|  |  |
| --- | --- |
|  | **Validation Methodology for Copernicus S3 OGVI/OTCI FAPAR and CCC Data Products Over FRM-Compliant Field Sites (VM)**VERSION 1.0University of SouthamptonEOLAB NPL25 June 2020 |
| Related image | This document was produced as part of the ESA-funded project “Fiducial Reference Measurements for Vegetation Phase 2 (FRM4VEG 2)” under ESA contract number: 4000129823/20/I-NS. |

##### Signatures

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Name** | **Organisation** | **Date** |
| **Written by** | Luke Brown | University of Southampton | 25/06/2020 |
|  | Jadunandan Dash | University of Southampton |  |
|  | Fernando Camacho | EOLAB |  |
|  | Niall Origo | NPL |  |
|  | Rosalinda Morrone | NPL |  |
|  | Joanne Nightingale | NPL |  |
| **Reviewed by (consortium)** |  |  |  |
| **Approved by (ESA)** |  |  |  |

##### Version History

|  |  |  |
| --- | --- | --- |
| **Version** | **Date** | **Publicly available or private to consortium** |
| 1.0 | 25/06/2020 | Private Consortium |

##### Table of Contents

[Signatures 1](#_Toc43388035)

[Version History 1](#_Toc43388036)

[Table of Contents 2](#_Toc43388037)

[List of Figures 3](#_Toc43388038)

[List of Tables 3](#_Toc43388039)

[Acronyms 4](#_Toc43388040)

[1 Introduction 6](#_Toc43388041)

[1.1 Purpose and Scope 6](#_Toc43388042)

[2 Products and Terminology 7](#_Toc43388043)

[2.1 Terminology 7](#_Toc43388044)

[2.1.1 FAPAR 7](#_Toc43388045)

[2.1.2 CCC 7](#_Toc43388046)

[2.2 Considered Products 7](#_Toc43388047)

[2.2.1 OLCI Global Vegetation Index (OGVI) 8](#_Toc43388048)

[2.2.2 OLCI Terrestrial Chlorophyll Index (OTCI) 9](#_Toc43388049)

[2.2.3 Product Quality Flags 11](#_Toc43388050)

[3 State-of-the-art Methods for FAPAR and CCC Validation 13](#_Toc43388051)

[3.1 Review of Validation Approaches 13](#_Toc43388052)

[3.1.1 Direct Validation 13](#_Toc43388053)

[3.1.2 Direct Validation: Networks and Projects 13](#_Toc43388054)

[3.1.3 Indirect Validation 16](#_Toc43388055)

[3.1.4 Indirect Validation: Networks and Projects 16](#_Toc43388056)

[3.2 Review of Current Direct Validation Methodology 18](#_Toc43388057)

[3.2.1 Site Characteristics 18](#_Toc43388058)

[3.2.2 Site Sampling Strategy 19](#_Toc43388059)

[3.2.3 ESU Sampling Strategy 20](#_Toc43388060)

[3.2.4 Upscaling Method 21](#_Toc43388061)

[3.2.5 Convex Hulls 22](#_Toc43388062)

[3.2.6 Aggregation Method 23](#_Toc43388063)

[3.2.7 Validation Metrics 23](#_Toc43388064)

[4 Validation Status and Challenges 25](#_Toc43388065)

[4.1 Validation Status of Considered Products 25](#_Toc43388066)

[4.2 Gaps and Challenges in Current Validation Protocols 26](#_Toc43388067)

[5 Recommended Validation Strategy 27](#_Toc43388068)

[5.1 Sampling Strategy 27](#_Toc43388069)

[5.2 Quality Control 27](#_Toc43388070)

[5.3 Upscaling Approach 28](#_Toc43388071)

[5.4 High Spatial Resolution Imagery 28](#_Toc43388072)

[5.5 Incorporation of Uncertainties in Upscaling 29](#_Toc43388073)

[5.6 Derivation of Quality Flags 30](#_Toc43388074)

[A Appendix 35](#_Toc43388075)

##### List of Figures

[Figure 1: Diagram illustrating the ‘two stage’ or ‘bottom up’ approach to validation proposed by the CEOS WGCV LPV sub-group [34]. 18](#_Toc43388018)

[Figure 2: Distribution of ESUs over the site, as suggested by [10]. 20](#_Toc43388019)

[Figure 3: Within ESU sampling schemes, as suggested by [10]. 21](#_Toc43388020)

##### List of Tables

[Table 1: OLCI L2 land quality flags [21]. 11](#_Toc43388025)

[Table 2: OLCI L2 land science flags [21]. 12](#_Toc43388026)

[Table 3: Studies making use of the BELMANIP network of sites for indirect validation 17](#_Toc43388027)

[Table 4: Categories of validation metrics, after [51]. 24](#_Toc43388028)

##### Acronyms

|  |  |
| --- | --- |
| **Abbreviation** | **Stands For** |
| ANN | Artificial neural network |
| BELMANIP | Benchmark Land Multisite Analysis and Inter-comparison of Products |
| BRF | Bidirectional reflectance factor |
| CCC | Canopy chlorophyll content |
| CCRS | Canada Centre for Remote Sensing |
| CEOS | Committee on Earth Observation Satellites |
| CGLS | Copernicus Global Land Service |
| CNES | Centre National d'Études Spatiales |
| CYCLOPES | Carbon Cycle and Change in Land Observational Products from an Ensemble of Satellites |
| DHP | Digital hemispherical photography |
| ECN | Environmental Change Network |
| ECV | Essential climate variable |
| EOLAB | Earth Observation Laboratory |
| EPA | Environmental Protection Agency |
| ESA | European Space Agency |
| ESU | Elementary sampling unit |
| ETM+ | Enhanced Thematic Mapper |
| FAPAR | Fraction of absorbed photosynthetically active radiation |
| FCOVER | Fraction of vegetation cover |
| FOV | Field-of-view |
| FPP | FRM Protocols and Procedures |
| FRM | Fiducial Reference Measurements |
| FRM4VEG | Fiducial Reference Measurements for Vegetation |
| GBOV | Ground Based Observations for Validation |
| GCOS | Global Climate Observing System  |
| GLASS | Global Land Surface Satellite |
| GLOBCARBON | Global Biophysical Products Terrestrial Carbon Studies |
| GPS | Global positioning system |
| HRV | High Resolution Visible |
| IGBP | International Geosphere Biosphere Programme |
| ImagineS | Implementation of Multiscale Agricultural Indicators Exploring Sentinels |
| INRA | Institut National de la Recherche Agronomique |
| IRLS | Iteratively reweighted least squares |
| JRC | Joint Research Centre |
| LAI | Leaf area index |
| LCC | Leaf chlorophyll concentration |
| LPV | Land Product Validation |
| LUT | Look-up-table |
| MERIS | Medium Resolution Imaging Spectrometer |
| MGVI | MERIS Global Vegetation Index |
| MODIS | Moderate Resolution Imaging Spectroradiometer |
| MODLAND | MODIS Land Science Team |
| MSI | Multispectral Instrument |
| MTCI | MERIS Terrestrial Chlorophyll Index |
| NEON | National Ecology Observatory Network |
| NPL | National Physical Laboratory  |
| NPP | Net primary productivity |
| ODR | Orthogonal distance regression |
| OGVI | OLCI Global Vegetation Index |
| OLCI | Ocean and Land Colour Instrument |
| OLIVE | Online Interactive Validation Exercise |
| OLS | Ordinary least squares |
| OTCI | OLCI Terrestrial Chlorophyll Index |
| PAR | Photosynthetically active radiation |
| PDGS | Payload data ground segment |
| PROVE | Prototype Validation Experiment |
| PSF | Point spread function |
| RTM | Radiative transfer model |
| RUT | Radiometric Uncertainty Tool |
| S3MPC | Sentinel-3 Mission Performance Centre |
| SAFARI-2000 | Southern Southern African Regional Science Initiative |
| SPOT | Satellite Pour l’Observation de la Terre |
| TM | Thematic Mapper |
| TOA | Top-of-atmosphere |
| VALERI | Validation of European Land Remote Sensing Instruments |
| WGCV | Working Group on Calibration and Validation |

