CCI+ Vegetation Parameters

D2.2 End-to-End ECV Uncertainty Budget (E3UB)

Simon Blessing, Dominique Jolivet

September 2023

Distribution list

Author(s)	:	Simon Blessing, Dominique Jolivet
Reviewer(s)	:	Else Swinnen, Christiaan Van der Tol
Approver(s)	:	Clément Albergel
Issuing authority	:	VITO

Change record

Release	Date	Pages	Description of change	Editor(s)/Reviewer(s)
V1.0	2023-09-14	all	First version	Simon Blessing, Dominique Jolivet/ Christiaan Van der Tol, Else Swinnen

Executive summary

The uncertainty budget, from the actual measurement with multiple sensors to the data product, of the vegetation parameters jointly retrieved for the CRDP-1 of the CCI+ Vegetation Parameters is explained in this document. Sources of uncertainty for LAI, *f*APAR, leaf Chlorophyll-A+B concentration ("Cab"), *f*APAR_Cab, and surface albedo are identified, quantified as far as possible, and it is documented whether the respective source is reflected in the uncertainty estimate provided with the CRDP-1. In addition, the existence and importance of uncertainty correlations is highlighted, which are available from the OptiSAIL retrieval system used for CRDP-1.

Table of Contents

List of Acronyms

BHR	Bi-Hemispherical Reflectance
BRDF	Bidirectional Reflectance Distribution Function
CCI+	Climate Change Initiative Plus
CRDP	Climate Resarch Data Package
DHR	Directional-Hemispherical reflectance
ECV	Essential Climate Variable
ED	External Document (as listed in section 1.2)
EOF	Empirical Orthogonal Function
fAPAR	*f*raction of Absorbed Photosynthetically Active Radiation
ID	Internal Document (as listed in section 1.2)
LAI	Leaf Area Index
NIR	Near Infra-Red range of the electromagnetic spectrum, here 700--2500 nm
PCA	Principal Components Analysis
PROSPECT	PROperties of leaf SPECTtra
RT	Radiative Transfer
SAIL	Scattering of Arbitrarily Inclined Leaves
TAF	Transformation of Algorithms in Fortran
TARTES	Two-streAm Radiative TransfEr in Snow
TIP	Two-stream Inversion Package
TOA	Top-Of-Atmosphere
TOC	Top-Of-Canopy
VIS	VISible range of the electromagnetic spectrum, here 400–700 nm
VP	Vegetation Parameters

List of Figures

1 Introduction

1.1 Scope of this document

This document describes the uncertainty characterization, estimation and/or propagation for each product in the ECV included in CCI vegetation parameters Climate Data Research Package of cycle 1 (CRDP-1). It can be regarded as complementary to the Product Validation and Intercomparison Report (VP-CCI_D4.1_PVIR_V1.0).
Here, Leaf Area Index (LAI) and *f*raction Absorbed of Photosynthetically Active Radiation (fAPAR) are retrieved together with many other parameters from optical sensors using OptiSAIL.
Details on the methodology to determine per-observation uncertainty products, and how they are presented to users of the CRDP-1, are provided in this document.

1.1.1 CRDP-1 as a true multi-sensor product

In CRDP-1 reflectance data from SPOT-4/VEGETATION (VGT1 hereafter), SPOT-5/VEGETATION-2 (VGT2 hereafter), and Proba-V/VEGETATION (PROBA-V hereafter) is used. While between VGT1 to VGT2 the data source is just switched with the availability of VGT-2 data, the overlap phase of VGT2 and PROBA-V is used to produce a true multi-sensor product. For this combination, the results are expected to benefit from the multi-sensor approach mainly through higher temporal sampling. Wherever orbital and instrumental differences occur, also the higher angular and wavelength sampling should improve the quality of the retrieval. This is because the radiative transfer models model the directional and spectral reflectance explicitly as a function of LAI (and other vegetation parameters).

1.2 Related documents

Internal documents

Reference ID	Document
ID1	Climate Change Initiative Extension (CCI+) Phase 2 New ECVs: Vegetation Parameters – EXPRO+ (ITT)
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_D4.1_PVIR_V1.2	Product Validation and Intercomparison Report (PVIR) CRDP-1, ESA CCI+ Vegetation Parameters http://climate.esa.int/media/documents/VP-CCI_D4.1_PVIR_V1.2.pdf
VP-CCI_D4.2_PUG_V1.2	Product User Guide (PUG) CRDP-1, ESA CCI+ Vegetation Parameters http://climate.esa.int/media/documents/VP-CCI_D4.2_PUG_V1.2.pdf

External documents

Reference ID	Document
ED-1	C3S ATBD Multi sensor CDR Surface Albedo v2.0

1.3 General definitions

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

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

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

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

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

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

2 Methodology

The methods for computing and propagating uncertainties referred to in this document, all assume that errors are small and Gaussian. All inversion steps use Bayesian inference by minimising a cost function (J) with two parts, a prior term, which contains the prior knowledge on the parameters, and a data term, which measures the misfit with the observations. The posterior parameter values and their uncertainties are the output of the inversion. The LAI is among the retrieved parameters, but fAPAR and surface albedo are computed in a diagnostic step following the retrieval using the posterior model parameters and the illumination geometry. The covariance matrix of the data (input to the inversion) is part of the data term, while the covariance matrix describing the prior knowledge of the parameters is part of the prior term. It is the curvature of this cost function at the minimum, which describes how well the combination of data and prior knowledge constrains the retrieved parameters, hence how uncertain they are. Therefore, the inverse of the Hessian (H) at this point gives the posterior covariance matrix ($\Sigma_{post} = H^{-1}$).
The propagation of uncertainties through diagnostic steps, which do not involve a model inversion, are done using the covariances and the Jacobians ($\Sigma_{\text{out}} = (d\,J/d\,p)\Sigma_{\text{in}}(d\,J/d\,p)^T$).

