CCI Vegetation

User Requirements Document
Cycle 1

Christiaan van der Tol

10 / 2022

Distribution list

Author(s) : Christiaan van der Tol

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	15 September 2022	all	First version	Christiaan van der Tol / Else Swinnen
V1.1	5 October 2022	all	Revised following RIDs	Christiaan van der Tol / Else Swinnen

Executive summary

This document describes the approach and results of the user requirement analysis for the Vegetation Parameters project of the ESA Climate Change Initiative (CCI). The project focusses on obtaining climate data records (CDR) of leaf area index (LAI) and fraction of absorbed photosynthetic radiation (FAPAR). The aim is to develop these data products to support research into the dynamic role of vegetation in the Earth's climate.

The user requirement study aims at maintaining a sustained dialogue with the end-user community. Through a literature review, participation in meetings, a survey and 11 detailed interviews in Months 1-6, users and applications have been identified, GCOS-200 requirements further specified, bottlenecks in existing products identified, and priorities and, recommendations for algorithm development and validation formulated.

The applications include climate reanalysis, phenology, the study of extreme events, land surface model development and intercomparison, local field studies, and early warning services. The user feedback revealed the strength and limitations of current data products in terms of consistency, quality, temporal and spatial resolution. An evaluation of the requirements resulted in a list of priorities for the project:

1. Provide transparency of the processing chain through documentation and ancillary data
2. Ensure physical consistency between LAI and FAPAR
3. Account for snow and soil effects
4. Compare with existing data products of LAI and FAPAR.
5. Derive sensor-independent products.
6. Consider clumping of the vegetation
7.Achieve a temporal resolution better than 10 days.
8. Retrieve of pigment content and green FAPAR
9. Assess consistency with other data products, such as land cover, burnt area.
10. Provide forward simulated SIF

In the following cycles, feedback will be obtained on the choices made during the project and on the quality and added value of the data products. Wider uptake and exploitation of vegetation ECV products will be stimulated.

Table of Contents

LIST OF ACRONYMS

CCI	Climate Change Initiative
CRD	climate data record
ECV	essential climate variables
GCOS	Global Climate Observing System
LAI	leaf area index
URD	User Requirement Document
PVASR	Product Validation and Algorithm Selection Report.
DGVM	Dynamic Global Vegetation model

LIST OF FIGURES

LIST OF TABLES

1 Introduction

1.1 Background and scope of this document

This document describes the approach and results of the user requirement analysis for the Vegetation Parameters project of the ESA Climate Change Initiative (CCI). The project focusses on obtaining climate data records (CDR) of leaf area index (LAI) and fraction of absorbed photosynthetic radiation (FAPAR). The aim is to develop these data products to support research into the dynamic role of vegetation in the Earth's climate.

Several data products of FAPAR and LAI are available already, most of them derived through radiative transfer model inversions, with either land cover specific or a generalized model conceptualization and parameterization. These products diverge, not only in magnitude but also in their performance in representing variability in space and time. Understanding this variability is critical for most applications in climate science. The user requirement study serves to identify the bottlenecks and challenges, leading to a priorities of innovations that can be achieved throughout the project.

The Global Climate Observing System (GCOS) programme has formulated high-level requirements for these and other essential climate variables (ECV) products in its Implementation Plan 2016, which will be updated in 2022. In this update, the differentiation in applications 'for modelling' and 'for adaptation' will be abandoned, and the requirements for modelling application will be guiding.In the user requirement study, these general requirements have been discussed with users and further specified. This report presents the outcome of the requirement study, which serves as input to the algorithm development plan (ADP) and the Product Validation Plan (PVP).

The overall objective of the user requirement study is *to enhance the impact and relevance of the project through a sustained dialogue with the end-user community*. More specifically, we aim to:

1. Identify the users of products of LAI and FAPAR and the applications they use the data for
2. Identify specific requirements for the products in the context of these applications
3. Identify the key bottlenecks in existing data products
4. Analyse the feasibility, technological gaps and identify priorities for innovation
5. Obtain feedback on the choices made during the project
6. Obtain feedback on the quality/ added value of the data products
7. Stimulate wider uptake and exploitation of vegetation ECV products, and build confidence in the products among the user community

The user requirement study is carried out throughout the project. Midway each of the tree cycles of the project, in 6 months into each year (M6, 18, 30), a version of the URD will be released. The current report (version 1.0) presents the results of the first iteration, in which objectives 1-4 are addressed. The consolidated user requirements are input to the ADP and an assessment of feasibility and innovation risks, eventually to the Product Validation and Algorithm Selection Report (PVASR). In the second and third iteration, objectives 2-4 are addressed in further detail (resulting in an update of the report), and objectives 5-7 are addressed as well.

In this document we present the methodology (Chapter 2) and results (Chapter 3), which includes user requirements, recommendations and directions for innovation.

1.2 Related documents

Internal documents

Reference ID	Document
ADP	Algorithm Development Plan
PVP	Product Validation Plan
PVASR	Product Validation and Algorithm Selection Report

External documents

Reference ID	Document
GCOS-200	GCOS 2016 implementation plan

1.3 General definitions

The **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].

The 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].

1.4 The background of GCOS requirements

The 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 background of GCOS requirements.

FAPAR as property of the vegetation quantifies the process of light absorption, which is the first step in photosynthesis, the primary driver and energy source of all terrestrial metabolism. Satellite based estimates of FAPAR can provide insight into the carbon sink on land, and into land-atmosphere exchanges due to its strong coupling to the energy and water budgets of the Earth surface.