# Introduction

## Purpose and Scope

This document forms part of deliverable D-40 of the European Space Agency (ESA) project ‘Fiducial Reference Measurements for Vegetation (FRM4VEG Phase 2)’. The purpose of the document is to provide an overview of a methodology for validating Copernicus Sentinel-3 OGVI/OTCI FAPAR and CCC data products over vegetated FRM-compliant field sites. The document is organized into 4 key sections:

* **Section 1** provides a summary of the VM-FC
* **Section 2** provides an overview of the considered FAPAR and CCC data products and relevant terminology.
* **Section 3** provides a review of the state-of-the-art methods for validation of satellite derived FAPAR and CCC data products.
* **Section 4** summarises the FRM-recommended validation methodology for validating satellite derived FAPAR and CCC data products using FRM-compliant field site data. This section will discuss the requirements and challenges for validating moderate resolution data products as well as demonstrate the handling of various error sources and their relative impact on the validation methods and results.
* **Appendix 1** provides a practical example of FRM-based validation methodology for validating S3 OGVI/OTCI FAPAR and CCC data products based on FRM4VEG Phase 1 field campaign data.

# Products and Terminology

## Terminology

### FAPAR

The fraction of absorbed photosynthetically active radiation (FAPAR) is recognised as an essential climate variable (ECV) by the Global Climate Observing System (GCOS) as it has a primary role in estimation of the carbon balance. FAPAR is generally defined as the fraction of photosynthetically active radiation (PAR) absorbed by vegetation, where PAR is the solar radiation reaching the vegetation in the wavelength region 400 nm to 700 nm. It is a dimensionless quantity varying from zero (over bare soil) to almost one for the largest amounts of green vegetation. Since FAPAR is mainly used as a descriptor of photosynthesis and evapotranspiration processes, only the green photosynthetic elements (leaves, needles, or other green elements) should be accounted for.

FAPAR depends on the illumination conditions, i.e. the angular position of the sun and the relative contributions of the direct and diffuse illumination. Both black-sky (assuming only direct radiation) and white-sky (assuming that all incoming radiation is in the form of isotropic diffuse radiation) FAPAR values may be considered. FAPAR may also be considered on an instantaneous (at the time of acquisition) or temporally integrated basis.

### CCC

As the key photosynthetic pigment within a plant, chlorophyll plays an important role in determining its physiological status. The content of chlorophyll within a vegetation canopy is therefore strongly related to its productivity and is a sensitive indicator of its health. Thus, estimates of canopy chlorophyll content (CCC) are a key input into models of terrestrial primary productivity and carbon exchange. CCC is defined as the product of leaf chlorophyll concentration (LCC) and leaf area index (LAI). LCC is expressed as the mass of chlorophyll per unit leaf area, whist LAI is a dimensionless quantity defined as the one-sided leaf area per unit ground area.

## Considered Products

Two L2 land products derived from Sentinel-3’s Ocean and Land Colour Instrument (OLCI) are considered: the OLCI Global Vegetation Index (OGVI), which corresponds to instantaneous green FAPAR, and the OLCI Terrestrial Chlorophyll Index (OTCI), a surrogate of canopy chlorophyll content (CCC). Both products are provided at OLCI’s native spatial resolution of 300 m, in addition to a reduced spatial resolution of 1 km, and benefit from the instrument’s three-day repeat cycle. The OGVI is designed to provide continuity to the Medium Resolution Imaging Spectrometer (MERIS) Global Vegetation Index (MGVI), whilst the OTCI is designed to provide continuity to the MERIS Terrestrial Chlorophyll Index (MTCI). The theoretical basis and validation status of each product is described in the following sections.

### OLCI Global Vegetation Index (OGVI)

Based on the MGVI, the OGVI exploits the fact that live green vegetation strongly absorbs solar radiation in the red region of the electromagnetic spectrum and strongly scatters it in the near-infrared region [1]. The product is designed to maximise sensitivity to instantaneous green FAPAR, whilst minimising sensitivity to perturbing factors such as atmospheric contamination and the soil background. The underlying algorithm, developed by the Joint Research Centre (JRC) and known as JRC-FAPAR, consists of two main steps [2]. In the first, ‘rectified’ reflectance values are computed from OLCI bands 17 (865 nm) and 10 (681.25 nm). The rectification procedure uses information from OLCI band 3 (442.5 nm) to supress atmospheric effects, whilst directional normalisation is carried out using the model of [3]. In the second step, the ‘rectified’ reflectance values are used to derive FAPAR [4].

Both steps are achieved with ratios of polynomials, making use of sensor-specific coefficients. Thus the OGVI is calculated from the rectified reflectance values as

where

and where are coefficients, whilst the rectified reflectance values are estimated as

where

and where

where is the simulated top-of-atmosphere (TOA) bidirectional reflectance factor (BRF) in band , whilst is the BRF normalised by the anisotropic function derived from the model of [3], is the solar zenith angle, is the observer zenith angle, is the relative azimuth angle, and , , , are coefficients [4]. The coefficients themselves are determined using a look-up-table (LUT) of one-dimensional radiative transfer model (RTM) simulations corresponding to a range of environmental and observational conditions. The semi-discrete vegetation model of [5], soil spectral library of [6], and atmospheric model of [7] are used for this purpose. The coefficients are optimised to minimise the difference between the OGVI and FAPAR, such that the OGVI is constrained to take on values as close as possible to the FAPAR associated with the simulations [4].

The OGVI is provided with per-pixel uncertainties, which are derived by propagating uncertainties in the input quantities through its calculation, such that

where are the TOA reflectance values in OLCI bands 3, 10, and 17, centred at 442.5 nm, 681.25 nm and 865 nm, respectively [4].

