practice the memory use is configurable for all algorithms.

Table 12: Memory used to process one date of one tile, per individual processing chain component.

Algorithm	Job	Minimum [MiB]	Maximum [MiB]	Mean [MiB]	Median [MiB]
OptiAlbedo+TIP	Optialbedo	51.07	1074.31	380.32	369.34
OptiAlbedo+TIP	TIP	8.00	1432.41	1373.26	1428.73
OptiSAIL	OptiSAIL	922.32	1955.31	1256.84	1244.64

OptiAlbedo and TIP always run sequentially, so we are only really interested in the process that uses the most memory, which from our measurements is clearly TIP, not OptiAlbedo. These combined results can be found in Table 13 below.

Table 13: Memory used to process on date of one tile, per processing chain.

Algorithm	Minimum [MiB]	Maximum [MiB]	Mean [MiB]	Median [MiB]
OptiAlbedo+TIP	51.07	1432.41	1373.26	1428.73
OptiSAIL	922.32	1955.31	1256.84	1244.64

The amount of memory needed by OptiAlbedo+TIP and OptiSAIL is comparable. Also keep in mind that, as described in the Computational Demand section (5.2.1), OptiSAIL is using multiple CPU cores to process a single timeseries date, and thus needs less memory per CPU core. Therefore, it is better able to use all available CPU cores in the cluster processing environment.

As shown in Table 14, for OptiSAIL, the maximum memory used increases for years where there are multiple sensors used (2013-2014), but this is not problematic.

Table 14: Maximum memory used to process one date of one tile, over all tiles and per year.

	Maximum [MiB]	
Job	**OptiAlbedo+TIP**	**OptiSAIL**
Year		
2000	1431.58	1383.12
2001	1431.00	1382.94
2002	1430.94	1348.70
2003	1430.73	1365.66
2004	1430.83	1365.89
2005	1431.73	1348.77
2006	1431.31	1348.74

2007	1431.20	1349.11
2008	1432.41	1349.11
2009	1432.27	1349.15
2010	1431.31	1349.13
2011	1431.49	1348.75
2012	1431.39	1348.72
2013	**1430.81**	**1852.81**
2014	**1431.87**	**1955.31**
2015	1431.66	1541.49
2016	1431.27	1541.23
2017	1431.46	1541.27
2018	1431.11	1540.46
2019	1431.58	1524.27
2020	1431.29	1541.49

The maximum memory used by OptiSAIL also increases for tiles that have a larger number of pixels to process (less water pixels), as shown in Table 15 below.

Table 15: Maximum memory used to process one date of one tile, over all years and per tile.

	Maximum [MiB]	
Job	**OptiAlbedo+TIP**	**OptiSAIL**
Tile		
X18Y02	1431.20	1851.82
X20Y00	1432.41	1731.98
X20Y01	1431.10	1955.31
X20Y02	1431.03	1818.24
X20Y03	1430.52	1714.63
X20Y04	1429.91	1680.09
X20Y06	1430.43	1594.01
X20Y07	1431.58	1576.77
X20Y08	1431.66	1611.16
X20Y09	1431.64	1611.47
X20Y10	1429.91	1611.11

5.2.3 Stability

From our processing experience, the OptiAlbedo+TIP and OptiSAIL implementations both consistenly perform as expected in the CCI Vegetation cluster processing environment. They both handle input data errors as expected and have elaborate logging for backward traceability.

Input data collection and algorithm configuration is almost identical for both algorithm implementations. The OptiAlbedo+TIP algorithm uses two sequential processing steps, whereas OptiSAIL has only one step, so processing workflow complexity is slightly less for OptiSAIL. In that sense, there is less opportunity for issues in the processing environment with OptiSAIL.

5.2.4 Data volume

OptiSAIL generates significantly more output layers than OptiAlbedo+TIP, resulting in a larger volume of the generated dataset. The total amount of disk storage used for the 21- year long timeseries of the 11 tiles (half of the transect) is shown in Table 16. Note that the output layers of OptiSAIL are configurable, and if necessary, only a limited set of output layers can be generated.

Table 16: Total amount of storage needed for the entire time series and calculated over 11 tiles.