LAI is important for climate research because it is a common state variable of Dynamic Global Vegetationmodel (DGVM) that simulate foliage growth with a model for the allocation of carbon assimilated through photosynthesis over root, shoot, stems and reproductive organs. Because the leaves and needles are responsible for light absorption through the pigments they contain, LAI is correlated to FAPAR, although such correlation may be weak or absent in cases where variations in leaf pigment content dominate (e.g., during senescence).

Because LAI is related to the structure of vegetation, it varies on a longer time scale than photosynthesis. The typical time scale of significant variations of LAI in time is in the order of days. For FAPAR, sub-daily fluctuations are possible due to changes in leaf orientation and chloroplast movement, but these are relatively minor and such processes are usually not considered in land surface models. Hence, the highest meaningful temporal resolution for present land surface models is 1 day. Table 1 shows the GCOS-200 requirements for FAPAR and LAI, and Figure 1 shows these requirements for several land surface ECV products graphically, along with the spatial and temporal dimensions at which the relevant land surface processes play a role.

Long time series of FAPAR and LAI contribute to improved understanding of the biosphere. Data products of FAPAR and LAI have been used in land surface and carbon models, dynamic global vegetation models (DVGMs) in various applications.

Table 1 GCOS-200 data requirements for FAPAR and LAI for climate modelling, with respect to space and time for the five applications mentioned in Section 3.1.

	Frequency	Resolution	Uncertainty	Stability per decade
FAPAR	1d	200/500 m	Max (10%; 0.05)	Max (3%; 0.02)
LAI	1d	250 m	Max (10%; 0.05)	Max (3%; 0.02)

The implementation plan lists actions to operationalize the retrieval of FAPAR and LAI products gridded and at global resolution, specifically:

- 10-day and monthly products at 5 km spatial resolution over time periods as long as possible; •
- 10-day FAPAR and LAI products at 50 m spatial resolution; •
- Daily products

The daily products are intended for a for characterization of rapidly greening and senescing vegetation. This is particularly relevant in areas with strong seasonality and snowfall and snow melt, i.e. the higher latitudes, at which polar orbiting overpasses are relatively frequent.

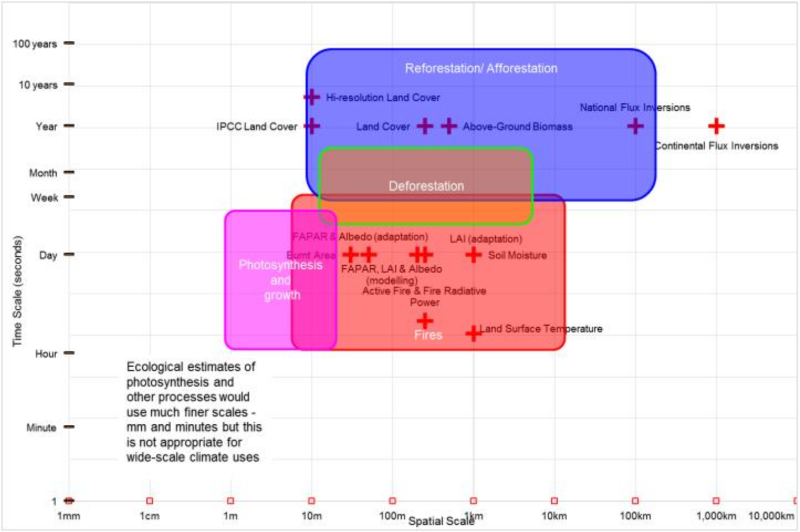

Figure 1 Time and spatial scales of applications in relation to land surface ECV products (GCOS-200)

The specific requirements may vary per application. The use of LAI and FAPAR can be roughly divided into the following application areas. Users in each of these categories have been involved in the user requirement study:

- Climate reanalysis, with an emphasis on data assimilation

- Development of Land Surface models, with emphasis on the handshaking between retrieved products and model state variables
- Specific aspects of the vegetation response to weather/climate extremes such as precipitation and temperature anomalies, often working with flux tower data (ICOS or similar) besides satellite data
- Development of monitoring and early warning services, with emphasis on near-real time availability, high spatial and temporal resolution.
- Monitoring of phenology (start, peak and end of season)

2 Methodology

2.1 User engagement

The overall approach is to seize multiple opportunities for dialogue with the climate change modelling community throughout the project. The user requirement study is carried out in three iterations, corresponding to the three years of the project.

The first year the focus is on the user requirements gathering (WP1.1), followed by an analysis of products by the users and gathering of user feedback (WP1.2). The user requirements are used in the ADP and the PVP (WP1.3).

The user engagement activities in the first 6 months of the project consist of:
- Obtaining an overview of user requirements, key issues, bottlenecks, and opportunities from the scientific literature
- Sending out a survey to users of existing FAPAR and LAI products and potential users of the CCI products, by e-mail and social media.
- Inviting lead scientists for an interview to complement the survey
- Participation in a GCOS requirements meeting, CMUG meetings and online sub-group discussions.

