CCI+ Vegetation Parameters

Algorithm Development Plan

update 2

Christiaan Van der Tol, Simon Blessing, Else Swinnen

November 2025

Distribution list

Author(s) : Christiaan Van der Tol, Simon Blessing

Reviewer(s) : Else Swinnen

Approver(s) : Marin Tudoroiu

Issuing authority : VITO

Change record

Release	Date	Pages	Description of change	Editor(s)/Reviewer(s)
D1.2_V1.0	2022/09/15	All	Algorithm Development Plan added in the Annex	Simon Blessing, Christiaan Van der Tol/Else Swinnen
D1.2_V1.1	2022/10/05	All	Answer to RIDs on ADP implemented	Simon Blessing, Christiaan Van der Tol/Else Swinnen
D2.3_V1.0	2023/09/15	All	Drafting update 1 of ADP	Simon Blessing, Christiaan Van der Tol/Else Swinnen
D2.3_V1.1	2023/10/17		Update after review	Simon Blessing, Christiaan Van der Tol/Else Swinnen
D2.3_V2.0	2025/10/02	All	Update at the end of cycle 2	Christiaan van der Tol Simon Blessing, Else Swinnen
D2.3_V2.1	2025/11/20	9,11,15,18	Answer to RID's: minor corrections	Christiaan van der Tol

Table of Contents

LIST OF ACRONYMS

ADP	Algorithm Development Plan
ATBD	Algorithm Theoretical Basis Document
BRDF	Bi-directional reflectance function
CCI	Climate Change Initiative
CCI+	The extension of CCI over the period 2017-2024
CI	Clumping Index
CRG	Climate Research Group
ECV	Essential Climate Variable
ESA	European Space Agency
FAPAR	Fraction of the photosynthetically active radiation absorbed by vegetation
GCOS	Global Climate Observing System
LAI	Leaf Area Index
LAIeff	Effective Leaf Area Index
LAItrue	True Leaf Area Index
URD	User Requirements Document
TARTES	Two-streAm Radiative TransfEr in Snow model
TIP	Two-stream Inversion Package
VITO	Flemish Institute for Technological Research
VP	Vegetation Parameters

List of Tables

LIST OF FIGURES

1 Introduction

1.1 Purpose of this document

This Algorithm Development Plan (ADP) describes an analysis of the technical feasibility to meet the user requirements and address issues that were identified during the validation of the previous cycles. By analysing the trade-off between requirements and feasibility, a prioritisation is made of what ECV products should be developed to maximise benefits to the users.

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_V2.0	User Requirement Document: fAPAR and LAI, ESA CCI+ Vegetation Parameters https://climate.esa.int/media/documents/VP-CCI_D1.1_URD_V2.0.pdf
VP-CCI_D2.1_ATBD_V2.2	Algorithm Theoretical Basis Document: fAPAR and LAI, ESA CCI+ Vegetation Parameters http://climate.esa.int/media/documents/VP-CCI_D2.1_ATBD_V2.2.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_V2.1	Product Validation and Intercomparison Report: fAPAR and LAI, ESA CCI+ Vegetation Parameters http://climate.esa.int/media/documents/VP-CCI_D4.1_PVIR_V2.1.pdf

1.3 Summary

In the Algorithm Development Plan, an analysis is made of the technical feasibility to meet the user requirements. By analysing the trade-off between requirements and feasibility, we establish a prioritisation of what ECV products should be developed to maximise benefits to the users. We include a specification of the ECV products that are planned to be developed in the project. The Algorithm Development Plans are updated with the experiences and the user requirements from the previous cycles. This is the second update of these plans for the beginning of cycle 3. It is based on the previous plan, the ongoing evaluations from cycles 1 and 2, on the statement of work, and on the experiences with the implementation of the algorithms, assuming that, among others, temporal coverage, accuracy, and reliable uncertainty estimates are fundamental user requirements. This document is an update of APD V1.2 that describes the development of the OptiSAIL algorithm in the third cycle of the project.