For OptiSAIL, Hessian and Jacobian matrices are computed directly from the implementations of the models by using automatic differentiation (AD).
When various measurements with their respective uncertainties are combined, it is important to take their covariance into account. OptiSAIL is equipped to do this, but the typical signature of the TOC reflectance inter-band covariance still needs to be determined, before this feature can be exploited. Co-variances of the atmospheric correction from an operational chain are generally not available, also because of the high data-volume this would imply. OptiSAIL does however compute the correlation of the posterior uncertainty between all results. Figure 1 shows an example from 2014-05-16 for the central-European tile X18Y02 (in PROBA-V nomenclature) of quality-filtered LAI, fAPAR, their uncertainties, and Figure 2 shows their correlation, and the number of per-band-observations of VGT2 and PROBA-V which were combined to retrieve these values. A cut-off at a maximum of three observations per band per sensor was applied, hence the maximum of 24 (=2 sensors *4 bands per sensor * 3). The visibility of some of the swath edges is caused by the quality filtering (different degree of cloud contamination at the different overpass times; for quality filtering see VP-CCI_D4.2_PUG_V1.0). Note that it is not an effect of a jump in the value but of the different density of accepted inversions (the locations of missing values are coloured black in the LAI and fAPAR images in Figure 1).

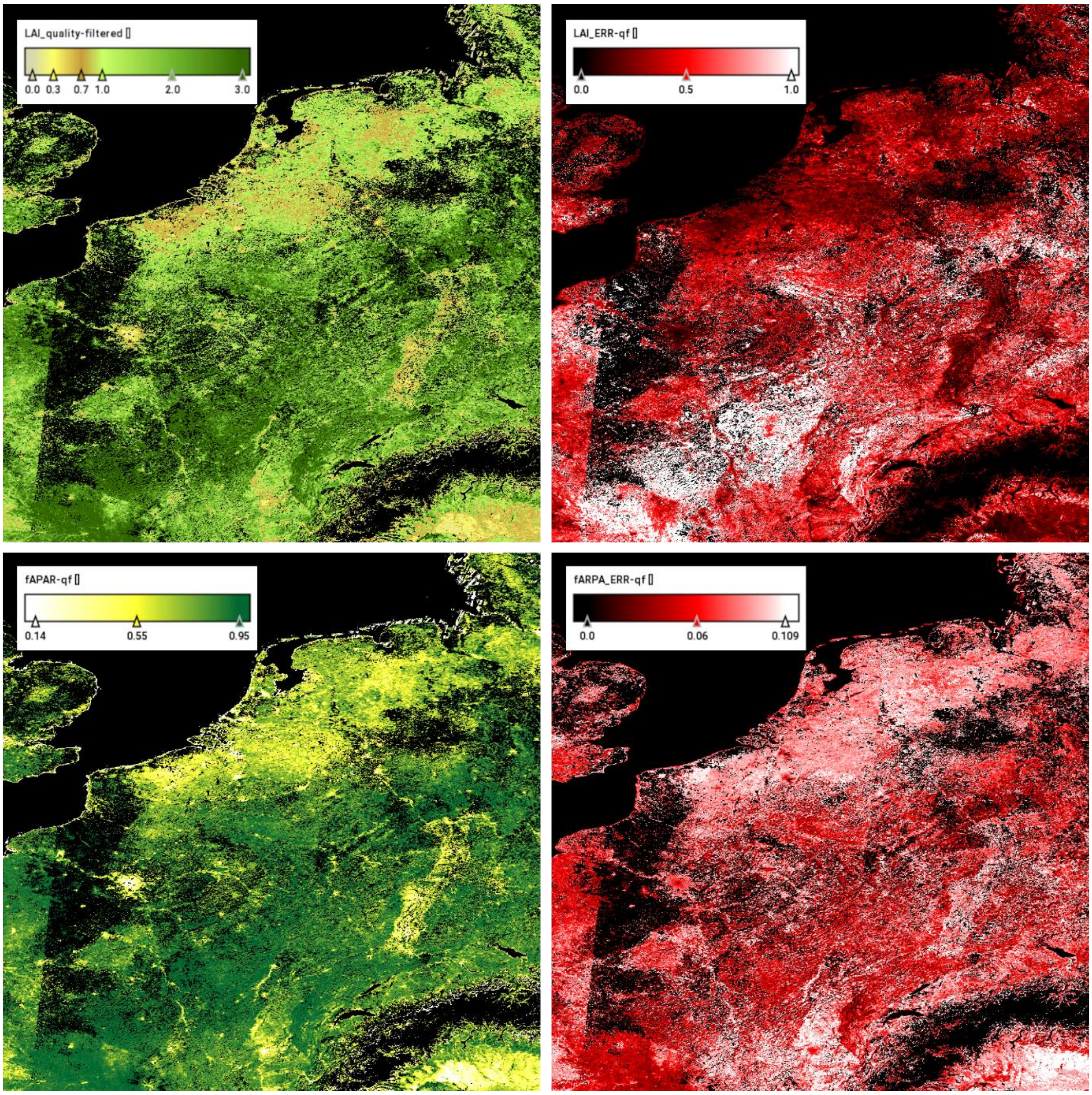


Figure 1: Quality-filtered ("-qf") OptiSAIL LAI, LAI uncertainty, fAPAR, fAPAR uncertainty for 2014-05-16 for the central-European tile X18Y02.

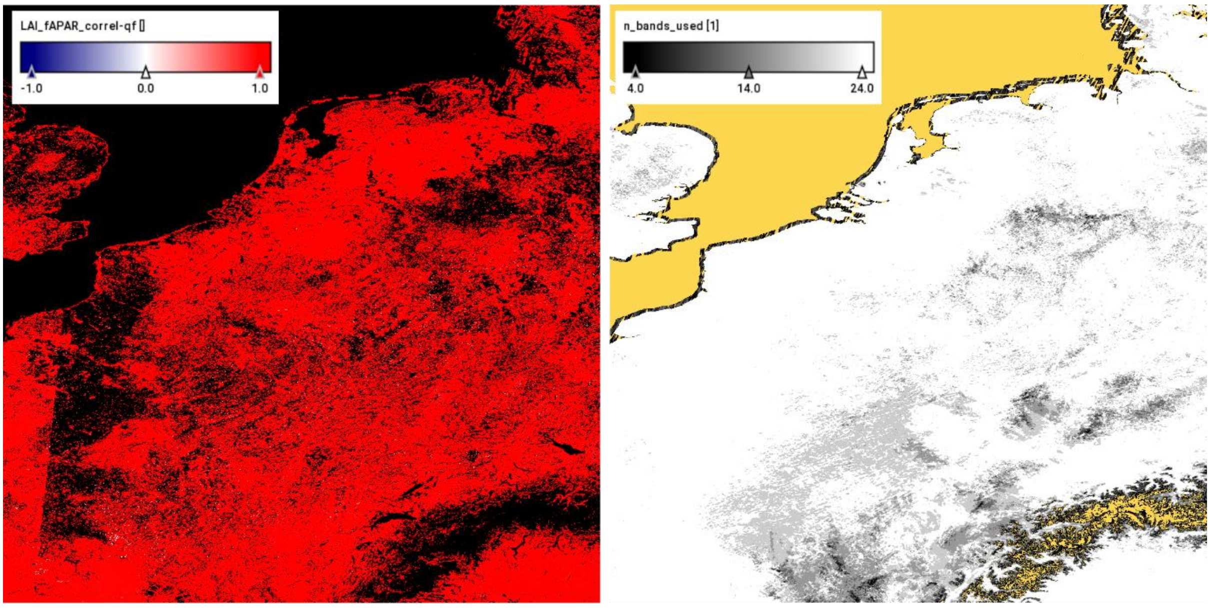


Figure 2: Correlation of the Quality-filtered ("-qf") OptiSAIL LAI uncertainty and fAPAR uncertainty (left) and the number of bands used (right) for 2014-05-16 for the central-European tile X18Y02.

3 TOC reflectances

