CCI+ Vegetation Parameters

Algorithm Development Plan

update 1

Christiaan Van der Tol, Simon Blessing, Else Swinnen

September 2023

Distribution list

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

Reviewer(s) : Else Swinnen

Approver(s) : Clement Albergel

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

Table of Contents

List of Acronyms 4
List of Tables 5
1 Introduction 6
1.1 Purpose of this document 6
1.2 Related documents 6
2 Algorithm Development Plan 7
2.1 Introduction 10
2.2 Discussion of user requirements 10
2.3 Adaption of outputs to user requirements 11
2.4 Cloud detection in multi-sensor retrieval 12
2.5 Snow detection 12
2.6 Previous retrieval as prior 12
2.7 Further options to stabilise the retrievals 13
2.8 Selection of observations out the available observed bands in a time window 13
2.9 Other opportunities 13
3 References 15

LIST OF ACRONYMS

ADP	Algorithm Development Plan
ATBD	Algorithm Theoretical Basis Document
CCI	Climate Change Initiative
CCI+	The extension of CCI over the period 2017-2024
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
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

Table 1: Development goals with risks and benefits 8

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 cycle(s). 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_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

2 Algorithm Development Plan

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 first update of these plans for the beginning of cycle 2. It is based on the previous plan, the ongoing evaluations from cycle 1, 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.

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 is on the addition of sensors in the retrieval step. The addition of further sensors in cycle 2 will likely require a characterisation of the independence of the individual observations, namely covariance information. However, innovations on the part of the atmospheric correction, such as the provision of full per pixel TOC reflectance covariance data, must remain subject of minor test experiments (subject to capacity) since they would create an inconsistency with the brokered datasets. If a typical covariance structure of the TOC reflectances can be identified, the efficient exploitation of it is implemented in OptiSAIL.

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
Physically-based algorithm	none	implemented	users prefer physically-based over machine learning (61/100)
Joint multi-sensor retrieval with the same algorithm	medium	while technically implemented, the different combinations of sensors will carry different information content, potentially leading to discontinuities, possibly in retrieved values, and very probably in uncertainty levels. Without further measures, lowest resolution product dictates final resolution	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.
Use previous retrieval as prior	low	temporal correlation in data, parallelisation only over spatial dimension	Faster processing; less noise at higher temporal resolution; higher temporal resolution <10 days required by users (70/100)
Extend cloud detection to multi-sensor	implemented	Implemented in cycle 1	better coverage, more stable and better-quality retrievals by avoiding cloud-contamination not detected in the cloud flags
Snow detection jointly with veg. ECV retrieval	none	included	identification of snow-influenced backgrounds, better quality retrievals
Adaption of outputs to user requirements	low	balance of preferences, done in cycle 1; to be continued upon user feedback on CRDP-1	improved usability
Retrieval of leaf pigments, leaf water content, surface soil moisture	none	implemented, but no validation foreseen in main contract	potentially useful data, more detailed regard of sources of uncertainty

Development goal/innovation	technical risk	risk description	benefit for the community
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;	Potentially improved quality of retrievals due to better characterisation of input uncertainties.
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.
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.
Provision of chlorophyll- and carotenoid-specify canopy absorption (fAPAR-green, fAPAR-Car)	none	Implemented in cycle 1	More specific information on fAPAR, relevant for estimation for energy budget of photosynthesis.

2.1 Introduction

The main ECVs produced in the project are LAI and fAPAR. However, depending on the application, there are different preferences from the users' side, regarding the exact definition and way of computation of these quantities. Moreover, as stated in the technical proposal, OptiSAIL is both modular and highly efficient in the generation of ancillary outputs, which will be exploited in the project. For example, fAPAR for pigments is also an output and can be exploited, the scattering of fluorescence can be implemented with little additional computational effort thus adopting developments in the model SCOPE. Updates of the leaf radiative transfer model PROSPECT (such as PROSPECT-PRO) can be implemented with little effort. Based on the user requirements of Task 1, the configuration of OptiSAIL will be updated in cycle 2, provided that test inversions turn out encouraging. For a description of the present status of the algorithms and their correspondence to the project goals, please refer to the ATBD [VP-CCI_D2.1_ATBD]. 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, and the bi-hemispherical reflectances (white-sky albedo) computed as intermediate output with OptiAlbedo.

2.2 Discussion of user requirements

A first evaluation of the D1.1 - User Requirements Document [VP-CCI_D1.1_URD], confirms the choice of a physically based retrieval method over machine learning approach, which is applied to multi-sensors jointly. This allows for the retrieval of time series of maximum length with good consistency as preferred by users. It must be noted, however, that different combinations of sensors carry different information content, hence it is not possible to produce a full sensor independent dataset. Higher information content will allow the algorithms to retrieve values which are further away from prior assumptions, and thus introduce discontinuities in the statistics. The stability of the retrievals is evaluated in the validation part of this project.

Most users stress that a 'true LAI' is needed, following the definition of m^2/m^2 of ground surface. Many radiative transfer models assume that the leaves are homogeneously distributed (turbid medium models or 1D models, such as OptiSAIL). If in reality the leaf area is concentrated on parts of the surface within the 1-km footprint (for example in crowns with gaps in between), then a LAI retrieved from such 1-D model tends to underestimate the true LAI, also shown in the [VP-CCI_D4.1_PVIR and VP-CCI_D2.4_PVASR].