In cycle 1, the algorithm development already covered the multi-sensor cloud contamination detection in OptiSAIL, along with other measures which improved quality and speed. The focus of cycle 2 was on the addition of sensors in the retrieval step. Further improvement to the algorithm was the introduction of a mixed prior approach that uses the previous retrievals. This improves the temporal smoothness of the data. An important aspect has been the testing of the consistency of the product when using different combinations of sensors as input. This appeared successful, with no sensor-induced discontinuities in the statistics of the data, although a trend in the bias with other datasets can be observed that seems to be related to some changes of sensor combinations. This confirms that the chosen approach of physically based inversion and inclusion of uncertainty information of the L1 products of the sensors was suitable.

In both cycle 1 and cycle 2, user requirements have been collected and prioritized. Most of these have been implemented during these cycles. Most of the developments proposed in an earlier version of this ADP have been implemented during cycle 2 of the project, as described in the current version of the ATBD [VP-CCI_D2.1_ATBD]. A few improvements have not been implemented yet, or the implementation has not been fully completed. For example, the inclusion of clumping correction is only partly addressed. In addition, a few new requirements have been received from the user community that needs to be addressed as well (additional variables).

The validation of CRDP-2 showed improvements compared to CRDP-1 in terms of completeness and smoothness. Two new aspects were identified that need to be addressed before processing CRDP-3. These include:

- An apparent discontinuity (increase) in the mean of both LAI and fAPAR between with increasing number of sensors.
- A lower overall mean LAI than in CRDP-1, that is low even considering that the product represents effective LAI.

Table 1 summarises the updated plan for technical developments, their risks, and their benefit to the user community. In the subsequent sections, the items of this table are described in more detail.

Table 1: Development goals with risks and benefits

Development goal/innovation	**technical risk**	**risk description**	**benefit for the community**	**Priority**
Physically-based algorithm	none	Implemented in cycle 1	users prefer physically-based over machine learning (61/100)	NA
Joint multi-sensor retrieval with the same algorithm	none	Implemented in cycle 2	Consistent product with similar interpretation over a long time series and a wide range of sensors, which is more important to users than high spatial resolution (70/100). Allows for low-latency operational line which uses only low-latency sensors.	NA
Use previous retrieval as prior	none	Implemented in cycle 2	Faster processing; less noise at higher temporal resolution; higher temporal resolution <10 days required by users (70/100)	NA
Extend cloud detection to multi-sensor	none	Implemented in cycle 1	better coverage, more stable and better-quality retrievals by avoiding cloud-contamination not detected in the cloud flags	NA
Snow detection jointly with veg. ECV retrieval	None	Included	identification of snow-influenced backgrounds, better quality retrievals	NA
Investigation cause of discontinuities (increases in mean LAI), in particularafter 2018	Medium	Cause can be in the L1 sensor data, which cannot be addressed in the project. Approach: Leaving sensors out, and sensitivity to position of bands and band weighting. In the worst case, leaving out a sensor	Temporal consistency is the most important user requirement. Discontinuities due to sensor or methodology would limit the usefulness for climate studies.	Highest (1)
Investigation of the cause of low LAI	Medium	Problem may be inherent model limitation (SAIL and/or PROSPECT), which cannot be addressed in the project	Better interpretation of the data. Better use of the data in DGVM's. Consistency with other data products.	High (2)
Use of clumping	medium	Model	improved usability, closer to GCOS	High (3)

Development goal/innovation	technical risk	risk description	benefit for the community	Priority
corrected true LAI		compatibility issues with requirements on true LAI from effective LAI, solutions available	definition	
Stricter GCOS definition: green LAI	medium	Post-processing step is necessary, alternative available methods each have limitations. Full physically based separation of senescent from green LAI not feasible.	Improved usability, closer to GCOS definition.	Low (4)
Retrieval of leaf pigments, leaf water content, surface soil moisture	none	Implemented in CCN	potentially useful data, more detailed regard of sources of uncertainty	NA
Use of correlation information in observational covariance matrix.	medium	A typical correlation pattern needs to exist and to be identified, as this information is only available for test cases; some loss of computational speed.	Potentially improved quality of retrievals due to better characterization of input uncertainties.	Low
Performance improvement for when correlated inputs are used (exploiting of block diagonal matrix structure)	none	Implemented in cycle 1	Potentially improved quality of retrievals due to better characterisation of input uncertainties. Better computational performance makes the use of input correlations viable.	NA
Performance improvement through selective use of observations from time window in case of abundance	None	Implemented in cycle 1	Better computational performance allows for more development sub-cycles and better product maturity.	NA
Provision of chlorophyll- and	None	Implemented in cycle 1	More specific information on fAPAR, relevant for estimation for energy	NA