Algorithm	Size [GiB]	Size [TiB]
OptiAlbedo+TIP	229.45	0.22
OptiSAIL	3489.32	3.41

The amount of storage required for 11 tiles is relatively stable for all 21 years, as can be seen in Table 17. The year 2020 has only half the amount of data, because the Proba-V mission ends at 2020-06-30.

Table 17: Amount of storage needed per year calculated for 11 tiles.

	Storage [GiB]	
Algorithm	**OptiAlbedo+TIP**	**OptiSAIL**
Year		
2000	11.08	169.49
2001	10.99	167.17
2002	10.93	166.07
2003	10.90	164.63
2004	11.03	167.46
2005	10.93	166.14
2006	10.94	167.08
2007	10.79	164.06
2008	10.99	167.70
2009	10.89	167.28
2010	10.94	166.93
2011	10.90	166.16
2012	10.98	167.27
2013	10.98	170.38
2014	11.53	179.36
2015	11.67	173.42
2016	11.95	178.73
2017	11.71	174.74
2018	11.65	175.85
2019	11.61	176.70
2020	6.07	92.69

Tiles with a larger number of pixels to be processed take more space, as shown in Table 18.

Table 18: Amount of storage needed per tile calculated over the entire time series.

	Size [GiB]	
Algorithm	**OptiAlbedo+TIP**	**OptiSAIL**
Tile		
X18Y02	21.39	334.64
X20Y00	9.65	142.09
X20Y01	16.60	254.06
X20Y02	25.53	398.45
X20Y03	17.88	276.30
X20Y04	11.81	146.46
X20Y06	24.87	375.94
X20Y07	26.80	424.83
X20Y08	26.18	412.81
X20Y09	26.05	383.11
X20Y10	22.69	340.62

Table 19 shows the estimated storage needed for all 397 non-water tiles in the CCI Vegetation global tile grid, by extrapolation.

Table 19: Total amount of storage needed for global processing.

Algorithm	**Size (TiB)**
OptiAlbedo+TIP	8.09
OptiSAIL	122.98

After repackaging the OptiSAIL data for distribution, and thus removing non-essential output parameters, the storage requirements decrease significantly. The non-essential parameters include by-products of the retrieval such as soil and leaf properties and their covariances. For the tests, a 21-year long timeseries was processed for 23 tiles (full transect + 1 extra tile), and the total data volume on disk was 247 GiB. Table 20 shows the estimated storage needed for all 397 non-water tiles in the CCI Vegetation global tile grid, by extrapolation.

Table 20: Total amount of storage needed for the global time series after repackaging of the data.

Algorithm	**Size [TiB]**
OptiSAIL + repackaging	4.16

The storage requirements are given for the 1 km resolution data set. For the 300 m data set, these figures should be multiplied by 9.

5.2.5 Implementation risks

The processing workflows for OptiAlbedo+TIP and OptiSAIL are largely identical. They both use the same type of input data, configuration files, and expect a very similar runtime environment, so there is no increase in implementation risk for choosing one algorithm over the other.

5.3 Qualitative user requirements

The user's wish to include effects of soil and snow is driven by their experience that algorithms that take this into account are more suitable to represent LAI and fAPAR in the early and late stages of the growing cycle. In sparse vegetation or vegetation types at the onset of the growing season, the 'freedom' in representing soil background spectra can prevent unrealistic LAI values, which is crucial for the detection of long-term changes in phenology: The identification of the start and end of the growing season, SOS and EOS, rely on the rise and fall of LAI or fAPAR in the seasonal cycle. Both algorithms consider the effects of snow and soil reflectance. In OptiAlbedo, BRDF kernels for the soil and snow are included. The explicit inclusion of a radiative transfer model for soil and snow reflectance (TARTES) in OptiSAIL results in by-products of snow depth and soil reflectance parameters. Although these are not validated, they can be further inspected, and can be physically meaningful output diagnostics.