### OLCI Terrestrial Chlorophyll Index (OTCI)

The interaction of incoming radiation with vegetated surfaced results in a distinctive reflectance spectrum. As the key photosynthetic pigment in plants, chlorophyll absorbs much of the incoming radiation in the visible region of the electromagnetic spectrum, leading to low reflectance. In the near-infrared, optical properties are mostly controlled by leaf structure, and reflectance in this region of the electromagnetic spectrum is high due to internal scattering within the leaf. The boundary between strong absorption due to the presence of chlorophyll and strong reflectance due to leaf structure is known as the red-edge, and its position (i.e. the location of the maximum rate of change) is strongly related to CCC [8]. As CCC increases, the red-edge moves towards longer wavelengths.

Where continuous spectra are available, the red-edge position can be determined as the maximum of the first derivative [8], [9] whilst a range of techniques have also been applied to estimate the position of the red-edge from discontinuous spectra. These include higher order curve fitting [10], an inverted Gaussian model [11], [12], linear interpolation [13], [14], and Lagrangian interpolation [15][16]. Drawbacks associated with these techniques include saturation at high CCC and difficulties in automating some procedures (e.g. the need for manual confirmation of the maximum of the first derivative when two peaks occur in the case of Lagrangian interpolation) [17]. Taking advantage of the MERIS red-edge bands, the MTCI was designed to provide a computationally efficient surrogate of CCC that remains sensitive to high CCC values [17]. Based on the MTCI, the OTCI is calculated from equivalent OLCI bands in and around the red-edge, and takes the form

where , , and are reflectance values in the OLCI bands centred at 753.75 nm, 708.75 nm, and 681.25 nm, respectively, after correction for gaseous absorption and Rayleigh scattering [18]. OTCI values are only provided for identified by the scene classification as land, and for which the following conditions are satisfied:

* (

Values are also restricted to the range of 0 to 6.5 [18].

Like the OGVI, the OTCI is provided with per-pixel uncertainties, which are derived by propagating uncertainties in the input quantities through its calculation, following the approach of [19] and as described in [20]. Thus, the standard uncertainty in the OTCI is determined as

where , , and are standard uncertainties in Rayleigh corrected reflectance values in OLCI bands 12, 11, and 10, respectively [18].

### Product Quality Flags

Both the OGVI and OTCI are provided with a number of quality flags to provide information to the user about issues with the input data and the conditions under which the observation was made (Table 1). These include a scene classification to identify cloud, land, snow/ice, and water pixels. In addition to the generic flags shared by the products, science flags are also provided, which provide information specific to each of the products (Table 1).

Table 1: OLCI L2 land quality flags [21].

|  |  |
| --- | --- |
| **Name** | **Description** |
| INVALID | Invalid flag: instrument data missing or invalid |
| WATER | Clear sky water |
| LAND | Clear sky land |
| CLOUD | Cloudy pixel |
| CLOUD\_AMBIGUOUS | Potentially cloudy pixels |
| CLOUD\_MARGIN | A margin around CLOUD and CLOUD\_AMBIGUOUS of 2 pixels in RR and 4 pixels in FR products |
| SNOW\_ICE | Possible sea-ice or snow contamination |
| INLAND\_WATER | Fresh inland waters flag: based on Level-1 land\_water flag |
| TIDAL | Pixel is in shallow water based on Level-1 land\_water flag |
| COSMETIC | Cosmetic flag (from Level-1B): missing data filled in by interpolation |
| SUSPECT | Suspect flag (from Level-1B): transmission errors means measurements may be unreliable |
| HISOLZEN | High solar zenith: 70 ° |
| SATURATED | Saturation flag: saturated within any band from 400 to 754 nm or in bands 779, 865, 885 and 1020 nm |
| WVFAIL | Suspect values derived for the water vapour over land: see ATBD SD-03-C02 for details. Set when the following internal flags are raised: ORINPWV\_F or OROUTWV\_F or L\_WV\_FAIL |
| OGVI\_FAIL | Suspect values derived for the OGVI (FAPAR): see ATBD SD-03-C13 for details - set when the following internal flags are raised: ORINP1\_F or OROUT1\_F |
| OTCI\_FAIL | Suspect values derived for the OTCI: see ATBD for SD-03-C14 details - set when the following internal flags are raised: ORINP2\_F (OTCI input out of range) or OROUT2\_F (OTCI output out of range) or LRAYFAIL\_F (problems deriving Rayleigh reflectance) |

Table 2: OLCI L2 land science flags [21].

|  |  |
| --- | --- |
| **Name** | **Description** |
| LRAYFAIL | Problems deriving the Rayleigh reflectance over the land: see ATBD SD-03-C15 for details |
| OGVI\_CLASS\_BAD | Flag bad data from OGVI spectral tests |
| OGVI\_CLASS\_WS | Flag water or deep shadow from OGVI spectral tests |
| OGVI\_CLASS\_CSI | Flag Cloud, snow or ice from OGVI spectral tests |
| OGVI\_CLASS\_BRIGHT | Flag bright from OGVI spectral tests |
| OGVI\_CLASS\_INVAL\_REC | Flag invalid rectification |
| OTCI\_BAD\_IN | Input data bad quality: (at least one of B12, B11, B10 is not Valid) or (B12-B11)<Threshold1 or (B11-B10)<Threshold2 |
| OTCI\_CLASS\_ANG | View angle flag: OTCI\_CLASS\_IN OK AND view angle> Threshold (TBD) AND sun angle > Threshold (TBD) |
| OTCI\_CLASS\_CLSN | Cloud and snow flag: input data quality flag OK but cloud shadow or partial snow |

# State-of-the-art Methods for FAPAR and CCC Validation

## Review of Validation Approaches

### Direct Validation

Direct validation refers to a set of approaches involving the comparison of satellite-derived products with independent in situ reference measurements. Because destructive sampling is laborious, time-consuming, and impractical at most sites, the use of indirect techniques is typically favoured for performing in situ measurements. These techniques make use of a range of optical instruments, and include:

* Ceptometry
* Digital hemispherical photography (DHP)
* Use of optical chlorophyll meters
* Use of the LI-COR LAI-2200C Plant Canopy Analyser (and the previous LAI-2000 and LAI-2200 variants)

The theory behind these techniques is described in the ‘FRM Protocols and Procedures for Surface Reflectance Fraction of Absorbed Photosynthetically Active Radiation, and Canopy Chlorophyll Content (FPP) document’. A major challenge in the direct validation of satellite-derived vegetation products is the heterogeneity of the terrestrial landscape. Because in situ measurements are point-based, straightforward comparison is possible only in highly homogeneous environments, as the number of samples that can be obtained to adequately represent a site is necessarily limited by logistical constraints.

### Direct Validation: Networks and Projects

A summary of some of the major direct validation projects carried out over the last two decades is provided below. The list is not exhaustive and seeks only to describe those projects involving campaigns over a network of multiple sites.

#### BigFoot

The BigFoot project was developed to support validation of the Moderate Spatial Resolution Imaging Spectroradiometer (MODIS) Land Science Team (MODLAND) products, following the success of early campaigns such as the Jornada Prototype Validation Experiment (PROVE) [22]. The products considered within the BigFoot project included land cover, LAI, FAPAR, and net primary production (NPP). The project involved the collection of in situ measurements at four 5 km x 5 km sites in the United States, covering deciduous broadleaf forest, cropland, grassland and evergreen needleleaf forest. Each site was also equipped with an eddy covariance flux tower. The sampling design was explicitly designed to capture spatial variation within plots of 25 m x 25 m, enabling in situ measurements to be upscaled to the flux tower footprint and MODIS spatial resolution, making use of Landsat Enhanced Thematic Mapper (ETM+) data [23].