OptiSAIL uses TOC reflectances derived from various optical sensors as input. TOC reflectances are obtained by applying SMAC (Rahman and Dedieu, 1994), a Simplified Method for the Atmospheric Correction of satellite measurements in the solar spectrum to the TOA reflectances. The choice of the SMAC algorithm is supported by the following arguments:

- It is operational and largely used in the land community and already implemented in the Copernicus Global Land Service and Copernicus Climate Change Service processing lines.
- It is a robust, generic algorithm; thus, it minimizes the dependence on the sensor which is a good thing when one wants to build a multi-sensor long time series with limited biases.
- The formulation of the algorithm is analytical and is adapted to an error propagation analysis

At that point, numerous sources of uncertainty are involved already:

1. Uncertainty of the measurements of the optical sensors
2. Uncertainty due to the aggregation and the geolocation
3. Uncertainties of the ancillary data (ozone, water vapour, surface pressure and Aerosol Optical Thickness) used in the atmospheric correction. Ancillary data are from the Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2). More information can be found in https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/
4. Uncertainty due to the atmospheric correction model.

For CRDP-1, the TOC reflectances from Proba-V and SPOT-VEGETATION that were processed in the frame of the C3S_312b contract for the multi-sensor data are used. The description of this data can be found in ED-1.
The sources of uncertainty 1 and 3 are accounted for in the uncertainty budget. How they are propagated is detailed in Section 3.4.3 of ED-1 while how they are characterized is presented in section 3.4.4 and Table 19 of ED-1.

The sources of uncertainty 2 and 4 are not accounted for in the uncertainty budget.

4 LAI

LAI uncertainty from OptiSAIL is computed from the posterior covariance Matrix of the model parameters as described in the ATBD, thus taking into account the uncertainties reported for the TOC reflectances in the cost function.