User engagement activities in the months 6-30 include:
- Presentation of results and user engagement at relevant conferences and workshops
- Critical user review of the datasets by the CRG and beta users
- Potentially inviting key users to contribute to a review paper

<u>Iteration 1 (Months 1-6).</u>

In the first iteration, a survey was distributed among (lead) scientists in the field of climate and land surface modelling, followed by a 1-1 interview. The survey addressed applications and data consistency issues reported in the literature. The interviews followed the structure and included the topics of the survey but allowed for a more open discussion on requirements for specific applications, preferences, and user experiences with data products. In line with the first four objectives of the user requirement study, the survey questions addressed the questions:

a) do we include all potential users: which scientific applications should the products target?
b) which are the strengths and weaknesses of existing datasets in view of these applications?
c) which are specific requirements for these applications?
d) which are priorities for development?

The survey was distributed among the network of the CRG and the team, the GCOS network, it was posted on the website and social media. In total, 25 scientists completed the survey (September 2022), and the survey will remain open, and 11 scientists have been interviewed. Although the sample is small, they represented the application domains mentioned in Section 1.4. The interviewed scientists have affiliations in Europe, Asia, and North America. The survey respondents had affiliations at universities (35%), research institutes (30%), companies (20%), government (10%) and space agencies (5%). Out of these, 70% is a regular user of data products of FAPAR and LAI, with the remaining 30% an occasional user (less than twice a year).

The survey results do not provide statistically significant numerical outputs due to limited sampling size, representativeness of the respondents for the total user community, and differences in group size of the research groups that were interviewed. Nevertheless, requirements and priorities that converge among members of the user community emerged.

3 Results

3.1 Applications

A further specification of the applications and focal areas has been identified. Each of the applications poses different requirements on data quality, spatial and temporal resolution, time span, data latency and data access.

1. **Analysis and forecast of long-term (changes in) the carbon land sink and energy budget of the Earth.** In these analyses the land surface is represented by dynamic global vegetation models (DGVM), such as LPJ, ORCHIDEE, JULES, SiB, CLM, which simulate processes in plant communities as a function of their physical environment: the meteorology, land cover and hydrology (Albergel et al., 2020; Fang et al., 2013.; Kaminski et al., 2012; Wu et al., 2018). LAI and/or FAPAR are state variables of these model, or state variable in these models are closely related to LAI and FAPAR.
2. Analysis of feedback mechanisms involved in climatic and/or weather extremes, such as droughts. In these analyses, DGVM’s are commonly used as well, but the geographical area and time period of interest is different from application1 (Cammalleri et al., 2019; Nunes et al., 2012).
3. Monitoring and early warning systems, such as Monitoring Agricultural ResourceS (MARS), focus on monitoring services for agriculture, yield prediction, fire risk with low data latency (Baruth et al.,2008).
4. Analysis of phenology and anomalies therein, such as changes in start of season (SOS), peak of season (POS), and of season (EOS) or length of season (LOS), either carried out with sec vegetation indicators (indices or derived products) or supported with DGVM’s (Bórnez et al., 2020; Macbean et al., 2015).
5. Land surface model intercomparison, where LAI and FAPAR products serve as a benchmark (Lafont et al., 2012).
6. Dedicated scientific studies in local study areas and at flux towers (e.g., ICOS, FLUXNET). These include research projects funded by (national) science organizations, and focus on improving process understanding, developing new measurement techniques, or calibration and validation of satellite data products and use of field experiments. Products of LAI and FAPAR are among the remote sending derived vegetation data products that are used in conjunction with field data (Balzarolo et al., 2015; Maes et al., 2020).

In the use of the products in DGVM’s (applications 1, 2, and 4, 5), a distinction can be made between using the data for validation or as a sanity check of the model on the one hand, and data assimilation on the other hand. In the first case, the data are used independent of the model run. The comparison between model and measurement serves to identify discrepancies in the model simulation of spatial patterns, latitude, humidity and altitude gradients, anomalies, and (seasonal) periodicities, and identify possible shortcomings of models in representing observed vegetation responses. In the second case, the data of LAI and/or FAPAR used in a Bayesian framework to narrow the posterior ensemble of the model output (Bonan et al., 2020).

In the following cycles of the project, we will make the products available to the community. Most users have expressed willingness to serve as beta-users. Preliminary datasets may only be used for validation or model intercomparison, while application in data assimilation requires a mature product due to the investment in computational resources and manpower in such exercise.

3.2 Existing products

The users identified about a dozen of alternative products of LAI and FAPAR that they have used regularly or occasionally (Table *2*). The most widely used products are MODIS15A, GEOV1 and GEOV2. In addition to global datasets, GEOLAND, GEOLAND2, THEIA and LSA-SAF products are also used (e.g. by EUMETSAT), and considered high-quality products, but these do not have global coverage.

The most widely used LAI and FAPAR data sets are the MODIS based MCD15A2 products, which is based on lookup tables of radiative transfer models (Table *2*). Biome specific (8 biomes) parametrizations are used in the design of the LUT, while an NDVI based empirical data product of LAI and FAPAR is provided as alternative.

JRC-TIP is based on a two-stream, turbid medium radiative transfer model concept, albeit in two steps (estimation of spherical albedo, followed by retrieval of LAI and FAPAR). Because of the approach, a strong correlation exists between LAI and FAPAR, and the product uncertainties are not biome specific (Mota et al., 2021).

Similarly, CYCLOPES is a radiative transfer model inversion of atmospherically corrected reflectance, using a trained neural network for inversion. Furthermore, CYCLOPES includes a simple correction of vegetation cover in the pixel, i.e., a LAI per unit of vegetation-covered area. The GCLS products (GEOv1,2,3) are blended products, a weighted MCD15A2 and CYCLOPES are used to train a neural network for retrievals from SPOT-VGT and PROBA-V (and Sentinel-3).