#### Southern African Regional Science Initiative (SAFARI-2000)

The Southern African Regional Science Initiative (SAFARI-2000) was developed to investigate the linkages between land and atmospheric processes in the Southern African region. Within the project, several field sites were established along the International Geosphere Biosphere Programme (IGBP) Kalahri Transect, covering equatorial/subtropical forest and arid scrubland [24]. In a series of campaigns, in situ LAI measurements were performed over a 1 km x 1 km area (enabling comparison to MODIS). Measurements were conducted every 25 m along three transects of 750 m, whilst intensive sampling occurred every 50 m within a nested 250 m x 300 m grid [25].

#### Validation of Land European Remote Sensing Instruments (VALERI)

The Validation of Land European Remote Sensing Instruments (VALERI) project was developed by the Centre National d'Études Spatiales (CNES) and the Institut National de Recherche Agronomique (INRA). It was designed to provide high spatial resolution reference maps of LAI, FAPAR, and the fraction of vegetation cover (FCOVER) for the purpose of satellite-derived vegetation product validation, making use of in situ measurements. The project focused on the development of an effective methodological framework, and offered a pool of instrumentation for performing in situ measurements [10]. By leveraging a large network of international partners, a total of 53 campaigns were carried out at 33 sites within the VALERI project, covering 21 counties across Africa, Asia, Europe, North America, Oceania, and South America. The in situ measurements collected in each campaign were processed by INRA to derive high spatial resolution reference maps covering a 3 km x 3 km area.

#### Implementation of Multiscale Agricultural Indicators Exploiting Sentinels (ImagineS)

The Implementation of Multiscale Agricultural Indicators Exploring Sentinels (ImagineS) project was designed to support the Copernicus Global Land Service (CGLS), and focussed on developing the processing chains required to generate global LAI, FAPAR and FCOVER products from PROBA-V at a spatial resolution of 300 m. To assess the performance of these processing chains, in situ measurements were performed in 46 campaigns covering 20 sites between 2013 and 2016, in collaboration with a range of international partners. Local teams followed data collection protocols designed to comply with the recommendations of the Committee on Earth Observation Satellites (CEOS) Working Group on Calibration and Validation (WGCV) Land Product Validation (LPV) sub-group, and as in the VALERI project, these in situ measurements were processed by the Earth Observation Laboratory (EOLAB) to derive high spatial resolution reference maps covering a 3 km x 3 km area. Upscaling was achieved using Landsat 8 Operational Land Imager (OLI) data.

#### DIRECT 2.0

The CEOS WGCV LPV DIRECT database is one of the largest collections of in situ measurements for validating satellite-derived vegetation products. The database represents a collection of in situ data obtained in the previously described projects, in addition to efforts by Boston University, the University of Alberta, the Canadian Centre for Remote Sensing (CCRS), EOLAB, the Environmental Protection Agency (EPA), and ESA [26], [27]. Now in its second version (DIRECT 2.0), it contains data from 242 campaigns over 140 sites, spanning from 2000 to 2017.

#### Ground Based Observations for Validation of Copernicus Land Products (GBOV)

The most recent direct validation initiative is the Ground Based Observations for Validation of Copernicus Land Products (GBOV) project. The project was recently initiated by the European Commission’s Joint Research Centre (JRC), and aims to develop and distribute robust in situ datasets for the purpose of direct validation. Within the project, raw observations from a range of existing networks are collected and processed to provide datasets suitable for validating global albedo, land surface temperature, soil moisture, surface reflectance, and vegetation products. In the case of the vegetation variables, in situ measurements are currently derived from raw DHP collected by the National Ecological Observatory Network (NEON) in the United States. These in situ measurements are then upscaled using Landsat 8 OLI and Sentinel-2 Multispectral Instrument (MSI) data. In the first year of the project, 744 3 km x 3 km high spatial resolution reference maps suitable for validating LAI, FAPAR and FCOVER products have been produced, covering 20 sites between 2013 and 2017.

### Indirect Validation

Indirect validation refers to the inter-comparison of satellite-derived products generated using different instruments and/or retrieval algorithms. These activities enable agreement between the different products to be assessed, providing information on their relative performance. Such information is typically provided by means of:

* Difference maps
* Frequency and cumulative distribution plots
* Transect and temporal profiles
* Product to product scatter plots

A key advantage of indirect validation is the diverse global sampling that can be achieved with relative ease. When compared to direct validation, indirect validation enables a greater number of sites, vegetation conditions, and time periods to be covered. Nevertheless, its utility depends upon the extent to which the traceability of the data can be described (i.e. how data are processed from L1, in addition to the retrieval algorithms and assumptions adopted by each product). The definitions of the products in question must also be carefully considered, as they may represent similar, but ultimately different quantities.

### Indirect Validation: Networks and Projects

#### Benchmark Land Multisite Analysis and Inter-comparison of Products (BELMANIP)

The importance of representing global conditions in indirect validation efforts was recognised by the CEOS WGCV LPV sub-group, who, in 2006, established a network of sites for product inter-comparison under the Benchmark Land Multisite Analysis and Inter-comparison of Products (BELMANIP) initiative. Sites were selected from existing networks and projects, ensuring that for each latitudinal band of the globe, a representative proportion of biomes were sampled. Additional constraints were imposed to ensure that sites were flat, homogeneous over a 10 km x 10 km area, and minimally covered by water bodies and urban settlements [28]. An updated list (BELMANIP2) was released in 2009, bringing the total number of sites to 445.

Several indirect validation exercises have been carried out over the BELMANIP sites, facilitating inter-comparison between a wide range of satellite-derived vegetation products, including:

* Carbon Cycle and Change in Land Observational Products from an Ensemble of Satellites (CYCLOPES)
* ECOCLIMAP
* Global Biophysical Products Terrestrial Carbon Studies (GLOBCARBON)
* Global Land Surface Satellite (GLASS)
* MODIS
* MGVI
* GEOV1

These studies have demonstrated, in some cases, substantial differences between the products. For example, [26] found that whilst the CYCLOPES and MODIS products reflect realistic spatial and temporal variations, the ECOCLIMAP and GLOBCARBON products demonstrate comparatively poor consistency. The best agreement was demonstrated over cropland sites, whilst the worst agreement was demonstrated over forests due to differences in each algorithm’s representation of canopy structure. When compared to the MODIS product, [29] found that the CYCLOPES product is characterised by smoother seasonality, but suffers from underestimation of higher values over forest environments. However, at lower values, the MODIS product tends to overestimate, particularly in the case of FAPAR. Similar results were presented by [30] and [6], whilst [27] demonstrated that by using a retrieval algorithm trained on fused CYCLOPES and MODIS data, the GEOV1 product is able to overcome some of these issues. In terms of the MGVI, [27] observed underestimation when compared to the GEOV1 product, and similar results were presented by [30] when comparing the MGVI to both the GEOV1 and MODIS products. Of the products considered by [31], the GLASS and MODIS products, which share input data from the same instrument, demonstrated the best agreement.