At present, neither proposed algorithmic chain retrieves clumping or a land cover fraction. Further investigation is needed to whether (and how) an approximate land-cover specific or land surface model specific conversion can be derived from land-cover maps by the users. Given the coarse nature of the observations and with disregard of the land-cover maps, also no differentiation is made in the algorithms for pixels with mixed vegetation (e.g., Savannah). The retrieved values are for homogenous vegetation at the pixel level and consistent with the model assumptions of two-stream and PROSAIL, respectively.

This issue, combined with the user requirement of achieving physical consistency with land surface models (LSM), is a gordian knot. Most LSMs include a simplified radiative transfer scheme to calculate leaf photosynthesis, using a constant conversion from true to apparent LAI, but these approaches differ among LSMs. This implies that it is not possible to achieve physical consistency with all LSMs. The use of biome-specific radiative transfer parameterization with ancillary (land cover) used by some existing products, which rely on additional assumptions on the structure of these biomes and a land cover classification map, has been identified as a major drawback by some users. Hence, biome specific parameterizations through land cover classification tend to be conservative with respect to land cover. For this reason, we decided for a uniform parameterization for all land cover types that does not require ancillary land cover specific inputs or parameter values.

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.
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). Furthermore, we have identified a recently 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 (Kovacs et al., 2023). 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.

2.3 Adaption of outputs to user requirements

In order to allow the users to reconstruct the background spectrum, the spectral basis functions for the soil model are included in the metadata of the OptiSAIL retrievals. The data also include the correlations among individual data layers. The CRG can investigate the value of these correlations for users. 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 computation of fAPAR based on the chlorophyll absorption (fAPAR-green or fAPAR_Cab) and the Carotenoids is implemented (e.g., Zhang et al., 2005; Yang et al., 2021; Croft et al., 2020). 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 will assess the added value of fAPAR-Chl for phenology climatology and shifts in the timing of the growing season.

After the publication of CRDP-1, we will actively seek user feedback, analyse the feedback and use it to make informed decisions on the algorithm development along the lines discussed in the next sections.
To inform us about choices in the temporal window, use of priors, or recommendations for temporal filtering as a post-processing step:

- The temporal consistency for applications of phenology (change detection) and consistency with forward simulated LAI and fAPAR in DVGM's. We do not expect users to use the data in data assimilation experiments, as they may prefer to wait for CRDP-

2 (released foreseen in coming months), but we do expect first analyses by comparing our products to forward simulations of vegetation parameters.
- The consistency with fire products and biomass, and the suitability of the data for change detection due to for example deforestation or land use change.

To inform us about informing the users about post-processing of clumping:
- A comparison with a machine learning based inversion of the SCOPE model that uses FCV as an additional parameter, without adding land cover specific priors (Kovacs et al., 2023).

To inform us about the evolution of the product portfolio:
- The added value of the additional output layers we produced, including the albedo, fAPAR-Chl and chlorophyll content.

2.4 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. (2021) 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.

2.5 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.

2.6 Previous retrieval as prior

This technique 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. However, it requires changes in the operational implementation, since previously independent tasks get a sequential dependency. That means 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. We will actively seek feedback from users and work with the CRG on the consistency of the CRDP-1 with such events.

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.

2.7 Further options to stabilise the retrievals

LAI, leaf chlorophyll content, and leaf inclination angle show a certain interdependence which may allow for a different 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. However, if this should not suffice, 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

2.8 Selection of observations out 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.

2.9 Other opportunities

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 this could be extended with carotenoid content and fAPAR-Car.

3 References

Blessing, S.; Giering, R. Simultaneous Retrieval of Soil, Leaf, and Canopy Parameters from Sentinel-3 OLCI and SLSTR Multi-spectral Top-of-Canopy Reflectances. Preprints 2021, 2021090147. https://doi.org/10.20944/preprints202109.0147.v1.

Croft, H., Arabian, J., Chen, J. M., Shang, J., & Liu, J. (2020). Mapping within-field leaf chlorophyll content in agricultural crops for nitrogen management using Landsat-8 imagery. Precision Agriculture, 21(4), 856-880.

Feng, M., J.O. Sexton, C. Huang, J.G. Masek, E.F. Vermote, F. Gao, R. Narasimhan, S. Channan, R.E. Wolfe, J.R. Townshend (2013) Global surface reflectance products from Landsat: Assessment using coincident MODIS observations, Remote Sens. Environ., 134, 276-293, 10.1016/j.rse.2013.02.031

Jacquemoud, S.,F. Baret (1990). PROSPECT: A model of leaf optical properties spectra. Remote Sensing of Environment, 34(2), 75-91, 10.1016/0034-4257(90)90100-Z.

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.

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.

Reitz, O., Bogena, H., Neuwirth, B., Sanchez-Azofeifa, A., Graf, A., Bates, J., & Leuchner, M. Environmental Drivers of Gross Primary Productivity and Light Use Efficiency of a Temperate Spruce Forest. Journal of Geophysical Research: Biogeosciences, e2022JG007197.

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.

Zhang, Q., Xiao, X., Braswell, B., Linder, E., Baret, F. and Moore III, B., 2005. Estimating light absorption by chlorophyll, leaf and canopy in a deciduous broadleaf forest using MODIS data and a radiative transfer model. Remote Sensing of Environment, 99(3), pp.357-371.)