4.1 Sources of uncertainty of OptiSAIL LAI

In addition to the TOC reflectance uncertainty, the following sources of uncertainties influence the retrieval of OptiSAIL LAI:

(a) Correlation of TOC reflectance uncertainties

All bands of a sensor which are using the same observation geometry suffer in a similar way from algorithmic and parametric uncertainties of the atmospheric correction. Therefore, their uncertainties are correlated.

(b) TOC reflectance geolocation and spatial sampling

Satellite products with known geolocation issues are avoided. However, especially in higher latitudes where the product grid is much smaller than 1 km^2 per pixel, geolocation uncertainties of the input data are expected to lead to a notable level of noise in the retrieval.

(c) Algorithmic uncertainty of the OptiSAIL simulation models (4SAILH, PROSPECT-D, Ross-Li BRDF, TARTES, soil model)

The models used for the simulation of the reflectance spectra within OptiSAIL are to a high degree idealised in order to allow for a minimum of ancillary data and very high computational speed.

4.2 Quantification of uncertainty of OptiSAIL LAI

(a) Correlation of TOC reflectance uncertainties

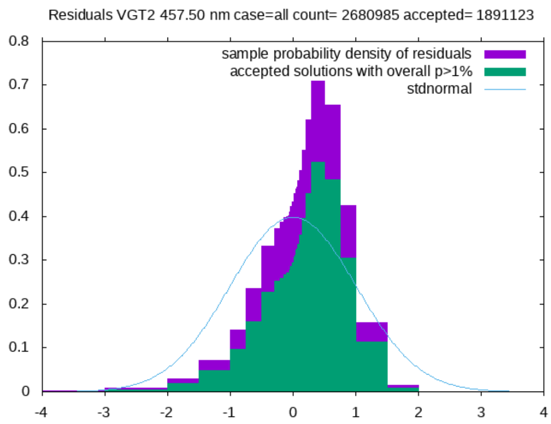
The correlation of the TOC reflectance uncertainties is unknown, but it is expected to be high, up to or above 0.5, and therefore relevant, especially for neighbouring bands at short wavelengths.

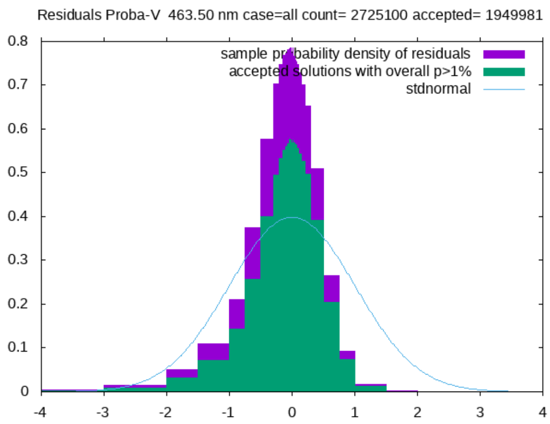
(b) TOC reflectance geolocation and spatial sampling

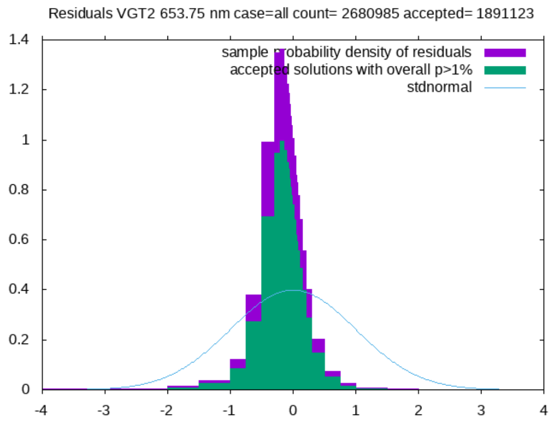
The effect of this source of uncertainty is unknown but expected to be sensor/platform dependent and spatially inhomogeneous, with a stronger impact over spatially inhomogeneous surfaces and higher latitudes.

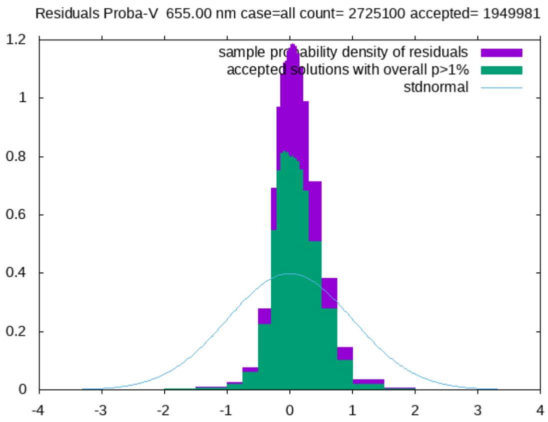
(c) Algorithmic uncertainty of the OptiSAIL simulation models (4SAILH, PROSPECT-D, Ross-Li BRDF, TARTES, soil model)