Table : Studies making use of the BELMANIP network of sites for indirect validation

|  |  |  |
| --- | --- | --- |
| **Study** | **Considered products** | **Considered variables** |
| [32] | CYCLOPES, ECOCLIMAP, GLOBCARBON, MODIS C4 | LAI |
| [33] | CYCLOPES, ECOCLIMAP, MODIS C4 | LAI, FAPAR |
| [27] | CYCLOPES, GEOV1, GLOBCARBON, MGVI, MODIS C5 | LAI, FAPAR, FCOVER |
| [5] | GEOV1, MGVI, MODIS | FAPAR |
| [6] | CYCLOPES, GEOV1, GLASS, MODIS C6 | LAI, FAPAR |

#### Online Interactive Validation Exercise (OLIVE)

In parallel to establishing a network of sites for indirect validation activities, efforts to make validation datasets open and accessible to the community have also been made. The Online Interactive Validation Exercise (OLIVE) tool developed by the CEOS WGCV LPV sub-group is one example [30]. OLIVE is an online platform designed for the validation of global leaf area index (LAI), fraction of absorbed photosynthetically active radiation (FAPAR), and fraction of vegetation cover (FCOVER) products. Containing data from five operational products over the BELMANIP2 sites, it enables users to generate online reports, facilitating:

* Evaluation of product continuity, temporal consistency and stability
* Generation of temporal profiles for the time period of interest
* Evaluation of product statistical distribution at the biome, continental and global scale
* Comparison with similar products (scatterplots and statistics for pairs of products)

## Review of Current Direct Validation Methodology

To address the disparity in spatial scale between in situ measurements and the satellite-derived vegetation products themselves, the ‘two stage’ or ‘bottom up’ approach was proposed by the CEOS WGCV LPV sub-group. The approach relies on the use of high spatial resolution imagery to upscale in situ measurements to the moderate spatial resolution of the satellite-derived product of interest. It consists of performing in situ measurements within elementary sampling units (ESUs) that approximate the extent of a pixel of high spatial resolution imagery. Fine scale variability is characterised by repeat measurements within the ESU, whereas the high spatial resolution imagery is used to characterise broad-scale variability over the site. By combining the two sources of data, a high spatial resolution reference map can be derived, which can then be aggregated to the required moderate spatial resolution, enabling validation (Figure 1).



Figure : Diagram illustrating the ‘two stage’ or ‘bottom up’ approach to validation proposed by the CEOS WGCV LPV sub-group [34].

### Site Characteristics

To successfully apply the ‘two stage’ or ‘bottom up’ approach, [10] suggest that validation sites should be relatively homogeneous, covering an area of at least 3 km x 3 km. This suggestion is reiterated by [34], who suggest that whilst a site of 1 km x 1 km may be suitable in homogeneous areas, such an extent is likely too small given positional uncertainties and the instrument’s point spread function (PSF). It should be noted that the 3 km x 3 km extent was recommended at a time when most considered products had a spatial resolution of approximately 1 km. Although modern instruments such as Sentinel-3’s OLCI have an improved spatial resolution (e.g. 300 m), the 3 km x 3 km criterion is still considered a good rule of thumb to ensure backwards compatibility of validation data with those older generation instruments still in operation. In addition to its size and homogeneity, [10] also recommend sites that are relatively flat to avoid terrain effects. In terms of the distribution of sites, it is suggested that they should sample the variability of biomes and vegetation conditions encountered over the globe [10].

### Site Sampling Strategy

The number and distribution of ESUs required over a site are a function of its heterogeneity. According to [35], the number of ESUs required can be determined as

where is equal to 2 in the case of the 95% confidence level, is the expected accuracy, is equal to , and is the allowable error. Thus, with an expected accuracy of 85% and an allowable error of 5%, the number of samples required to reach the 95% confidence level would be 204. In reality this number is unlikely to be attainable in a single field campaign due to logistical constraints. As such, the typical minimum number of ESUs recommended ranges from twenty to thirty [10], [34].

To achieve a good spatial distribution, [10] recommend that the 3 km x 3 km site should be split into 1 km x 1 km cells, and that three to five ESUs should be established in each (Figure 2). Both [34] and [10] recommend that ESUs should be selected based on the land cover of the site to achieve proportional representation of all dominant land cover types, and [36] suggest that at least three ESUs should be established per land cover type (including bare areas), covering, low, medium and high values of the parameter of interest. Additional criteria to be considered in the placement of ESUs are related to logistical constraints, including accessibility. They should be located close enough together to maintain efficient in situ sampling, but far enough apart to minimise spatial autocorrelation [34]. Similarly, they should be located at a reasonable distance from borders to avoid adjacency effects [10], [36].

### ESU Sampling Strategy

ESU extent is primarily driven by the high spatial resolution imagery used for upscaling. For example, [10] refer to ESUs of 20 m x 20 m, corresponding to the spatial resolution of Satellite Pour l’Observation de la Terre (SPOT) High Resolution Visible (HRV) data. Similarly, [34] refer to ESUs of approximately 30 m x 30 m, reflecting the spatial resolution of Landsat Thematic Mapper (TM) and ETM+ data, but note that the actual extent will depend on the field-of-view (FOV) of the in situ measurements, and may reflect a small cluster of pixels in reality. When planning a campaign, the ESU extent may be selected to account for positional uncertainties in the high spatial resolution imagery. When the positional uncertainty is known, the minimum required extent can be calculated according to [37] as:

where is the required extent, is the spatial resolution, and is the positional uncertainty (in pixels). For example, if Sentinel-2 MSI data with a spatial resolution of 10 m and a positional uncertainty of approximately 0.5 pixels are to be used in upscaling, the minimum extent of the ESUs should be 20 m x 20 m.



Figure : Distribution of ESUs over the site, as suggested by [10].

Depending on the type and density of vegetation, a number of potential sampling schemes may be adopted within an ESU, including random, systematic, and transect based approaches. Both the square and cross sampling schemes (Figure 3) proposed by [10] were shown to provide similar performances. Transect based approaches are most appropriate in sparse canopies [10]. In the case of row structured crops, small transects should be arranged between rows to characterise the row effect [36], [38]. The optimal number of individual in situ measurements is driven by the heterogeneity of the ESU and the measurement footprint of the instrument. Work by [39] and [40] suggests that between 10 to 15 are adequate in the case of DHP and the LAI-2200C. This is supported by [10] and [36], who recommend at least 12 samples, and reiterated by [41]. In terms of optical chlorophyll meter measurements, existing validation protocols typically recommend selecting several leaves per plant, enabling variations at different levels of the canopy to be accounted for [42], [43]. They also recommend a minimum of four to seven measurements per leaf, in order to account for variations across the leaf surface [42], [44], [45].



Figure : Within ESU sampling schemes, as suggested by [10].