Fang et al. (2014) concluded that MCD15A2 and GEOv1 are rather consistent. This can at least partly be explained by the dependence of these datasets, and MCD15A2 was used to train the NN retrieval from SPOT-VGT in GEOv1.

Table 2 Existing datasets for LAI and FAPAR

	% of survey respondent using this product	**# Publications using the product (WoS)**	**Sensors**		**Method**	**Clumping**	**Biome specifici**	**Reference**
MCD15A2H, LAI	80	1958	MODIS	1 km / 8 d	3D RTM LUT/ biome specific		X	(Knyazikhin et al., 1998)
MCD15A2H, FAPAR		533	MODIS	1 km/8 d			X	
CGLS (GEOv1-3)	37	338	SPOT-VGT	~1 km/ 10 d	NN calibration to MODIS and Cyclopes		*	(Baret et al., 2013)
GLASS-LAI	4	138	MODIS/ AVHRR	250 m / 8d	GRNN calibration to MODIS and Cyclopes	X	X	Ma and Liang, (Ma & Liang, 2022)2022
CYCLOPES	-	99	POLDER	1/112 deg/ 10 d	Clumping			(Baret et al., 2007)
QA4ECV-LAI	10	78	AVHRR	0.05deg/ 1 d	NN calibration to MCD15A2H			(Franch et al., 2017)

JRC-FAPAR	15	35	theSeaWiFS	2km/10d				(Gobron et al., 2006)
GLOBCARBON	-	15	AVHRR	0.05deg/ 1 d	NN calibration to MCD15A2H		X	(Plummer et al., 2006)
GLOBMAP	-	26		0.5 km/ 8 d			X	(Liu et al., 2012)
JRC-TIP	-	14	MODIS MISR		Albedo / AD			(Pinty et al., 2011)
MTLAI / LSA-SAF LAI**	10	27	MSG SEVIRI	12/3km				(García-Haro et al., 2019)
C3S	40	6	PROBA-V	10d/1km				(Blessing & Giering, 2010)

*inherited from MODIS 15 A, ** not global (incl. Africa, Europe, Brazil)

The users consider the spatial resolution of the products (in most cases 1 km) as a strong point, although some users (3/10) consider the spatial resolution require higher resolution, notably users in the domains of monitoring, early warning, agriculture and specific research at flux tower sites. The accessibility of the data, the length of the time series are also strengths of most of the products. Some consider the consistency of the time series as a strength.

The following bottlenecks of FAPAR and LAI products have been identified

Inter-product diverges in terms of mean, standard deviation and seasonal cycle, and long-term trends (Fang et al., 2013; Mota et al., 2021), In some cases, the differences among data products of LAI and FAPAR is larger than the differences among time series of spectral indices (NDVI) derived from alternative sensors, which led some users decide to use NDVI time series as a qualitative signal instead of a data product of LAI and FAPAR.

Discrepancies between the data products and field data of LAI and FAPAR. Scientists working in field research expressed most concerns with the correspondence with independent field data. For products that use a land-cover specific algorithm, the underlying land cover classification does not always match with reality.

Documentation. Limited documentation and clear metadata pose limitations on the correct interpretation of the data in some cases. The metadata include explicit description of the underlying assumptions on leaf optical properties, canopy structure and representation of clumping, black or white sky FAPAR, definition of ancillary data (such as flags).

Handshaking with the land surface models. While most users emphasize that the true rather than the effective LAI is required in the models, the exact use of FAPAR and LAI diverges among land surface models (JULES, ORCHIDEE, CLM, SiB). Some of these models use LAI as a variable in a simplified radiative transfer model to estimate photosynthesis. However, these model representations diverse among LSMs. For example, both JULES and ISBA include a representation of clumping, but these are not identical. To facilitate the use of the products, it is essential that the algorithm is transparent and that it is possible to achieve a handshake between the LSMs and the product.

Temporal resolution

The decadal temporal resolution is the lowest limit for phenology studies. Gaps in the time series due to cloud cover and residual contamination by snow limit the applicability in phenology.

Long term consistency
The (dis)continuity between sequential satellite missions is a constraint for climate modelling. A specific example is a discontinuity in GEOV2 (but not in GEOV1) between SPOT-VGT and PROBA-V. In addition, degrading sensor sensitivity may play a role, but the magnitude is not known.

Data latency
For services and early warning systems, the data latency is a limitation. For these applications near-real time data provision is necessary (latency of approximately 2 days). For other applications, a latency of 3 months is sufficient. The users would be interested in including 2021 and 2022 in the data sets of this project, in order to be able to analyse extreme events in those years.

3.3 Requirements

In

Table 3, we summarize general requirements on temporal and spatial aspects of the data set.

Table 3 Data requirements with respect to space and time for the five applications mentioned in Section 3.1.