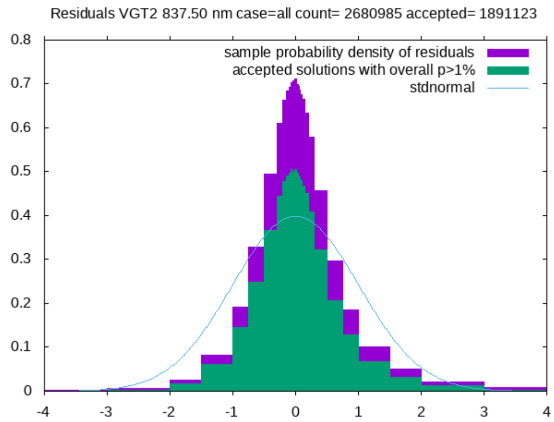
The effect of the algorithmic uncertainties in these models is complex. Algorithmic uncertainties owing to the model formulation as investigated by Berger et al. (2018) are taken into account by adding a variance corresponding to 6% of the observed reflectances to account for the model error of the simulation component of OptiSAIL ($\sigma_{data}^2 = \sigma_{r_{toc}}^2 + (0.06 r_{toc})^2$). The estimate of Berger et al. is based on an analysis of the accuracy by which measured reflectance spectral can be reproduced. They use a threshold value of 0.01 in mean absolute reflectance error in their analysis. This would be 6 % of a reflectance of 0.17 when expressed in relative terms. Investigation of cost function residuals justify the current choice of 6 % for the CRDP-1 production setup (Figure 3). The value may be adapted with increasing refinement of the uncertainty budget.

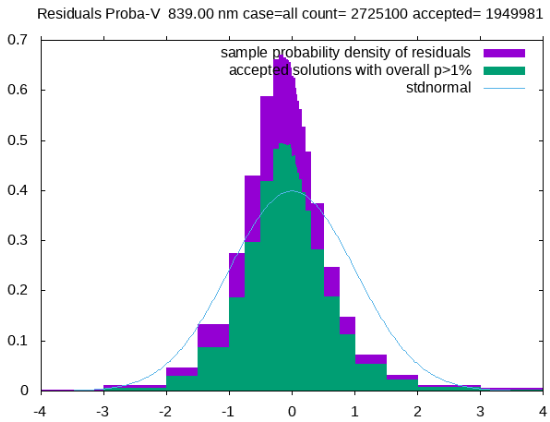

Residuals VGT2 457.50 nm case=all count= 2680985 accepted= 1891123
sample probability density of residuals
accepted solutions with overall p>1%
stdnormal


Residuals Proba-V 463.50 nm case=all count= 2725100 accepted= 1949981
sample probability density of residuals
accepted solutions with overall p>1%
stdnormal


Residuals VGT2 653.75 nm case=all count= 2680985 accepted= 1891123
sample probability density of residuals
accepted solutions with overall p>1%
stdnormal


Residuals Proba-V 655.00 nm case=all count= 2725100 accepted= 1949981
sample probability density of residuals
accepted solutions with overall p>1%
stdnormal


Residuals VGT2 837.50 nm case=all count= 2680985 accepted= 1891123
sample probability density of residuals
accepted solutions with overall p>1%
stdnormal


Residuals Proba-V 839.00 nm case=all count= 2725100 accepted= 1949981
sample probability density of residuals
accepted solutions with overall p>1%
stdnormal

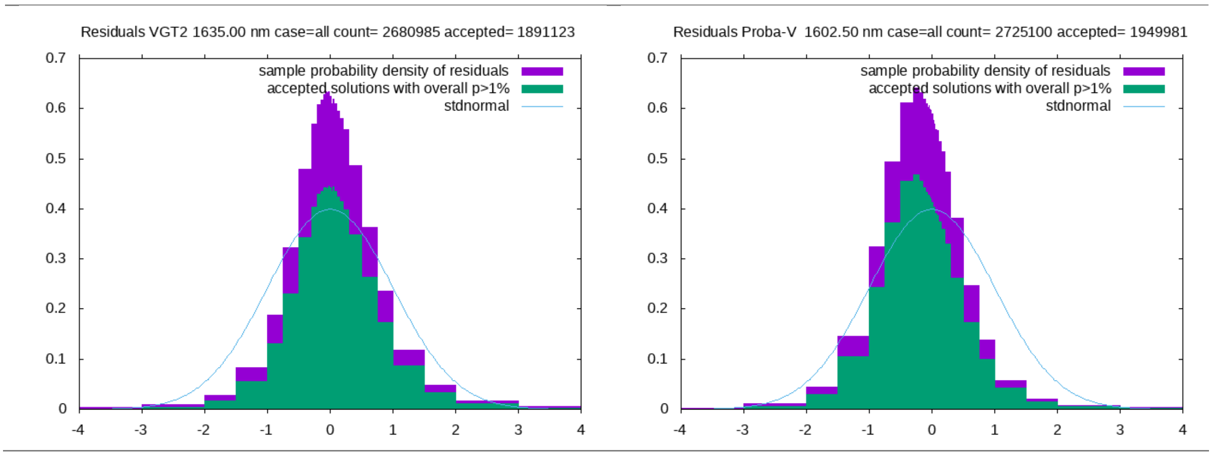


*Figure 3: **In purple**: Residuals per sensor and band, contributing to the cost function vector after retrieval for the central-European tile X18Y02 (in PROBA-V nomenclature) for 2014-05-16. Left column: VGT2, right column PROBA-V; bands as indicated in sub-captions. **In green**: subset of inversions accepted by a chi-square test. Light blue line: density of the Gaussian distribution which would be expected if all assumptions were ideally fulfilled.*

4.3 End-to-end uncertainty budget in OptiSAIL LAI

Uncertainties, as far as they are known and quantifiable, are propagated through the whole processing chain for every single retrieved grid cell and are part of the product. See the validation report for comparisons with other products and measurements.