### Upscaling Method

Once in situ measurements and near-coincident high spatial resolution imagery are acquired, several upscaling approaches are available to derive a high spatial resolution reference map. The most widely used involves establishing an empirical transfer function relating the in situ measurements to the high spatial resolution imagery. This may be achieved, for example, by multiple regression incorporating radiance/reflectance values in several spectral bands. Alternatively, the transfer function may be based on a vegetation index [10], [34], [41]. In addition to ordinary least squares (OLS), several studies make use of robust regression techniques to establish a transfer function, including the Theil-Sen estimator [41], [46] and iteratively reweighted least squares (IRLS) [47]. These techniques are advantageous as they are less sensitive to outliers. Because of the empirical nature of the transfer function, atmospheric correction is not considered mandatory if atmospheric characteristics can be considered constant over the site [34]. Other approaches include RTM inversion, machine learning/hybrid techniques, and geostatistical methods.

In the RTM inversion approach, the in situ data are used both to parameterise RTM simulations and validate the retrievals [10], [41], [48]–[50]. In some cases, they may also be used to correct for biases in the retrieval results [10]. The machine learning approach incorporates an array of non-parametric regression methods, including, for example, artificial neural networks (ANNs) [10], [41]. It is suggested by [41] that such methods require a greatly increased number of samples for model training, reducing their utility in field campaigns with smaller numbers of ESUs. This limitation is somewhat overcome in hybrid approaches, which make use of RTM simulations to train a machine learning algorithm [43]. As with RTM inversion, the in situ measurements may be used to parameterise the RTM simulations, validate the retrievals, and correct for biases in the results. Finally, geostatistical methods involve the use of interpolation techniques such as kriging to upscale in situ measurements, making use of models of spatial autocorrelation [10], [41]. It is suggested by [41] that the additional gain in performance offered by such methods is small, demonstrated by the fact that a spatial covariate (i.e. the high spatial resolution imagery) is typically required to achieve adequate results.

High spatial resolution imagery from various multispectral instruments has been adopted for upscaling, including data from Landsat programme’s TM, ETM+ and OLI instruments, in addition to the SPOT HRV instruments. The spectral characteristics of these instruments are well suited to upscaling in situ measurements of LAI, FAPAR and FCOVER. However, their lack of red-edge bands has limited their utility for upscaling in situ measurements of CCC. Because no freely available source of high spatial resolution imagery was available, validation efforts previously required costly airborne hyperspectral data acquisition. The recent launch of the Sentinel-2 missions represents a solution to this challenge. The utility of the Sentinel-2 MSI’s red-edge bands for upscaling in situ measurements of CCC was recently demonstrated by [43].

### Convex Hulls

Because upscaling methods may have limited extrapolation capabilities, analysis is often restricted to the convex hull of the training data. According to [41], three types of convex hull can be considered:

* The spatial convex hull, which restricts analysis areas characterised by similar conditions as the sampled ESUs
* The temporal convex hull, which restricts analysis to similar time periods as the sampled ESUs
* The thematic convex hull, which restricts analysis on the basis of the joint distribution of auxiliary variables (e.g. the multispectral convex hull if multispectral data is used in upscaling)

In terms of the multispectral convex hull, [47] defines both a ‘strict’ and ‘large’ convex hull. Areas outside of the ‘strict’ convex hull but within the ‘large’ convex hull are derived by assuming 5% noise in the high spatial resolution imagery. It is suggested that although these technically areas represent extrapolation, they may nevertheless provide good prediction capabilities.

### Aggregation Method

Although [34] argue that the aggregation method adopted in validation studies should account for the apparent point spread function (PSF) of the satellite-derived product under validation, they also suggest that this step is, at present, poorly addressed. Because detailed knowledge on the instrument’s geometric characteristics would be required, accounting for the PSF in aggregation is not a straightforward process. The apparent PSF is influenced by instrument observational characteristics such as viewing geometry, particularly in the case of wide-swath instruments. Indeed, [43] implemented weighted mean downsampling aggregation, making use of a simple Gaussian PSF model, and found the method to perform poorly when compared to the standard mean value downsampling aggregation method. Because of these difficulties, few studies have explicitly considered the PSF. Instead, validation studies have typically performed comparison over multiple product pixels (i.e. 3 x 3), in the hope of minimising PSF effects.

### Validation Metrics

A range of metrics are commonly used to quantify the accuracy and uncertainty of satellite-derived vegetation products, however, there is a general a lack of agreement on which metrics should be reported. According to [51], validation metrics can be grouped into four categories: error indices, correlation based measures, dimensionless indices, and pattern indices (Table 4). It is suggested that as overall measures of accuracy, the root mean square error (RMSE) and coefficient of determination (*r*2) are useful as they are widely used and ease comparison with previous analyses, whilst the slope and intercept are useful as indicators of bias. To enable comparisons between variables with different units, the relative RMSE (RRMSE) is also recommended. In addition to these metrics, a range of plots should be produced to facilitate visual assessment of agreement (e.g. scatter plots, box plots of residuals), incorporating error bars where appropriate.

Table 4: Categories of validation metrics, after [51].

|  |  |  |  |
| --- | --- | --- | --- |
| **Category** | **Statistical metric** | **Abbreviation** | **Range** |
| Error indices | Root mean square error | RMSE | Variable units |
|  | Mean absolute error | MAE | Variable units |
|  | Mean bias error (bias) | B (MBE) | Variable units |
|  | Simulation bias | SB | Variable units |
|  | Standard error of predictioncorrected for bias | SEPC | Variable units |
| Correlation based indices | Coefficient of determination | *r*2 | 0 to 1 |
|  | Pearson correlation Coefficient | r | -1 to 1 |
|  | Slope | m | Range of variable values |
|  | Intercept | B | Range of variable values |
| Dimensionless indices | Nondimensional error index | NDI | 0 to ∞ |
|  | Normalised RMSE | NRMSE | 0 to ∞ |
|  | Relative RMSE | RRMSE | 0 to ∞ |
|  | SD to RMSE ratio | RDP | 0 to ∞ |
|  | Relative bias | Rel. B | -∞ to ∞ |
|  | Relative percentage error | RPE | 0 to ∞ |
|  | Nash–Sutcliffe efficiency | NSE | -∞ to 1 |
|  | Willmott’s index of agreement | d | 0 to 1 |
|  | Refined index of agreement | dr | -1 to 1 |
|  | Agreement coefficient | AC | 0 to 1 |
| Pattern indices | Range-based fixed pattern index | PI | Variable units |
|  | F-based fixed pattern index | PI-F | Variable units |

# Validation Status and Challenges

## Validation Status of Considered Products

Existing validation activities related to the OLCI L2 land products have been coordinated by the Sentinel-3 Mission Performance Centre (S3MPC). The S3MPC forms part of the mission’s payload data ground segment (PDGS) and is responsible for quality control of all Sentinel-3 data products. Within the S3MPC, routine quality control has been performed by assessing the temporal consistency of the products and comparing them to a 10-year climatology established using the equivalent MERIS products. In the case of the OGVI, inter-comparison with FAPAR derived from MODIS using the same algorithm has also been carried out [52]. These activities have been performed over 55 ‘supersites’ proposed by the CEOS WGCV LPV sub-group, in addition to 11 ‘core’ validation sites, which were selected by the Sentinel-3 Validation Team (S3VT).