	Time			Space	
	Span	**Time step**	**Latency**	**Span**	**Resolution**
Climate	Decades	Monthly	Years	Global	km-deg
Weather Extremes	Multiple years	(multi-) daily	~3 months	Regional	1000 m
Phenology	Multiple years	(multi-) daily	~3 months	Global	1000 m
Monitoring services	-	(multi-) daily	Near Real Time	Regional	300 m
Flux sites	Multiple years	(multi-) daily	~3 months	Local, API access	<300 m

3.4 Priorities identified by users

The users were confronted with a number of dilemmas, in which 100 points had to be distributed over pairs of requirements that are from a technical point of view difficult, if not impossible, to meet both at the same time:

Length of the time series versus spatial resolution
Due to the availability of satellite data, the length of the time series and the spatial resolution cannot be both maximized in the same dataset. While the preference obviously depends on the application (See Sections 3.1 and 3.2), the users who contributed to the survey prefer long and consistent (70/100) over spatial resolution. In terms of consistency, two aspects are most important: sensor dependence of the data set (long-term consistency), and scientific and retrieval consistency with other data products and between LAI and FAPAR. At least, it must be possible to trace differences among alternative LAI products to the underlying assumptions.

Uncertainty versus temporal resolution
Because the number of observations scales with the time window that is used in the retrieval, the uncertainty has a negative relationship with the time window. Most users prefer a higher temporal

resolution over a lower uncertainty. Furthermore, users interested in disturbances (abrupt events) prefer data that are not smoothed in time.

Radiative transfer versus statistical and empirically based retrieval
Concerning the algorithm, the majority of users have a preference for retrieval with a physically based model over retrieval with machine learning algorithms. However, this does not hold for all users. Concerns with radiative transfer models are the unrealistic representation of vegetation, and the low correlation with field measured equivalents of LAI and FAPAR, which may be considerably better with a well-trained statistical model (in addition to an RTM) such as GEOv1 or GLASS.

3.5 User recommendations

The users were requested to provide recommendations, in an field of the survey. The following recommendations have been provided:

General
- Absolute transparency about the algorithm and assumptions, especially on issues of clumping, white/black sky
- Ensure physical consistency with definitions of these products used land surface models (and with other products) or document the product in such a way that land surface models can be adapted towards a more consistent (i.e., in terms of radiative transfer) coupling of state variables with the offered data products.
- Carry out intercomparison experiments

Algorithm
- Consider the effect of the soil
- De-couple effect of seasonally varying chlorophyll content and LAI on FAPAR, or provide a ‘green FAPAR’ product
- If we use a land-cover specific algorithm, ensure consistency with CCI-Land Cover
- Perform a correction for the effect of snow, which is in particular relevant at the start of the growing season.
- Do not filter out potentially discontinuities in the time series that may be real (e.g., due to harvest or fire).
- Strive for a product that is less dependent on the sensor.
- The dependence of the FAPAR on the viewing and solar geometry is considered in the algorithm. Because of differences among LSMs, there may not be an ideal choice for all applications. It is therefore of great importance that the algorithm is well documented, and that it is designed such (including metadata) that it is possible to carry out transformations on the data afterwards to make them agree with the LSM.

Validation
- Test and document the consistency with other data products that are available, and the sensor dependence of the data set
- Compare radiative transfer derived data to statistically and empirically derived values, in order to avoid a data product that does not represent the reality in the field. This is particularly relevant for biomes for which the radiative transfer model may not apply.

Data access
- Provide subsets at specific sites (e.g., FLUXNET), and an API for data access

3.6 Innovations

The users have been requested to comment on and rate the importance of possible innovations. In the dialogue of the interviews, the other potential innovations have been discussed as well.

All participants in the survey identified 'accounting for clumping' as at least of limited importance, and 80% rated it as important or very important. Clumping affects the ratio between true and apparent LAI, where the latter is an effective optical LAI assuming complete vegetation cover in the footprint of the observation, which some of the current data products account for, such as MODIS15A2. Despite the importance, some users have expressed concerns with using clumping factors as they may further complicate the interpretation of the dataset, contribute to ill-posed retrievals, not be realistic enough, and not necessarily improve the consistency with the state variables of LAI and FAPAR in land surface models.

Solutions for mixed pixels have been identified as important or very important by 67% of the respondents.

Differentiation by pigments (e.g., chlorophyll) has been identified as important (27%) or very important (21%) by 48%

Consistency of the algorithm with the radiative transfer of SIF used in the SCOPE model, or the possibility to include SIF in the retrieval has been identified as important (38%) and very important (11%) by 49%.

Another innovative direction is the direct use of Level-1 satellite data in data assimilation with DGVMs. This makes it possible to avoid the step of using higher level satellite data such as LAI and FAPAR altogether. However, this development is still in an early stage, and the computational demands of radiative transfer models are prohibitive.

Users identify the need for a differentiation between both black-sky and white-sky FAPAR, with a preference for 'blue sky' FAPAR product is needed that takes into account the fraction of diffuse and direct radiation. Land surface models may differ in radiative transfer representation (e.g., CLM, SiB, ORCHIDEE, ISBA and JULES). For example, ISBA makes use of a constant fraction of diffuse radiation and would benefit from a black and white sky FAPAR product, however, in the fraction of diffuse radiation is treated as a constant.

4 Priorities and feasibility for this project

Priority 1. Provide transparency of the processing chain through documentation and ancillary data
Feasibility: high
Two requirements for reproducibility include clarity of the documentation and completeness of the ancillary data. The clarity of the algorithm description will be evaluated by collecting feedback from beta users in the user requirement study in cycles 2 and 3.
Reproducibility of the algorithm also requires that we document ancillary data, for example the soil spectra and pigment contents that are retrieved along with LAI and FAPAR. While this is technically feasible and it requires very little additional computation power, it does require storage space.