Development goal/innovation	technical risk	risk description	benefit for the community	Priority
carotenoid-specify canopy absorption (fAPAR-green, fAPAR-Car)			budget of photosynthesis.	
Provision of brown pigment (Cs) content	None	Data available, but not foreseen to distribute publicly and unvalidated. Side study in collaboration with CRG in which the data can be made available	Application in global ecology and nutrient recycling	Low

2 Discussion of user requirements

2.1 Consistency of the time series

The main ECVs produced in the project are LAI and fAPAR. Since the start of the Vegetation -CCI project, the question of the added value of a new product of these ECVs has been frequently asked by members of the scientific community, considering the number of available products.

To provide unique data products, the project has decided for a direct physically based inversion of a radiative transfer model. Even though not all users have confidence in data products derived from radiative transfer models due to past experience and obvious limitations, its advantages outweigh the disadvantages. First, it enables a joint sensor approach in which observations of different missions can be combined, considering the uncertainty of each observation. The absence of sensor-specific calibration or training avoids discontinuities in the time series as new sensors are included. Second, the models provide consistency between LAI and fAPAR, because fAPAR is a diagnostic output of the same model for which LAI is retrieved. Third, the approach provides values for the uncertainty of the data products, propagated from the satellite sensor measurements. This is essential information for data assimilation.

The user requirement investigation [VP-CCI_D1.1_URD], confirmed the choice of a physically based retrieval method over machine learning approach, which is applied to multi-sensors jointly. However, different combinations of sensors carry different information content, hence it is not possible to produce a dataset that is truly 'sensor independent'. Higher information content will allow the algorithms to retrieve values which are further away from prior assumptions, and thus the introduction of sensors in the time series with higher information content may also introduce discontinuities in the statistics.

The stability of the retrievals has been evaluated in the validation part of this project. Discontinuities in the mean and a decreasing bias compared to MODIS was discovered with increasing number of sensors. For this reason, further investigation is needed to identify the root cause. Considering the importance of the user requirement, this investigation has priority in the algorithm development.

For some users, true LAI rather than effective LAI is very important, following the GCOS definition, expressed in m^2/m^2 of ground surface. 1D models such as OptiSAIL do not account for heterogeneity in the horizontal direction. Retrievals from satellite data with such models result in an effective LAI, resembling the LAI of a surface with horizontally homogenously distributed leaves that produces the same reflectance as the real vegetation canopy in which the leaves are more concentrated. Due to the shape of the LAI-reflectance relationship, the effective LAI is usually lower than the true LAI, especially in highly structured or sparse vegetation. The effective LAI (LAIeff) from OptiSAIL can be defined as the LAI-parameter of a 1-D radiative transfer 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), where the true LAI (LAItrue) is the one-sided green leaf area per unit of ground cover. Differences between LAIeff and LAItrue arise due to clumping or incomplete vegetation cover within the pixel (Figure 1). The 1-D nature of OptiSAIL can by design not account for the optical effects of vegetation clumping or clustering or vegetation within a pixel (~1km2). This is also shown in the [VP-CCI_D4.1_PVIR and VP-CCI_D2.4_PVASR]. Models exist that represent this complexity explicitly, but these are not invertible.

LAIeff and LAItrue have different applications. LAIeff is useful for the quantification of light interception, plant-light interactions, and photosynthesis. An important feature is that it is consistent with fAPAR. It is also useful for the estimation of rainfall interception. The true LAI is a state variable in dynamic vegetation growth models, and it is relevant for (carbon) stock modelling, respiration estimation, biomass estimation, and vegetation water content (Fang, 2021).