Overall, good temporal consistency has been observed, and both the OGVI and OTCI appear consistent with the MGVI and MTCI climatology over the majority of sites (within the expected range of interannual variability). Nevertheless, the OGVI has been reported to demonstrate a slight positive bias with respect to the MGVI climatology, whilst over some sparsely vegetated sites, some degree of underestimation of the OTCI with respect to the MTCI climatology has been observed [53]. When compared to FAPAR derived from MODIS, the OGVI demonstrates good agreement [52]. Analysis of products acquired during the Sentinel-3A and -3B tandem phase indicate some differences between the products generated from the two OLCI instruments, and further investigation is required [53].

In terms of direct validation, two dedicated campaigns have been carried out in the case of the OTCI, whilst in situ FAPAR data from several sites have been used to validate the OGVI [43], [52]. The OTCI direct validation campaigns have involved collection of in situ CCC measurements over a) a deciduous broadleaf forest site, and b) a vineyard dominated Mediterranean site. These in situ measurements have then been upscaled to the moderate spatial resolution of the OTCI using MSI data. In both cases, strong correlations between the OTCI and CCC were observed [43], [52]. In terms of the OGVI, in situ FAPAR data reveal it appears to be subject to some degree of underestimation, particularly at the peak of the growing season. However, rather than being caused by deficiencies in the product, some degree of underestimation is expected due to the definition of the in situ measurements, which correspond to the total rather than green FAPAR.

## Gaps and Challenges in Current Validation Protocols

Although much progress in the validation of satellite-derived vegetation products has been made over the last two decades, several gaps and challenges remain. Major gaps and challenges in current validation protocols and procedures are identified and summarised below:

* Individual field campaigns fail to adequately characterise vegetation temporal dynamics, and some biomes are underrepresented in existing direct validation efforts
* Little consideration has been given to in situ measurement uncertainties and their propagation through subsequent processing and analysis steps (i.e. in situ measurements are simply assumed to represent the truth)
* The uncertainties associated with in situ measurements are not currently utilised in the development of empirical transfer functions (although robust regression approaches are used to supress outliers)
* High spatial resolution reference maps are not currently provided with per-pixel uncertainties (although quality flags are produced)
* Protocols and procedures for the validation of satellite-derived CCC products are less developed than those for FAPAR (due in part to the historical lack of freely available high spatial resolution imagery incorporating appropriate spectral bands)
* The traceability of the satellite-derived products in question is not always clear, whilst the definitions and assumptions associated with in situ measurements do not always match those of the satellite-derived products themselves

# Recommended Validation Strategy

## Sampling Strategy

For logistical reasons including the need to avoid destructive sampling, indirect, optical techniques are recommended to perform in situ measurements. As discussed in Section 3.2.3, the optimal number of individual in situ measurements within an ESU is driven by its heterogeneity and the measurement footprint of the instrument. In line with previous recommendations, it is recommended that each ESU should contain at least thirteen sampling points. Additionally, it is recommended that several further measurements should be performed at randomly distributed sampling points, providing redundancy if there are any anomalous or non-representative measurements that need to be discarded during quality control. The centre of each ESU should be located using a handheld global positioning system (GPS) receiver, with a minimum accuracy of a few metres, enabling the ESU to be identified in the high spatial resolution imagery.

Although the heterogeneity of the site will determine the required number of ESUs, a minimum of twenty are recommended. ESUs should be distributed across the site, stratified by land cover to ensure all dominant vegetation types are represented. In agricultural environments, it is recommended to sample at least five ESUs per field to characterize the intra-field variability. To avoid adjacency effects, ESUs should be located at least 50 m from borders, and should be surrounded by pixels with the same type of vegetation. It is recommended that at least five additional ESUs are established over areas with little to no vegetation cover to constrain the lower boundary of the transfer function used for upscaling. Knowledge of land cover characteristics will enable identification of these areas. In agricultural environments in particular, a field survey may be required so that up-to-date information on land cover characteristics at the time of the campaign itself is available.

## Quality Control

Prior to upscaling, quality control should be carried out to screen both in situ measurements and high spatial resolution imagery for poor quality data. In the case of the high spatial resolution imagery, the quality flags provided with the product should be used to discard pixels contaminated by radiometric saturation or cloud cover. In the case of the in situ measurements, any measurement with a value of greater than two standard deviations from the ESU mean should be examined to assess its validity. If the measurement is determined to be spurious, it should be discarded prior to further analysis [45].

## Upscaling Approach

An upscaling approach based on the use of empirical transfer functions is recommended. Although more complex approaches are available, they have several drawbacks, some of which are discussed in Section 3.2.4. In the case of RTM inversion and hybrid approaches, these include the need to select and parameterise an RTM to accurately simulate the canopy of interest, whilst in the case of machine learning approaches, they include the need to sample an extensive number of ESUs.

In terms of the transfer functions themselves, it is suggested by [41] that a transfer function with two predictor variables would require 1024 samples to achieve the same precision as one based on a single predictor variable with only 10 samples. As such, models with a single predictor variable should be favoured unless their performance is significantly worse than those with a greater number of predictor variables. Transfer functions based on vegetation indices are therefore recommended. Following the recommendations of [41], vegetation indices should be selected based on known relationships with the variable of interest, enabling the form of the transfer function to be selected a priori.

## High Spatial Resolution Imagery

It is recommended to use L1 Sentinel-2 MSI data (i.e. top-of-atmosphere reflectance) for upscaling. At the peak of the growing season, when the condition of the vegetation can be considered stable, it is recommended that MSI data are acquired within one week of in situ data collection [10], [54]. If in situ data are collected during the start or end of the growing season, when more rapid changes in biophysical or biochemical characteristics may occur, a more restrictive temporal constraint will be required. This will depend on the phenological characteristics of the site itself. The rationale for using L1 data is that per-pixel uncertainty information is not available for atmospherically corrected L2 surface reflectance products produced using the Sen2Cor algorithm, whereas per-pixel uncertainties can be easily derived for L1 MSI data using the Radiometric Uncertainty Tool (RUT) developed by [55]. As described in Section 3.2.4, because of the empirical nature of the transfer functions, atmospheric correction is not mandatory. The RUT currently incorporates the following uncertainty sources:

* Instrument noise (shot, thermal etc. noise introduced by the detectors)
* Out-of-field straylight systematic (telescope out-of-field light that results in a positive bias)
* Crosstalk (focal plane (optical) and front-end electronics (electrical) inter-band signal)
* Analogue-to-digital conversion quantisation (at MSI’s video chain unit)
* Dark signal stability (residual thermal fluctuations of the detector offset along the orbit)
* Gamma knowledge (knowledge on the correction for non-linearity and non-uniformity)
* Diffuser absolute knowledge (knowledge on the diffuser reflectance factor)
* Diffuser temporal knowledge (estimated effect of diffuser degradation)
* Diffuser cosine effect (cosine correction knowledge as a consequence of angular noise)
* Diffuser straylight residual (residual of the correction of the stray-light during in-flight diffuser calibration)
* L1C image quantisation (effect of the finite resolution of the L1C reflectance factor)

If data from bands with different spatial resolutions (i.e. 10 m and 20 m) are to be used, the higher spatial resolution band should be aggregated to that of the lower spatial resolution one. To correctly propagate the radiometric uncertainties through the aggregation procedure, the ‘select/deselect’ approach described by [56] should be adopted. This involves running the RUT twice, enabling uncertainty sources that are uncorrelated in space to be separated from those that are correlated in space. The mean of the two outputs (i.e. all uncertainty components and only those correlated in space) should then be taken as the radiometric uncertainty in the aggregated pixel values.