Figure 4 shows the propagated uncertainty as presented in the product as scatter plot against the estimated LAI value for the same date and tile as in the introductory example.

Note that the reported uncertainty is the uncertainty for the retrieved model parameter of LAI. Uncertainties due to model representation errors are not included in this budget.

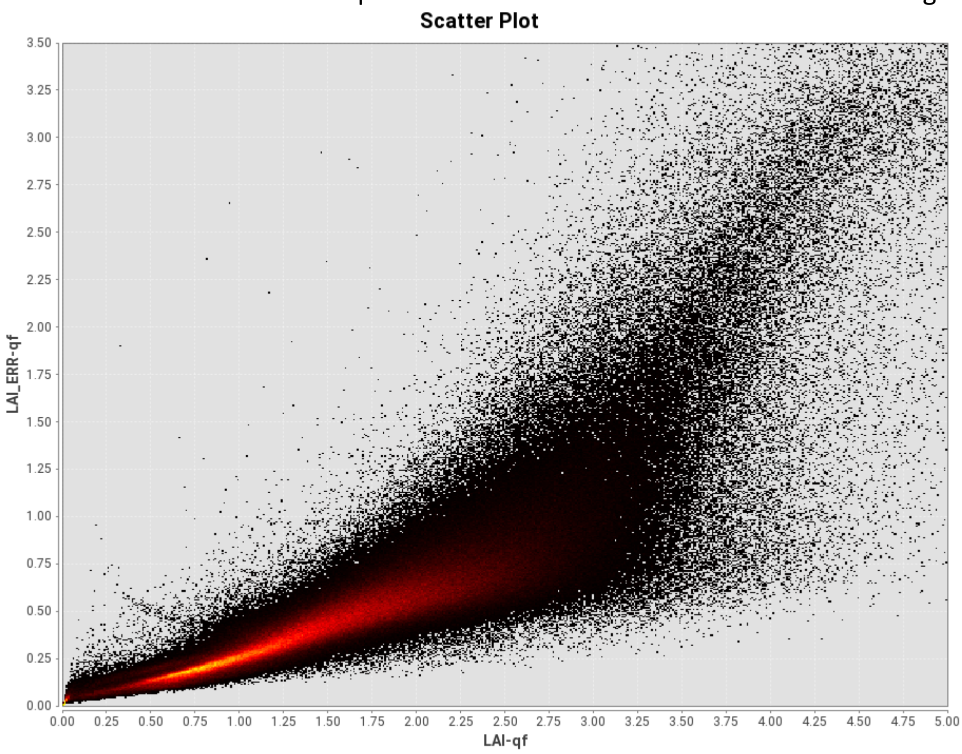


Figure 4: Scatter plot of OptiSAIL LAI uncertainty over the estimated quality-filtered LAI value from 2014-05-16 for the central-European tile X18Y02 (in PROBA-V nomenclature).

5 fAPAR

OptiSAIL retrieves LAI and fAPAR together. Therefore, the treatment of uncertainties is very similar except for the final step. fAPAR is computed from the retrieved parameters of the model by doing a hyper-spectral simulation. For its uncertainty, the full posterior covariance matrix of the parameters is propagated to fAPAR, using the Jacobian of the model (see ATBD for formulas). Similar considerations as for LAI apply (see OptiSAIL LAI section 4 above).
Figure 5 shows the propagated uncertainty as presented in the product as scatter plot against the estimated fAPAR value for the same date and tile as in the introductory example.