Priority 2. Ensure physical consistency between LAI and FAPAR
Feasibility: high
The physical consistency between LAI and FAPAR can be achieved by using an identical radiative transfer scheme for the retrieval of both. The algorithms initially proposed already qualify in this respect, and no additional development would be needed to meet this requirement.

Priority 3. Account for snow and soil effects
Feasibility: medium
Accounting for soil background and snow cover is highly relevant for phenology studies, where accurate LAI and FAPAR retrievals are required at the beginning and the end of the growing season when the soil is partly exposed. Snow is also a state variable of some land surface models, and retrieval of (below canopy) snow cover can be of added value. A challenge is that the freedom of surface elements in the model parameterization lead to less constraint inversions if the number of independent observations is limited. The algorithm should strike the balance between well-constraint retrievals, and a right level of detail in describing the variability of soil and snow background spectra.

Priority 4. Compare with existing data products of LAI and FAPAR.
Feasibility: high
Assessing the added value through intercomparison with available products in an essential component of the product evaluation: It will reveal inconsistencies that can be traced to their root causes (input data and/or algorithm).

Priority 5. Derive sensor-independent products.
Feasibility: medium
Although the users stress the importance products that do not depend on the properties of the sensors, this can only be achieved partially. The availability of data will dictate how well the inversions are constraint. Furthermore, this project relies on available information on sensor degradation over the lifetime of the instruments. The sensor dependence can be reduced in two ways: by using a spectrally and (viewing) angularly continuous radiative transfer model, and (2) by assessing the influence of input data availability in overlapping periods.

Priority 6. Consider clumping of the vegetation
Feasibility: medium-low
Users stress the importance of providing true LAI (i.e. comparable to field data) rather than effective LAI. Two scales of clumping can be identified: (1) at larger scale partial vegetation cover plays a role, where vegetation is distributed according to land cover within a pixel, and (2) the 3D structure of the vegetation in twigs and crowns. In both cases, a turbid medium representation would underestimate the true LAI.
Accounting for clumping involves a number of challenges, notably:

(1) There is no consensus in the scientific community on the model to use, and the choice of a model may result in the loss of handshaking with specific vegetation models
(2) Additional parameters are needed, and this results in the poorly constrained inversions. Coarse resolution data do not contain information that allow the differentiation of a lower fractional cover from a lower overall LAI.

In theory, the second problem can be overcome by including information from higher spatial resolution, such as a land cover classification, possibly in combination with vegetation type specific radiative transfer models. However, this may compromise priorities 1 (reproducible and transparent algorithm) and 5 (sensor-independent algorithm) due to dependence on ancillary higher level data products as input. Alternatively, the magnitude of the problem can be investigated and tools for post processing of the data products could be provided that enable the smart use of the FAPAR and LAI products in combination with vegetation CCI land cover product.

Priority 7. Achieve a temporal resolution better than 10 days.

Feasibility: medium

Achieving a temporal resolution better than 10 days is listed as GCOS action (T40), and a wish of users who focus on phenology. A physical limitation is the cloud cover, but a multi-sensor approach increases the chance of cloud-free images. A high temporal resolution is particularly relevant for higher latitudes, where snow and phenology indices (SOS, EOS) are important indicators, and luckily, the data availability of polar orbiters is higher at high latitudes. The compromise between a short time window for the retrieval window and the accuracy of the products needs to be investigated.

Priority 8. Retrieval pigments and green FAPAR

Feasibility: medium

The interest in using chlorophyll and other pigments in land surface modelling (Croft et al., 2017; Luo et al., 2019). Technically, a physically based retrieval algorithm such as initially proposed for this project (OptiSAIL) can provide pigment concentrations with little additional computational effort, but it requires additional storage capacity.

A challenge of retrieving such additional data product is to validate its accuracy at the spatial scale of the data products. However, it would be possible to (1) evaluate the products at smaller spatial scale, using data of dedicated field experiments in which TOC reflectance and pigment content were both measured, and (2) intercompare with products such as derived by (Croft et al., 2017). The retrieval of ‘green FAPAR’, e.g. the FAPAR of Chlorophyll only, can be achieved as well, but this would require a (minor) upgrade of the radiative transfer code to accommodate such output product, similar to SCOPE (Yang et al., 2021).

Priority 9. Assess consistency with other data products, such as land cover, burnt area.

Feasibility: high

The users identify the need for consistency with other data products, in particular land cover. Including such products as input the algorithm compromises priorities (1) and (5), but the consistency with other data products can be included in the science studies included in the project in cycles 2 and 3 (see PMP).

Priority 10. Provide forward simulated of SIF

Feasibility: medium

The use of a radiative transfer model for the inversion also enables the inclusion of solar-induced chlorophyll fluorescence (SIF) as additional output. The technical challenge is that SIF is wavelength and geometry (illumination-viewing) dependent. A further user consultation would be needed in the second cycle to arrive at a scientifically meaningful product, for example a SIF product for the Sentinel-5P geometry.