In DVGMs, the fAPAR is calculated from LAI, while the LAI is a model state variable that is updated between time steps. Each DVGM uses a different radiative transfer model for this purpose, with different (or no) representations of clumping, leading to divergence in the way LAI is used. The importance of considering the effect of clumping on gross primary productivity has been shown by (Braghiere et al., 2019).

Figure 1. Examples of clumping at different scales: clumping of needles in shoots, and crown clumping in sparse vegetation (Dehesa)

To meet the user requirement, a post-processing tool to obtain true LAI from the effective LAI is therefore required at minimum. This post-processing script has been developed during cycle 2, as described below. In case this is insufficient for the user, then further development of a true LAI from the effective LAI may be required, for example by including the post-processing step in the algorithm.

Alternatively, we have also identified a published dataset in which a similar radiative transfer inversion (SCOPE) was carried out including the retrieval of fractional vegetation content (FVC) from Sentinel-3 without land cover specific parameterization (Kovács et al., 2023). This is relevant for one particular type of clumping, notably sparse vegetation (Figure 1, left panel) Considering the importance of true LAI for the users, we will investigate how well such retrieval can be constraint by the observations (whether equifinality can be avoided), and compare the two datasets.

It is in theory possible to use biome-specific 3-D models to retrieve LAItrue directly, rather than deriving it in a post-processing step using a CI product. However, this has two major drawbacks. The parameterization would be much more complex and cannot be not consistent with all DVGM's or earth system model (ESMs), because these use different biome specific parameterizations of the radiative transfer. Second, such retrieval is poorly constrained and requires additional assumptions on the structure of biomes and a land cover classification map, and this has been identified as a major drawback by some users of other products. Biome specific parameterizations through land cover classification tend to be conservative with respect to land cover. More sophistication on the model level would require more parameters to be estimated, with all the adverse consequences such as, worse performance, longer temporal aggregation window. It may also potentially lead to convergence problems due to under-determination. Furthermore, the use of a land-cover input layer may have adverse effects on the long-term consistency of the CDR, due to different qualities of historic land cover classifications and in case of land cover changes.

For all these reasons, we decided for a uniform parameterization for all land cover types that does not require ancillary land cover specific inputs or parameter values.

2.2 Additional variables

The requests from the scientific community have been evolving during the past years, along with the increasing availability of multi and hyperspectral data. New requests received during cycle 1 include the separation of fAPAR by pigments that are involved in photosynthesis or photoprotection (fAPARchl and fAPARcar). During cycle 2, a request was received for estimates of the content of brown pigments not involved in photosynthesis. Furthermore, the need for approaches that help the interpretation of solar induced fluorescence (SIF) has emerged. Because OptiSAIL is both modular and highly efficient in the generation of ancillary outputs, these requests can be addressed with sufficient satellite data as input.

For example, during cycle 1 of the project, fAPAR for pigments has been produced as an output. During cycle 2 of the project, in the context of a CCN, the scattering of fluorescence has been implemented, using parts of the code of the model SCOPE. The modularity of OptiSAIL also enables following the developments of the leaf radiative transfer model PROSPECT (such as PROSPECT-PRO).

Another example of such ancillary outputs is the bi-hemispherical and directional-hemispherical reflectance (black-sky albedo) for the VIS, NIR and SW spectral range computed with OptiSAIL from the retrieved parameter set.

By providing the ancillary data, users will be able to post-process the data according to their wishes, for example using LUT that link PROSAIL to 3D RTM for specific vegetation types (e.g., (Miraglio et al., 2020).

It was decided not to publicly distribute these additional layers, because these layers are not validated. However, the data can be made available through Terrascope to interested users. Feedback from the use of these layers will be collected by the CRG.

3 Adaptation of the algorithm and output to user requirements

The requirements on temporal consistency and consistency between LAI and fAPAR have been achieved in Cycle 2, by means of measures described in Sections 3.1 to Section 3.7 below. Other adaptations are presented in Sections 3.8 to 3.10.