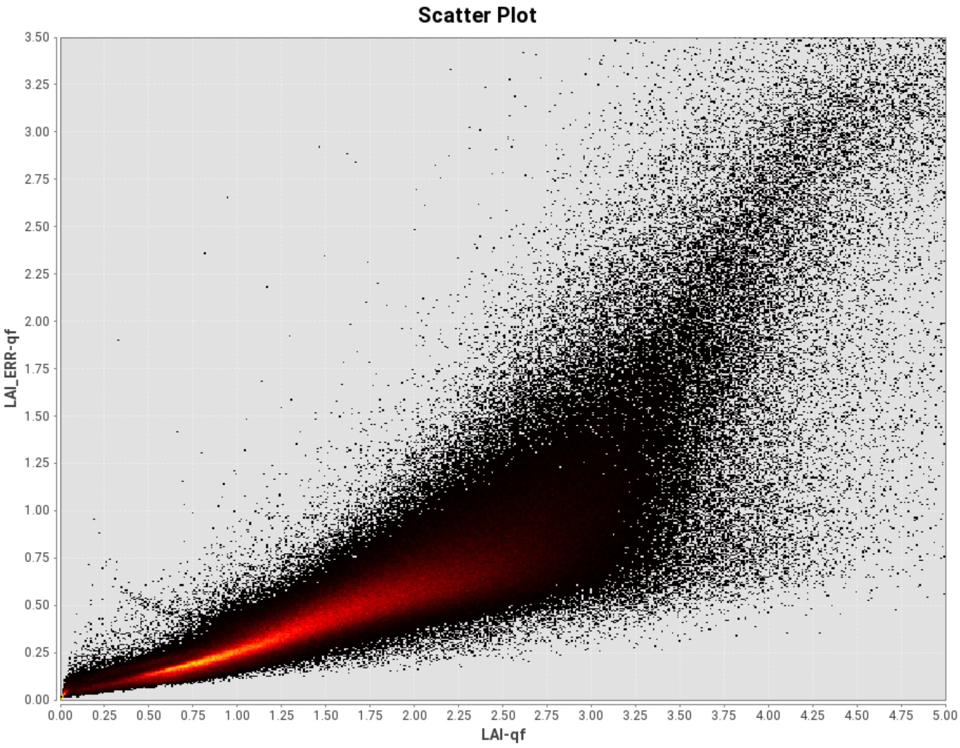


Figure 5: Scatter plot of OptiSAIL fAPAR uncertainty over the estimated quality-filtered fAPAR value from 2014-05-16 for the central-European tile X18Y02 (in PROBA-V nomenclature).

6 Chlorophyll-A+B (Cab)

Chlorophyll-A+B is one of the leaf pigments, whose effect on the leaf optical properties is simulated with PROSPECT-D inside the OptiSAIL retrieval system as mass per leaf area. It is retrieved together with the other model parameters, and its uncertainty budget is computed similar to the one of LAI. Cab uncertainty has a strong anticorrelation with LAI (see Figure 7). Not so much for low LAI, because the lower the LAI, the more the spectrum in the visible range is dominated by soil rather than the leaf absorption spectrum and Cab becomes less well determined. But for medium and higher LAI, a canopy with lower LAI and higher Cab can to some extent (at least in the visible part the spectrum) have a similar reflectance signature as a canopy with higher LAI and lower Cab. This can be exploited when computing the canopy chlorophyll content (CCC) as the product of the two (CCC=LAI*Cab), where the anticorrelation leads to a reduction of the uncertainty (compared to the uncorrelated case) in the error propagation. Note that for low LAI, Cab is not well determined and is retrieved as a value near the prior of 60 ug.cm-2 with an uncertainty near the Cab prior uncertainty of 25 ug.cm-2, as shown in Figure 6.

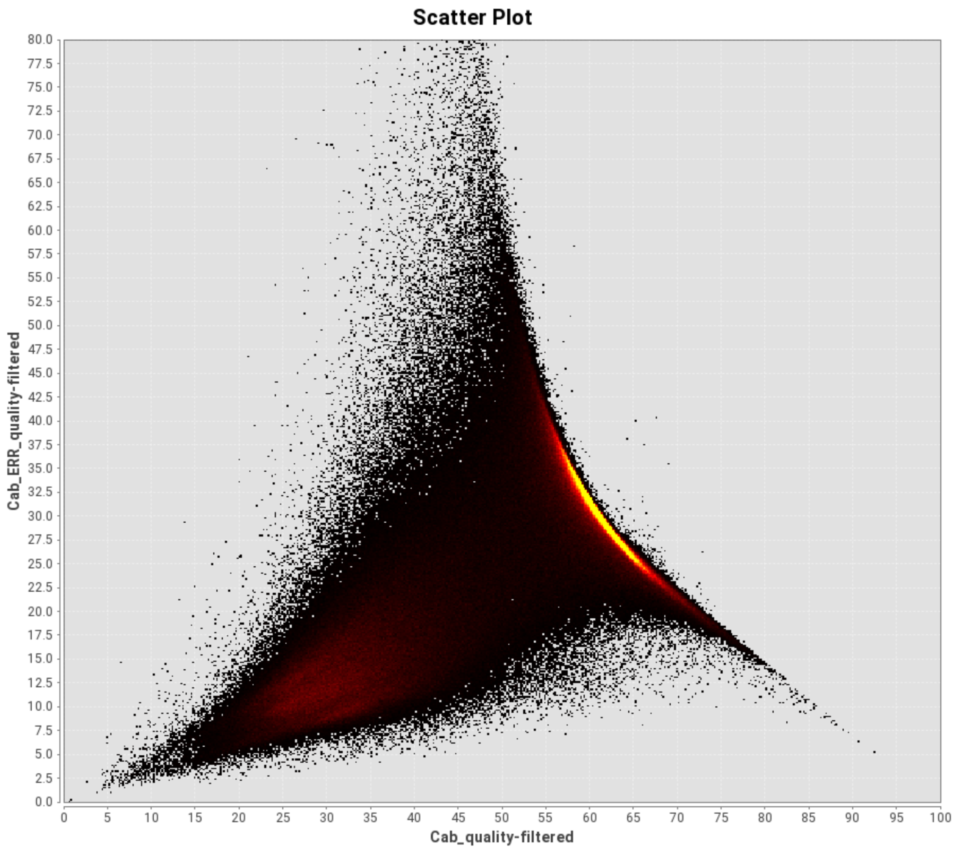


Figure 6: Scatter plot of OptiSAIL Cab uncertainty over the estimated quality-filtered Cab value from 2014-05-16 for the central-European tile X18Y02 (in PROBA-V nomenclature, units on both axes are ug.m-2).