The user interest in differentiation of fAPAR by pigments is driven by recent developments of canopy chlorophyll products, which are expected to have a closer relation with photochemical and biochemical processes than total fAPAR. OptiSAIL can de-couple the effects of seasonally chlorophyll content and varying LAI on FAPAR, through the additional output of chlorophyll content, and the specific absorption spectra of the PROSPECT model. The fraction of PAR absorbed by chlorophyll is a diagnostic variable computed in a forward modelling step with PROSAIL following the retrieval. OptiAlbedo-TIP cannot provide this output.

Several land surface models, such as CLM (Lee et al., 2015) and Bethy (Norton et al, 2018) now include the forward simulation of solar induced chlorophyll fluorescence (SIF), which enables the assimilation of satellite observations of SIF for model exchange estimates. These models include a photochemical, biophysical and biochemical representation of the relationship between SIF and photosynthesis in leaves. For the scaling from leaf to the stand and for representing the directionality of SIF, transfer functions trained with the model SCOPE have been used. Parazoo et al. (2020) demonstrated that wide discrepancies exist between SIF representations among models, partly due to illumination-observation geometry and (re-)absorption of SIF. OptiSAIL has the potential to extend the radiative transfer representation with SIF as well. This would open the possibility to provide geometry and reabsorption and wavelength normalizations of SIF from observations of different platforms, thus providing homogenization of SIF data. OptiAlbedo does not provide this opportunity.

Land surface models differ in the way they use FAPAR, depending on the treatment of direct and diffuse radiation. White sky fAPAR (for diffuse radiation) is independent on the solar angle, black sky fAPAR (for direct solar radiation) varies with the solar angle, and blue sky fAPAR varies with both solar angle and the fraction of diffuse radiation. With OptiSAIL it is possible to provide different fAPAR estimates in a diagnostic step (an additional forward simulation), such as diurnally integrated blue or black sky fAPAR, or a black sky fAPAR for a specific solar angle can be provided. OptiAlbedo-TIP cannot provide this output.

The validation demonstrated that clumping of vegetation causes the LAI derived with both TIP or OptiSAIL to underestimate true LAI in some land cover types, including needleleaf forests. For both algorithms, either the implementation of a clumping representation or a posterior correction for clumping is technically feasible. The required efforts to implement this aspect is similar for Opti-Albedo-TIP and OptiSAIL.

6 Conclusion and selection

From the validation point of view, OptiSAIL prototype algorithm reaches higher or similar performance than TIP in most of the validation criteria evaluated (see a summary in Table 21):

- Both products provide similar completeness but considering the quality flag OptiSAIL is less restrictive than TIP using SPOT/VGT input data.
- TIP and OptiSAIL are temporally consistent with reference datasets. TIP provides in some cases unrealistic zero values (mainly over grassland/shrubland cases). OptiSAIL shows, on the other hand, some outliers not identified by quality flags.
- The intra-annual precision shows better results for OptiSAIL as TIP provides some noise. Both products show similar inter-annual precision, with median absolute anomalies of around 5%.
- In the comparison with reference ground datasets (DIRECT V2.1, GBOV V3 and AMMA) and reference satellite product (CGLS V2), OptiSAIL provides higher correlation, better accuracy, and lower uncertainties that TIP for both LAI anf fAPAR products, resulting in larger fraction of samples meeting GCOS goal and threshold requirements. TIP tends to provide lower values than references, mainly for the higher ranges. LAIeff values of OptiSAIL are more realistic than LAIeff values of TIP either comparing to ground LAIeff or satellite LAI values (considering typicall clumping index reported in the literature for forests).

While it takes substantially longer to process using OptiSAIL compared to OptiAlbedo+TIP, the processing is within acceptable bounds. It should be feasible to generate a 21-year timeseries globally within a reasonable time frame. Additional processing resources can be allocated to speed up the processing. Memory requirements and stability are comparable between the algorithms. While the final data volume will be significantly larger for OptiSAIL, due to the larger number of output layers generated, the total expected volume is acceptable. In addition, the choice of output layers is configurable. After repackaging for distribution (removal of non-essential parameters) the OptiSAIL data volume is comparable to OptiAlbedo+TIP.