3.1 Temporal stability

To resolve the issue of the discontinuity in 2018 that was discovered during the validation, the following steps will be taken. In a diagnosis step, scenarios of excluding single sensors will be carried out to identify if a specific sensor is responsible for the discontinuity: (1) Excluding OLCI on Sentinel-3, and (2) excluding PROBA-V. If neither of these is found responsible, then tests will include excluding VIIRS and AVHRR sensors.

The identification of a sensor may give an indication of the cause. As a second step, the sensitivity of the retrieval of LAI to the position of the bands will be investigated. Due to model deficiencies, the outcome of the retrieval may differ depending on the wavelengths of the observations. Because this varies among sensors, a correction may be required by providing wavelength dependent weights to the information. If this solution is not sufficient, then the sensor responsible for the discontinuity will be excluded from the time series.

If a sensitivity of the retrieval on the position of the bands is demonstrated, then that would suggest a model deficiency. Such model deficiency must be analysed further and addressed outside the scope of the project.

3.2 Low values for the LAI

The relatively low values of the LAI compared to CRDP-1 as found in the validation will be addressed by evaluating the band sensitivity analysis described in Section 3.1. It is possible that the inclusion of additional bands has lowered the LAI compared to CRPD-1, if the retrieval is sensitive to the position of the bands or if there is any other model deficiency.

Other tests have already been conducted on a Sentinel-3 synergy dataset described in (Blessing et al., 2024), including assessing the influence of the mixed prior, the sensor combination, the number of sensors, the BRDF of the soil, the cloud model and the filter for the brightest pixel. These tests indicated that neither of these isolated factors is responsible for the lower overall LAI in CRDP-2 than CRDP-1.

3.3 Cloud detection in multi-sensor retrieval

OptiSAIL cloud contamination detection has been extended from using one single cloud thickness parameter in the Sentinel-3-synergy (SY_2_SYN) processing from (Blessing et al., 2024) to multiple cloud contamination parameters, one per sensing geometry and time, in order to improve coverage in areas with frequent cloud-cover. This was already used for the CRDP-1 processing. This may also allow for a higher temporal resolution due to more observations which can potentially be included in the retrieval by ignoring or relaxing the cloud mask.

3.4 Snow detection

OptiSAIL has sub-canopy snow detection by design by the inclusion of the snow reflectance model TARTES. The detection of snow is important because of its high impact on the radiative transfer in the canopy, thus affecting data quality in high latitudes and in winter. Status maps coming with the TOC reflectance may be inconsistent between sensors and are typically not sensitive to snow under the canopy.

3.5 Previous retrieval as prior

This use of the results of the previous retrieval as prior has been demonstrated in the literature for related retrieval systems, for example (Yang et al., 2021) retrieved stable LAI time series by applying temporal covariance in the Soil-Plant-Atmosphere radiative transfer (SPART) model. It is technically feasible for OptiSAIL. It has the potential to reduce noise, speed up the processing, improve the accuracy of the retrieved quantities, and/or allow for a higher temporal resolution. The operational implementation has been adopted to accommodate the sequential dependency of the retrievals. A consequence is that parallelisation can only be done in the spatial domain, but no longer in the temporal domain. It also creates the need of a spin-up period in the processing and a careful choice of the strength of the constraint by the prior, based on the uncertainty estimate, its correlation with other parameters, and the temporal variability of the constrained parameter: (Yang et al., 2021) showed that the output is sensitive to the choice of parameter values that quantify the strength of the temporal covariance. Furthermore, the use of previous retrievals as prior (depending on the strength of the prior) may reduce the ability of the algorithm to detect abrupt changes such as harvest or fire events. During Cycle 2, this temporal dependence has been implemented as described in the ATBD.

This technique is potentially able to fill gaps in the observational time series. It will have to be decided, also based on interaction with the CRG, whether gap filling with increased uncertainty values is preferred over a missing retrieval. Caused by the strictly sequential processing, such gap-filling will tend to have jumps at the end of the gap when new observations become available, because the new observations do not influence the filling of the preceding gap. For CRDP-2 it was decided to discard these values.