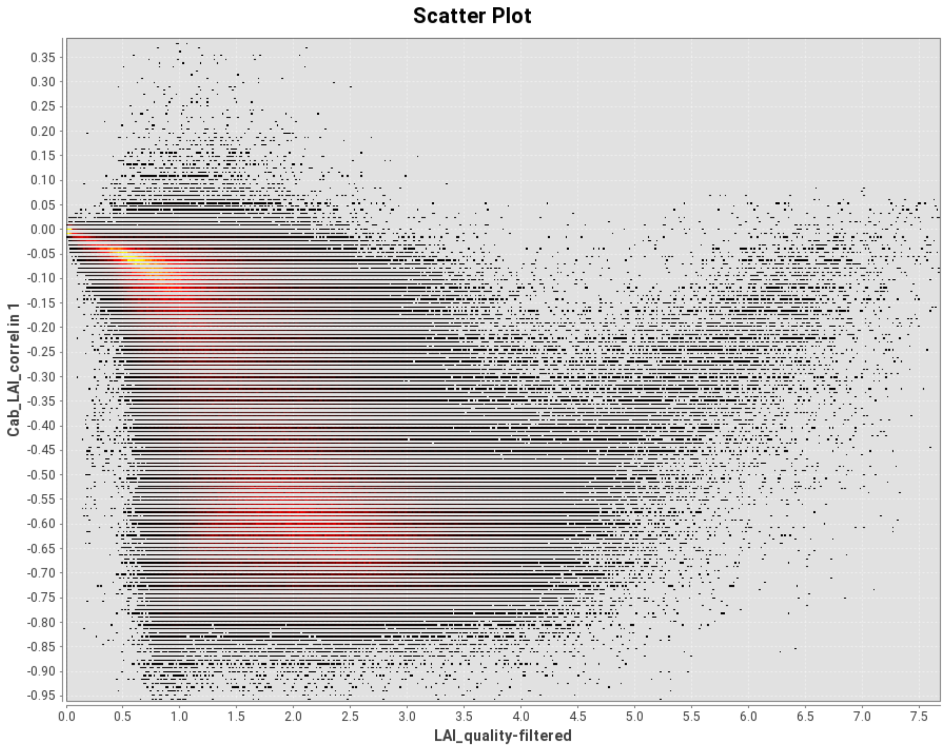


Figure 7: Scatter plot of Cab-LAI uncertainty correlation over the estimated quality-filtered LAI value from 2014-05-16 for the central-European tile X18Y02 (in PROBA-V nomenclature, both axes dimensionless). The horizontal striping is an artefact introduced by lossy data compression.

7 fAPAR_Cab

The uncertainty budget of fAPAR_Cab produced by the OptiSAIL algorithm, is technically very similar to that of OptiSAIL fAPAR, with the exception that the focus on the absorption by Chlorophyll A+B is expected to make this quantity more useful for the users interested in plant photosynthesis. This could be interpreted as a reduction of algorithmic uncertainty, as compared to total vegetation fAPAR.

8 Surface Albedo

OptiSAIL simulates surface albedo from the retrieved parameters, similar to the computation of fAPAR (cf. Section 5).

In OptiSAIL, surface albedo (bi-hemispherical and directional-hemispherical at local solar noon for the spectral broadbands VIS, NIR, SW) is computed from the retrieved parameters of the model by a hyper-spectral simulation. For its uncertainty, the full posterior covariance matrix of the parameters is propagated to surface albedo, using the Jacobian of the model (see ATBD for formulas). For algorithmic uncertainties, the same considerations as detailed for LAI and fAPAR below apply.

9 References

Baoxin Hu, Wolfgang Lucht, Xiaowen Li, and Alan H. Strahler: Validation of kernel-driven semiempirical models for the surface bidirectional reflectance distribution function of land surfaces. Remote Sensing of Environment, Volume 62, Issue 3, 1997, Pages 201-214, https://doi.org/10.1016/S0034-4257(97)00082-5.

Berger, K.; Atzberger, C.; Danner, M.; Wocher, M.; Mauser, W.; Hank, T. Model-Based Optimization of Spectral Sampling for the Retrieval of Crop Variables with the PROSAIL Model. Remote Sens. 2018, 10, 2063. https://doi.org/10.3390/rs10122063

Pinty, B., T. Lavergne, R. E. Dickinson, J.-L. Widlowski, N. Gobron, and M. M. Verstraete (2006), Simplifying the interaction of land surfaces with radiation for relating remote sensing products to climate models,J. Geophys. Res.,111, D02116, doi:10.1029/2005JD005952

Pinty, B., I. Andredakis, M. Clerici, T. Kaminski, M. Taberner, M. M. Verstraete, N. Gobron, S. Plummer, and J.-L.Widlowski (2011), Exploiting the MODIS albedos with the Two-stream Inversion Package (JRC-TIP): 1. Effective leaf area index, vegetation, and soil properties,J. Geophys. Res.,116, D09105, doi:10.1029/2010JD015372.