5 References

Albergel, C., Zheng, Y., Bonan, B., Dutra, E., Rodríguez-Fernández, N., Munier, S., Draper, C., de Rosnay, P., Muñoz-Sabater, J., Balsamo, G., Fairbairn, D., Meurey, C., & Calvet, J. C. (2020). Data assimilation for continuous global assessment of severe conditions over terrestrial surfaces. *Hydrology and Earth System Sciences*, *24*(9), 4291–4316. https://doi.org/10.5194/HESS-24-4291-2020

Balzarolo, M., Vicca, S., Nguy-Robertson, A. L., Bonal, D., Elbers, J. A., Fu, Y. H., Grünwald, T., Horemans, J. A., Papale, D., Peñuelas, J., Suyker, A., & Veroustraete, F. (2015). *Matching the phenology of Net Ecosystem Exchange and vegetation indices estimated with MODIS and FLUXNET in-situ observations*. https://doi.org/10.1016/j.rse.2015.12.017

Baret, F., Hagolle, O., Geiger, B., Bicheron, P., Miras, B., Huc, M., Berthelot, B., Niño, F., Weiss, M., Samain, O., Roujean, J. L., & Leroy, M. (2007). *LAI, fAPAR and fCover CYCLOPES global products derived from VEGETATION Part 1: Principles of the algorithm*. https://doi.org/10.1016/j.rse.2007.02.018

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 Sensing of Environment*, *137*, 299–309. https://doi.org/10.1016/J.RSE.2012.12.027

Baruth, B., Royer, A., Klisch, A., & Genovese, G. (n.d.). *THE USE OF REMOTE SENSING WITHIN THE MARS CROP YIELD MONITORING SYSTEM OF THE EUROPEAN COMMISSION*. Retrieved September 15, 2022, from http://mars.jrc.it/marsstat/Crop_Yield_Forecasting/crop_yield_f

Blessing, S., & Giering, R. (2010). *Algorithm Theoretical Basis Document PROBA-V CDR and ICDR LAI and fAPAR v2.0*.

Bonan, B., Albergel, C., Zheng, Y., Lavinia Barbu, A., Fairbairn, D., Munier, S., & Calvet, J. C. (2020). An ensemble square root filter for the joint assimilation of surface soil moisture and leaf area index within the Land Data Assimilation System LDAS-Monde: Application over the Euro-Mediterranean region. *Hydrology and Earth System Sciences*, *24*(1), 325–347. https://doi.org/10.5194/HESS-24-325-2020

Bórnez, K., Descals, A., Verger, A., & Peñuelas, J. (2020). Land surface phenology from VEGETATION and PROBA-V data. Assessment over deciduous forests. *International Journal of Applied Earth Observation and Geoinformation*, *84*. https://doi.org/10.1016/J.JAG.2019.101974

Cammalleri, C., Verger, A., Lacaze, R., & Vogt, J. v. (2019). Harmonization of GEOV2 fAPAR time series through MODIS data for global drought monitoring. *International Journal of Applied Earth Observation and Geoinformation*, *80*, 1–12. https://doi.org/10.1016/J.JAG.2019.03.017

Croft, H., Chen, J. M., Luo, X., Bartlett, P., Chen, B., & Staebler, R. M. (2017). Leaf chlorophyll content as a proxy for leaf photosynthetic capacity. *Global Change Biology*, *23*(9), 3513–3524.

Fang, H., Jiang, C., Li, W., Wei, S., Baret, F., Chen, J. M., Garcia-Haro, J., Liang, S., Liu, R., Myneni, R. B., Pinty, B., Xiao, Z., & Zhu, Z. (n.d.). Characterization and intercomparison of global moderate resolution leaf area index (LAI) products: Analysis of climatologies and theoretical uncertainties. *J. Geophys. Res. Biogeosci*, *118*, 529–548. https://doi.org/10.1002/jgrg.20051

Fang, H., Jiang, C., Li, W., Wei, S., Baret, F., Chen, J. M., Garcia-Haro, J., Liang, S., Liu, R., Myneni, R. B., Pinty, B., Xiao, Z., & Zhu, Z. (2013). Characterization and intercomparison of global moderate resolution leaf area index (LAI) products: Analysis of climatologies and theoretical uncertainties. *J. Geophys. Res. Biogeosci*, *118*, 529–548. https://doi.org/10.1002/jgrg.20051

Franch, B., Vermote, E. F., Roger, J. C., Murphy, E., Becker-reshef, I., Justice, C., Claverie, M., Nagol, J., Csiszar, I., Meyer, D., Baret, F., Masuoka, E., Wolfe, R., & Devadiga, S. (2017). A 30+ Year AVHRR Land Surface Reflectance Climate Data Record and Its Application to Wheat Yield Monitoring. *Remote Sensing 2017, Vol. 9, Page 296*, *9*(3), 296. https://doi.org/10.3390/RS9030296

García-Haro, F. J., Camacho, F., Martínez, B., Campos-Taberner, M., Fuster, B., Sánchez-Zapero, J., & Gilabert, M. A. (2019). Climate Data Records of Vegetation Variables from Geostationary

SEVIRI/MSG Data: Products, Algorithms and Applications. *Remote Sensing 2019, Vol. 11, Page 2103*, *11*(18), 2103. https://doi.org/10.3390/RS11182103

Gobron, N., Pinty, B., Aussedat, O., Chen, J. M., Cohen, W. B., Fensholt, R., Gond, V., Huemmrich, K. F., Lavergne, T., Mélin, F., Privette, J. L., Sandholt, I., Taberner, M., Turner, D. P., Verstraete, M. M., & Widlowski, J. L. (2006). Evaluation of fraction of absorbed photosynthetically active radiation products for different canopy radiation transfer regimes: Methodology and results using Joint Research Center products derived from SeaWiFS against ground-based estimations. *Journal of Geophysical Research: Atmospheres*, *111*(D13), 13110. https://doi.org/10.1029/2005JD006511