3.6 Further options to stabilise the retrievals

LAI, leaf chlorophyll content, and leaf inclination angle show a certain interdependence which may allow for different values of this parameter set to exhibit a similar spectral signature. We speculate that multi-angular observations with good coverage of NIR and SWIR bands may improve on the constraint on the canopy structure and hence on all parameters. Experience in CRDP-2 shows, that there may still be inconsistencies between bands, possibly both on model and observations side, which lead to an improvement that is smaller than expected. However, a number of options are available to prescribe a certain inter-dependence between the parameters and thus stabilise the retrieval. Examples for this approach are:

- Prefer solutions with higher leaf chlorophyll content by choosing a higher or stronger prior.
- Prefer solutions with green leaves over senescent leaves by using a Chlorophyll/Carotenoid ratio of 4 to 5 as a soft constraint.
- Use empirical functional dependency of leaf specific area and the structure parameter N as a soft constraint (Jacquemoud and Baret 1990)

Note that all these options are soft constraints on the retrieval, which still permit the full range of the retrieved parameters but provide additional prior information to the inversion. Their implementation will require an assessment of the strength of the constraint which is required to get a good balance between regularisation and liberty of the inversion. For CRDP-1, none of these options was used. For CRDP-2, the prior for chlorophyll was actually reduced, because the old default was considered to be quite high. The performance was good enough to abstain from the other measures suggested in this section.

3.7 Selection of observations from the available observed bands in a time window

Selecting a subset of observations from all observations available in the time window has two advantages. The first is a higher processing speed, because simulations only have to be done for the selected scenes. The second is a better temporal resolution of changes taking place at a timescale shorter than the window size, where the amount of good (e.g., flagged cloud-free) observations does permit this. For this, a pre-selection algorithm was implemented, which chooses for each observed band the N observations closest in time to the valid date (the window centre). For a given sensor combination this has no effect in the case of sparse data but caps the maximum number of observations which is used in the retrieval when multiple (more than N) observations of the same band are available. Further quality-oriented pre-selection algorithms are in place (see [VP-CCI_D2.1_ATBD]).

3.8 True LAI

A Jupyter Notebook is provided that provides instructions for the access to an existing clumping index product, and provides projection, co-location in time, and the computation of true LAI. At present, the tool is provided but the steps have to be made by the user of the data.

The tool makes use of a published, MODIS Aqua and Terra based algorithm for a 'clumping index' (CI) developed by (Wei et al., 2019; Wei & Fang, 2016) . The reason for choosing this product is that it makes use of the bi-directional reflectance distribution function of MODIS Aqua and Terra data, thus it is strongly driven by observations and it is a time series. However, the product still makes use of land cover specific coefficients (A and B in Eq 2). The corrected LAI is computed in the notebook as:

$$LAI_{true} = \frac{LAI_{eff}}{CI} \tag{1}$$

Where CI is derived from the MCD43A4 V006 BRDF product (Schaaf et al., 2002; Wang et al., 2018) , as the normalized difference between hotspot and darkspot (NDHD) (Chen et al., 2005; Leblanc et al., 2005) :

$$CI = A \cdot NDHD + B \tag{2}$$

Where

$$NDHD = \frac{\rho_h - \rho_d}{\rho_h + \rho_d} \tag{3}$$

The hotspot (ρ_h) and darkspot (ρ_d) are computed at daily time intervals from the BRDF kernels of the MCD43A4 product (Wei et al., 2019), which are fitted using 16-day moving retrieval window of MODIS Aqua and Terra data (Schaaf et al., 2002; Wang et al., 2018) .

In the current version of the Jupyter notebook, the user is instructed to use Google Earth Engine to compute CI with the above approach, using the tool developed by (Li & Fang, 2022) . However, the team is working on a more use friendly approach for the post-processing, or potentially the implementation of the computation of LAItrue in the processing chain.