For the qualitative user requirements, OptiAlbedo-TIP meets part of the requirements (accounting for snow and soil background to ensure sensitivity at the onset and end of the growing season), OptiSAIL meets all requirements and has the flexibility for an extended product portfolio.

In conclusion, OptiSAIL outperforms OptiAlbedo-TIP in the validation and in the prospects for an extended product portfolio, while staying within the bounds of computational feasibility. OptSAIL is selected as the algorithm for further use in the Vegetation CCI project.

Table 21: Summary of product prototype algorithm validation results

Criteria	TIP performances	OptiSAIL performances	Comments
Product completeness	±	±	• TIP and OptiSAIL similar distribution of gaps, located in wintertime (northern hemisphere) and equatorial areas. • OptiSAIL slightly better for SPOT/VGT when using quality flags.
Temporal consistency	±	+	• Both products provide similar temporal trends than ground data and CGLS V2. • OptiSAIL>TIP for higher values (more realistic). • TIP is noisier than OptiSAIL. • TIP provides unrealistic zero values in some cases (mainly grassland cases). • OptiSAIL provides some outliers not identified by quality flags.
Error evaluation: Direct validation vs DIRECT V2.1	±	±	LAI: • TIP: B=-0.7, RMSD=1.1, %goal/threshold=10/22 • OptiSAIL: B=-0.5, RMSD=0.9, %goal/threshold=17/28 • Large negative bias in TIP. OptiSAIL more realistic for higher values. fAPAR : • TIP: B=-0.05, RMSD=0.14, %goal/threshold=8/25 • OptiSAIL: B=0.08, RMSD=0.15, %goal/threshold=10/21 • TIP overestimates higher values. OptiSAIL provides slightly positive systematic bias.
Error evaluation: Direct validation vs GBOV V3	-	±	LAI (forest): • OptiSAIL shows better agreement. Large underestimation for TIP. Almost no retrievals within threshold levels (different definition). LAI (non-forest): • TIP: B=0.13, RMSD=0.33, %goal/threshold=17/33 • OptiSAIL: B=0.11, RMSD=0.29, %goal/threshold=25/48 • OptiSAIL slightly better for all validation & APU metrics. fAPAR: • TIP: B=-0.1, RMSD=0.19, %goal/threshold=4/9 • OptiSAIL: B=0.02, RMSD=0.14, %goal/threshold18/33 • TIP > GBOV V3 for low fAPAR and <GBOV V3 for high. • OptiSAIL > GBOV V3 mainly for non-forest and almost no mean bias for forest high values.
Error evaluation: Direct validation vs AMMA	-	±	LAI: • TIP: B=-0.2, RMSD=0.4, %Ogoal/threshold=16/31 • OptiSAIL: B=0.01, RMSD=0.31, %Goal/threshold=29/46 • Large negative bias in TIP. OptiSAIL more realistic. fAPAR: • TIP: B=-0.1, RMSD=0.17, %Ogoal/threshold=3/7 • OptiSAIL: B=0.05, RMSD=0.15, %Goal/threshold=6/13 TIP < AMMA systematically. Low bias for OptiSAIL (slightly >AMMA for high values).
Error evaluation: Product intercomparison vs CGLS V2	-	±	LAI: • Large differences in both cases (LAI vs LAIeff) but OptiSAIL better agrees and provides more realistic values considering typical clumping index values. fAPAR : • TIP: B=-0.05, RMSD=0.13, %Goal/threshold=7-9/14-17 • OptiSAIL: B , RMSD=0.08, %Ogoal/threshold=15/28-29 TIP < CGLS V2 for high values. OptiSAIL almost no mean bias. Analysis per biome: • OptiSAIL is more realistic for forest cases, reaching higher values than TIP.
Intra-annual precision	± -	±	• OptiSAIL is smoother than TIP.

Inter-annual precision	+	+	• Similar for both products (around 5%)
Computation time	+	-	The required computation time of OptiSAIL is substantially higher (one order of magnitude), but still acceptable
Memory usage	+	±	OptiSAIL memory use is approximately 4x as large as that of OptiAlbedo-TIP
Stability	+	+	The processing with both algorithms is stable and reliable
Storage space	+	±	OptiSAIL storage requirements is larger than OptiAlbedo-TIP, but after packing and selecting output variables similar.
Ancillary output	±	+	OptiSAIL has the possibility to provide all relevant ancillary outputs identified by users. OptiAlbedo-TIPP can provide part of the additional outputs.