Kaminski, T., Knorr, W., Scholze, M., Gobron, N., Pinty, B., Giering, R., & Mathieu, P.-P. (2012). Consistent assimilation of MERIS FAPAR and atmospheric CO 2 into a terrestrial vegetation model and interactive mission benefit analysis. *Biogeosciences*, *9*, 3173–3184. https://doi.org/10.5194/bg-9-3173-2012

Knyazikhin, Y., Martonchik, J. v., Myneni, R. B., Diner, D. J., & Running, S. W. (1998). Synergistic algorithm for estimating vegetation canopy leaf area index and fraction of absorbed photosynthetically active radiation from MODIS and MISR data. *Journal of Geophysical Research: Atmospheres*, *103*(D24), 32257–32275. https://doi.org/10.1029/98JD02462

Lafont, S., Zhao, Y., Calvet, J.-C., Peylin, P., Ciais, P., Maignan, F., & Weiss, M. (2012). Modelling LAI, surface water and carbon fluxes at high-resolution over France: comparison of ISBA-A-gs and ORCHIDEE. *Biogeosciences*, *9*, 439–456. https://doi.org/10.5194/bg-9-439-2012

Liu, Y., Liu, R., & Chen, J. M. (2012). Retrospective retrieval of long-term consistent global leaf area index (1981–2011) from combined AVHRR and MODIS data. *Journal of Geophysical Research: Biogeosciences*, *117*(G4), 4003. https://doi.org/10.1029/2012JG002084

Luo, X., Croft, H., Chen, J. M., Liming He, |, & Keenan, T. F. (2019). Improved estimates of global terrestrial photosynthesis using information on leaf chlorophyll content. *Glob Change Biol*, *25*, 2499–2514. https://doi.org/10.1111/gcb.14624

Ma, H., & Liang, S. (2022). Development of the GLASS 250-m leaf area index product (version 6) from MODIS data using the bidirectional LSTM deep learning model. *Remote Sensing of Environment*, *273*. https://doi.org/10.1016/J.RSE.2022.112985

Macbean, N., Maignan, F., Peylin, P., Bacour, C., Bréon, F.-M., & Ciais, P. (2015). Using satellite data to improve the leaf phenology of a global terrestrial biosphere model. *Biogeosciences*, *12*, 7185–7208. https://doi.org/10.5194/bg-12-7185-2015

Maes, W. H., Pagán, B. R., Martens, B., Gentine, P., Guanter, L., Steppe, K., Verhoest, N. E. C., Dorigo, W., Li, X., Xiao, J., & Miralles, D. G. (2020). *Sun-induced fluorescence closely linked to ecosystem transpiration as evidenced by satellite data and radiative transfer models*. https://doi.org/10.1016/j.rse.2020.112030

Mota, B., Gobron, N., Morgan, O., Cappucci, F., Lanconelli, C., & Robustelli, M. (2021). Cross-ECV consistency at global scale: LAI and FAPAR changes. *Remote Sensing of Environment*, *263*, 112561. https://doi.org/10.1016/J.RSE.2021.112561

Nunes, E. L., Costa, M. H., Malhado, A. C. M., Dias, L. C. P., Vieira, S. A., Pinto, L. B., & Ladle, R. J. (2012). Monitoring carbon assimilation in South America's tropical forests: Model specification and application to the Amazonian droughts of 2005 and 2010. *Remote Sensing of Environment*, *117*, 449–463. https://doi.org/10.1016/J.RSE.2011.10.022

Pinty, B., Andredakis, I., Clerici, M., Kaminski, T., Taberner, M., Verstraete, M. M., Gobron, N., Plummer, S., & Widlowski, J. L. (2011). Exploiting the MODIS albedos with the Two-stream Inversion Package (JRC-TIP): 1. Effective leaf area index, vegetation, and soil properties. *Journal of Geophysical Research: Atmospheres*, *116*(D9), 9105. https://doi.org/10.1029/2010JD015372

Plummer, S., Arino, O., Simon, M., & Steffen, W. (n.d.). *Mitigation and Adaptation Strategies for Global Change ESTABLISHING A EARTH OBSERVATION PRODUCT SERVICE FOR THE TERRESTRIAL CARBON COMMUNITY: THE GLOBCARBON INITIATIVE*.

Popp, T., Hegglin, M. I., Hollmann, R., Ardhuin, F., Bartsch, A., Bastos, A., Bennett, V., Boutin, J., Brockmann, C., Buchwitz, M., Chuvieco, E., Ciais, P., Dorigo, W., Ghent, D., Jones, R., Lavergne, T., Merchant, C. J., Meyssignac, B., Paul, F., … Willén, U. (2020). Consistency of Satellite Climate Data Records for Earth System Monitoring. *Bulletin of the American Meteorological Society*, *101*(11), E1948–E1971. https://doi.org/10.1175/BAMS-D-19-0127.1

Wu, M., Scholze, M., Voßbeck, M., Kaminski, T., & Hoffmann, G. (2018). *remote sensing Simultaneous Assimilation of Remotely Sensed Soil Moisture and FAPAR for Improving Terrestrial Carbon Fluxes at Multiple Sites Using CCDAS*. https://doi.org/10.3390/rs11010027

Yang, P., Prikaziuk, E., Verhoef, W., & van der Tol, C. (2021). SCOPE 2.0: A model to simulate vegetated land surface fluxes and satellite signals. *Geoscientific Model Development*, *14*(7), 4697–4712. https://doi.org/10.5194/GMD-14-4697-2021