3.9 Green LAI

The GCOS definition of LAI includes the word 'green' before LAI, to differentiate from senescent vegetation and from plant area index (PAI), the one-sided area of foliage and wood. OptiSAIL does not differentiate wood from foliage, nor green from senescent vegetation. A straightforward approach to limit the LAI to the green fraction, would be to discard the LAI if Cab drops below the threshold of 10 µg cm^{-2} of chlorophyll. However, because OptiSAIL retrieves a single value of Cab for all leaves (does not discriminate between brown and green leaves), this threshold will be too abrupt and not accommodate the gradually increasing fraction of senescent leaves. It is probably better to let the user define thresholds based on fAPARchl, fAPARcar. Alternatively, it is possible to differentiate leaves from each other, as in SenSCOPE or mSCOPE (Pacheco-Labrador et al., 2020; Yang et al., 2017) . However, this is not desired as it introduces additional degrees of freedom in OptiSAIL and may lead ill-posed problems.

3.10 Other variables

With respect to additional variables, the computation of fAPAR based on the chlorophyll absorption and the carotenoid absorption (fAPARchl, fAPARcar), and the chlorophyll content Cab has been implemented.

Recent studies highlight the importance of the seasonality of chlorophyll and the green portion of LAI for estimating gross primary productivity (Reitz et al., 2023) , thus we expect the user demands for fAPAR-green to grow in the near future. The CRG assessed the added value of fAPAR-Chl for phenology, climatology and shifts in the timing of the growing season.

Because the request for the retrieval of brown pigments (Cs) has been received for ecosystem modelling, it is the question whether more output layers should be made available. Validation of these output layers are not feasible in the project, and challenging in general considering that no fiducial reference measurements are available for fAPARchl, fAPARcar and Cs. They nevertheless enable possibilities for expert users.

To allow the users to reconstruct the background spectrum, the spectral basis functions for the soil model are already included in the metadata of the OptiSAIL retrievals. If the retrieved soil reflectance coefficients are also provided, then the users could reconstruct the soil reflectance spectra easily.

Furthermore, the data also include the correlations among individual data layers. In multi-variable assimilation of data in DGVMs of, for example, LAI, fAPAR and Cab, accounting for the correlation between these data will influence the model state updates. The value of these correlations have hitherto not been evaluated, but this could still be investigated by the CRG or other expert users.

OptiSAIL retrieves not only LAI and fAPAR, but also other vegetation parameters, such as leaf pigments and leaf water content. Even if these quantities may not be well determined in all situations, taking them into account gives a more realistic estimate of the overall retrieval uncertainty. While the validation of these quantities is beyond the validation activities foreseen in this project, interest has been voiced from the CRG, and is evident in the scientific literature on vegetation trait analysis (e.g. Kattenborn et al., 2017) to study these data and to confront them with in situ observations in a suitably scoped project, which, if done in a timely fashion, could feed back into the algorithm development of the present project. In the CRDP-1 we have included leaf chlorophyll content and fAPAR by chlorophyll. Depending on user feedback and after the inclusion of sensors with multiple bands in CRDP-2, this could be extended with carotenoid content and fAPAR-Car.

4 References

Blessing, S., Giering, R., & van der Tol, C. (2024). OptiSAIL: A system for the simultaneous retrieval of soil, leaf, and canopy parameters and its application to Sentinel-3 Synergy (OLCI+SLSTR) top-of-canopy reflectances. *Science of Remote Sensing*, *10*, 100148. https://doi.org/10.1016/J.SRS.2024.100148

Braghiere, R. K., Quaife, T., Black, E., He, L., & Chen, J. M. (2019). Underestimation of Global Photosynthesis in Earth System Models Due to Representation of Vegetation Structure. *Global Biogeochemical Cycles*, *33*(11), 1358–1369. https://doi.org/10.1029/2018GB006135

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*(4), 447–457. https://doi.org/10.1016/J.RSE.2005.05.003

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

Kattenborn, T., Fassnacht, F. E., Pierce, S., Lopatin, J., Grime, J. P., & Schmidtlein, S. (2017). Linking plant strategies and plant traits derived by radiative transfer modelling. *Journal of Vegetation Science*, *28*(4), 717–727.

Kovács, D. D., Reyes-Muñoz, P., Salinero-Delgado, M., Mészáros, V. I., Berger, K., & Verrelst, J. (2023). Cloud-Free Global Maps of Essential Vegetation Traits Processed from the TOA Sentinel-3 Catalogue in Google Earth Engine. *Remote Sensing*, *15*(13), 3404.