7 Risks and mitigation

In cycle 2, other sensors will be added to the processing chain. The output products are sensitive to the values of the uncertainty levels and the correlation between uncertainties. The estimation of uncertainties and in particular the correlations is a critical task. It may be necessary to make assumptions on the uncertainty correlations and adjustments (inflation) if not all sources of uncertainties are precisely quantifiable.
Cloud screening before and pre-selection of input is necessary, and in the algorithm for the CRDP-1, filters have been applied, and the screening can be optimized in cycle 2.

The computation speed is sensitive to the number of bands included. In cycle 1, this was limited to three observations per band and per sensor. With addition of new sensors, this may need to be adjusted.

The lack of postprocessing steps, primarily the temporal filtering and accounting for vegetation specific clumping, may influence the experience of the users who may be used to products that are gap filled, smoothed and for which land cover type specific prior information is used in the retrieval. Communication to the users through the climate research group will be necessary to mitigate this risk.

8 References

Baret, F., Weiss, M., Lacaze, R., Camacho, F., Makhmara, H., Pacholcyzk, P., Smets, B., 2013. GEOV1: LAI and FAPAR essential climate variables and FCOVER global time series capitalizing over existing products. Part1: Principles of development and production. Remote Sens. Environ. 137, 299–309. https://doi.org/10.1016/j.rse.2012.12.027

Chen, J.M., Menges, C. H., Leblanc. S. G., 2005. Global mapping of foliage clumping index using multi-angular satellite data. Remote Sensing of Environment 97 (2005) 447 – 457.

Fang, H. (2021). Canopy clumping index (CI): A review of methods, characteristics, and applications. *Agricultural and Forest Meteorology*, 303, 108374.

Féret, J. B., Gitelson, A. A., Noble, S. D., & Jacquemoud, S. (2017). PROSPECT-D: Towards modeling leaf optical properties through a complete lifecycle. *Remote Sensing of Environment*, *193*, 204-215.

Fernandes, R.A., Plummer, S.E., Nightingale, J., Baret, F., Camacho, F., Fang, H., Garrigues, S., Gobron, N., Lang, M., Lacaze, R., Leblanc, S.G., Meroni, M., Martinez, B., Nilson, T., Pinty, B., Pisek, J., Sonnentag, O., Verger, A., Welles, J.M., Weiss, M., Widlowski, J.-L., Schaepman-Strub, G., Román, M.O., Nicheson, J., 2014. Global Leaf Area Index Product Validation Good Practices. Version 2.0. In G. Schaepman-Strub, M. Román, & J. Nickeson (Eds.), Best Practice for Satellite-Derived Land Product Validation (p. 76): Land Product Validation Subgroup (WGCV/CEOS), doi:10.5067/do [WWW Document]. https://doi.org/10.5067/doc/ceoswgcv/lpv/lai.002

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 Sens. 12, 1017. https://doi.org/10.3390/rs12061017

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. J. Geophys. Res. *Biogeosciences* 113. https://doi.org/10.1029/2007JG000635

Harper, W. V., 2014. Reduced Major Axis regression: teaching alternatives to Least Squares. Proc. Ninth Int. Conf. Teach. Stat. 1–4. https://doi.org/10.1016/B978-0-12-420228-3.00013-0

Jiang, C., Ryu, Y., Fang, H., Myneni, R., Claverie, M., & Zhu, Z. (2017). Inconsistencies of interannual variability and trends in long-term satellite leaf area index products. *Global Change Biology*, *23*(10), 4133-4146.

Lee, J. E., Berry, J. A., van der Tol, C., Yang, X., Guanter, L., Damm, A., ... & Frankenberg, C. (2015). Simulations of chlorophyll fluorescence incorporated into the C ommunity L and M odel version 4. *Global change biology*, *21*(9), 3469-3477.