Leblanc, S. G., Chen, J. M., White, H. P., Latifovic, R., Lacaze, R., & Roujean, J. L. (2005). Canada-wide foliage clumping index mapping from multiangular POLDER measurements. *Canadian Journal of Remote Sensing*, *31*(5), 364–376. https://doi.org/10.5589/M05-020

Li, Y., & Fang, H. (2022). Real-Time Software for the Efficient Generation of the Clumping Index and Its Application Based on the Google Earth Engine. *Remote Sensing 2022, Vol. 14, Page 3837*, *14*(15), 3837. https://doi.org/10.3390/RS14153837

Miraglio, T., Adeline, K., Huesca, M., Ustin, S., & Briottet, X. (2020). Joint use of PROSAIL and DART for fast LUT building: Application to gap fraction and leaf biochemistry estimations over sparse oak stands. *Remote Sensing*, *12*(18), 2925.

Pacheco-Labrador, J., El-Madany, T. S., van der Tol, C., Martin, M. P., Gonzalez-Cascon, R., Perez-Priego, O., Guan, J., Moreno, G., Carrara, A., Reichstein, M., & Migliavacca, M. (2020). senSCOPE: Modeling radiative transfer and biochemical processes in mixed canopies combining green and senescent leaves with SCOPE. In *bioRxiv*. https://doi.org/10.1101/2020.02.05.935064

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).

Reitz, O., Bogena, H., Neuwirth, B., Sanchez-Azofeifa, A., Graf, A., Bates, J., & Leuchner, M. (2023). Environmental Drivers of Gross Primary Productivity and Light Use Efficiency of a Temperate Spruce Forest. *Journal of Geophysical Research: Biogeosciences*, *128*(2). https://doi.org/10.1029/2022JG007197

Schaaf, C. B., Gao, F., Strahler, A. H., Lucht, W., Li, X., Tsang, T., Strugnell, N. C., Zhang, X., Jin, Y., Muller, J. P., Lewis, P., Barnsley, M., Hobson, P., Disney, M., Roberts, G., Dunderdale, M., Doll, C., D'Entremont, R. P., Hu, B., … Roy, D. (2002). First operational BRDF, albedo nadir reflectance products from MODIS. *Remote Sensing of Environment*, *83*(1–2), 135–148. https://doi.org/10.1016/S0034-4257(02)00091-3

Wang, Z., Schaaf, C. B., Sun, Q., Shuai, Y., & Román, M. O. (2018). Capturing rapid land surface dynamics with Collection V006 MODIS BRDF/NBAR/Albedo (MCD43) products. *Remote Sensing of Environment*, *207*, 50–64. https://doi.org/10.1016/J.RSE.2018.02.001

Wei, S., & Fang, H. (2016). Estimation of canopy clumping index from MISR and MODIS sensors using the normalized difference hotspot and darkspot (NDHD) method: The influence of BRDF models and solar zenith angle. *Remote Sensing of Environment*, *187*, 476–491. https://doi.org/10.1016/J.RSE.2016.10.039

Wei, S., Fang, H., Schaaf, C. B., He, L., & Chen, J. M. (2019). Global 500 m clumping index product derived from MODIS BRDF data (2001–2017). *Remote Sensing of Environment*, *232*, 111296. https://doi.org/10.1016/J.RSE.2019.111296

Yang, P., Verhoef, W., Prikaziuk, E., & van der Tol, C. (2021). Improved retrieval of land surface biophysical variables from time series of Sentinel-3 OLCI TOA spectral observations by considering the temporal autocorrelation of surface and atmospheric properties. *Remote Sensing of Environment*, *256*, 112328. https://doi.org/10.1016/J.RSE.2021.112328

Yang, P., Verhoef, W., & van der Tol, C. (2017). The mSCOPE model: A simple adaptation to the SCOPE model to describe reflectance, fluorescence and photosynthesis of vertically heterogeneous canopies. *Remote Sensing of Environment*, *201*, 1–11.