Morisette, J.T., Baret, F., Privette, J.L., Myneni, R.B., Nickeson, J.E., Garrigues, S., Shabanov, N. V., Weiss, M., Fernandes, R.A., Leblanc, S.G., Kalacska, M., Sánchez-Azofeifa, G.A., Chubey, M., Rivard, B., Stenberg, P., Rautiainen, M., Voipio, P., Manninen, T., Pilant, A.N., Lewis, T.E., Iiames, J.S., Colombo, R., Meroni, M., Busetto, L., Cohen, W.B., Turner, D.P., Warner, E.D., Petersen, G.W., Seufert, G., Cook, R., 2006. Validation of global moderate-resolution LAI products: A framework proposed within the CEOS land product validation subgroup. IEEE Trans. Geosci. Remote Sens. 44, 1804–1814. https://doi.org/10.1109/TGRS.2006.872529

Nilson, T. (1971). A theoretical analysis of the frequency of gaps in plant stands. *Agricultural meteorology*, 8, 25-38.

Norton, A. J., Rayner, P. J., Koffi, E. N., & Scholze, M. (2018). Assimilating solar-induced chlorophyll fluorescence into the terrestrial biosphere model BETHY-SCOPE v1. 0: model description and information content. *Geoscientific Model Development*, *11*(4), 1517-1536.

Parazoo, N. C., Magney, T., Norton, A., Raczka, B., Bacour, C., Maignan, F., ... & Frankenberg, C. (2020). Wide discrepancies in the magnitude and direction of modeled solar-induced chlorophyll fluorescence in response to light conditions. *Biogeosciences*, 17(13), 3733-3755.

Pinty, B., Lavergne, T., Dickinson, R. E., Widlowski, J. L., Gobron, N., & Verstraete, M. M. (2006). Simplifying the interaction of land surfaces with radiation for relating remote sensing products to climate models. *Journal of Geophysical Research: Atmospheres*, 111(D2).

Sánchez-Zapero, J., Camacho, F., Martínez-Sánchez, E., Lacaze, R., Carrer, D., Pinault, F., Benhadj, I., Muñoz-Sabater, J., 2020. Quality Assessment of PROBA-V Surface Albedo V1 for the Continuity of the Copernicus Climate Change Service. Remote Sens. 2020, Vol. 12, Page 2596 12, 2596. https://doi.org/10.3390/rs12162596

Skiles, S. M., & Painter, T. H. (2019). Toward understanding direct absorption and grain size feedbacks by dust radiative forcing in snow with coupled snow physical and radiative transfer modeling. *Water Resources Research*, *55*(8), 7362-7378.

Verger, A., Sánchez-Zapero, J., Weiss, M., Descals, A., Camacho, F., Lacaze, R., Baret, F., 2023. GEOV2: Improved smoothed and gap filled time series of LAI, FAPAR and FCover 1 km Copernicus Global Land products. Int. J. Appl. Earth Obs. Geoinf. 123, 103479. https://doi.org/10.1016/J.JAG.2023.103479

Verhoef, W., Jia, L., Xiao, Q., & Su, Z. (2007). Unified optical-thermal four-stream radiative transfer theory for homogeneous vegetation canopies. *IEEE Transactions on geoscience and remote sensing*, *45*(6), 1808-1822.

Wang, Z., Schaaf, C., Lattanzio, A., Carrer, D., Grant, I., Roman, M., Camacho, F., Yang, Y., Sánchez-Zapero, J., 2019. Global Surface Albedo Product Validation Best Practices Protocol. Version 1.0. In Z. Wang, J. Nickeson & M. Román (Eds.), Good Practices for Satellite-Derived Land Product Validation (p. 45): Land Product Validation Subgroup (WGCV/CEOS). [WWW Document]. https://doi.org/doi: 10.5067/DOC/CEOSWGCV/LPV/ALBEDO.001

Weiss, M., Baret, F., Garrigues, S., Lacaze, R., 2007. LAI and fAPAR CYCLOPES global products derived from VEGETATION. Part 2: validation and comparison with MODIS collection 4 products. Remote Sens. Environ. 110, 317–331. https://doi.org/10.1016/j.rse.2007.03.001