## Incorporation of Uncertainties in Upscaling

Although upscaling approaches based on the use of empirical transfer functions are well established, few studies have explicitly considered in situ measurement uncertainties, or uncertainties in the high spatial resolution imagery. As discussed in Section 3.2.4 current approaches aim to establish transfer functions that are robust to outliers, making use of methods such as the Thiel-Sen estimator IRLS, but the individual uncertainties associated to each data point are not used to inform the fitting procedure. To address this, a modified upscaling approach is recommended.

Making use of the per-pixel uncertainties associated with the high spatial resolution imagery (i.e. provided by the Sentinel-2 RUT) in addition to those associated with the in situ measurements (see the ‘FRM Protocols and Procedures for Surface Reflectance, Fraction of Absorbed Photosynthetically Active Radiation and Canopy Chlorophyll Content (FPP)’ document), it is recommended that the transfer function is established using orthogonal distance regression (ODR). Unlike other methods, ODR is able to incorporate uncertainties in both predictor and response variables. When applied, uncertainties in the regression coefficients and high spatial resolution imagery should be propagated through the transfer function, enabling per-pixel uncertainties to be obtained along with the high spatial resolution reference map itself.

## Derivation of Quality Flags

It is recommended that a quality flag layer is provided along with the high spatial resolution reference map. As described in Section 3.2.5, a quality flag layer based on the multispectral convex hull may be used to identify areas where the transfer function is extrapolating. In this case, both the ‘strict’ and ‘large’ convex hull should be provided.

# References

[1] N. Gobron, B. Pinty, M. Verstraete, and Y. Govaerts, “The MERIS Global Vegetation Index (MGVI): Description and preliminary application,” *Int. J. Remote Sens.*, vol. 20, no. 9, pp. 1917–1927, Jan. 1999.

[2] N. Gobron, B. Pinty, M. M. Verstraete, and J.-L. Widlowski, “Advanced vegetation indices optimized for up-coming sensors: Design, performance, and applications,” *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 6, pp. 2489–2505, 2000.

[3] H. Rahman, B. Pinty, and M. M. Verstraete, “Coupled surface-atmosphere reflectance (CSAR) model: 2. Semiempirical surface model usable with NOAA advanced very high resolution radiometer data,” *J. Geophys. Res.*, vol. 98, no. D11, p. 20791, 1993.

[4] N. Gobron, *Ocean and Land Colour Instrument (OLCI) FAPAR and Rectified Channels over Terrestrial Surfaces Algorithm Theoretical Basis Document*. Ispra, Italy: European Commission Joint Research Centre, 2010.

[5] M. Weiss *et al.*, “On line validation exercise (OLIVE): A web based service for the validation of medium resolution land products. application to FAPAR products,” *Remote Sens.*, vol. 6, no. 5, pp. 4190–4216, 2014.

[6] K. Yan *et al.*, “Evaluation of MODIS LAI/FPAR Product Collection 6. Part 2: Validation and Intercomparison,” *Remote Sens.*, vol. 8, no. 6, p. 460, May 2016.

[7] E. F. Vermote, D. Tanre, J. L. Deuze, M. Herman, and J.-J. Morcette, “Second Simulation of the Satellite Signal in the Solar Spectrum, 6S: an overview,” *IEEE Trans. Geosci. Remote Sens.*, vol. 35, no. 3, pp. 675–686, May 1997.

[8] D. N. H. Horler, M. Dockray, and J. Barber, “The red edge of plant leaf reflectance,” *Int. J. Remote Sens.*, vol. 4, no. 2, pp. 273–288, Jan. 1983.

[9] P. J. Curran, J. L. Dungan, and H. L. Gholz, “Exploring the relationship between reflectance red edge and chlorophyll content in slash pine,” *Tree Physiol.*, vol. 7, no. 1-2-3–4, pp. 33–48, Dec. 1990.

[10] F. Baret *et al.*, *VALERI: a network of sites and a methodology for the validation of medium spatial resolution land satellite products*. Avignon, France: Institut National de la Recherche Agronomique, 2005.

[11] G. F. Bonham-Carter, “Numerical procedures and computer program for fitting an inverted gaussian model to vegetation reflectance data,” *Comput. Geosci.*, vol. 14, no. 3, pp. 339–356, Jan. 1988.

[12] J. R. Miller, W. Jiyou, M. G. Boyer, M. Belanger, and E. W. Hare, “Seasonal patterns in leaf reflectance red-edge characteristics,” *Int. J. Remote Sens.*, vol. 12, no. 7, pp. 1509–1523, Jul. 1991.

[13] G. Guyot, F. Baret, and D. J. Major, “High spectral resolution: Determination of spectral shifts between the red and infrared,” *Int. Arch. Photogramm. Remote Sens.*, vol. 11, no. 1, pp. 750–760, 1988.

[14] F. M. Danson and S. E. Plummer, “Red-edge response to forest leaf area index,” *Int. J. Remote Sens.*, vol. 16, no. 1, pp. 183–188, Jan. 1995.

[15] T. P. Dawson and P. J. Curran, “Technical note A new technique for interpolating the reflectance red edge position,” *Int. J. Remote Sens.*, vol. 19, no. 11, pp. 2133–2139, Jan. 1998.

[16] T. Majasalmi, M. Rautiainen, P. Stenberg, and H. Rita, “Optimizing the sampling scheme for LAI-2000 measurements in a boreal forest,” *Agric. For. Meteorol.*, vol. 154–155, pp. 38–43, 2012.

[17] J. Dash and P. J. Curran, “The MERIS terrestrial chlorophyll index,” *Int. J. Remote Sens.*, vol. 25, no. 23, pp. 5403–5413, Dec. 2004.

[18] J. Dash, *Algorithm Theoretical Basis Document: OLCI Terrestrial Chlorophyll Index (OTCI)*. Southampton, United Kingdom: University of Southampton, 2012.

[19] T. Miura, A. R. Huete, and H. Yoshioka, “Evaluation of sensor calibration uncertainties on vegetation indices for MODIS,” *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 3, pp. 1399–1409, May 2000.

[20] Working Group 1 of the Joint Committee for Guides in Metrology, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement*. Paris, France: Bureau International des Poids et Mesures, 2008.

[21] ESA, “Land Products Quality and Science Flags.” [Online]. Available: https://sentinel.esa.int/web/sentinel/technical-guides/sentinel-3-olci/level-2/quality-and-science-flags-lp. [Accessed: 17-Mar-2019].

[22] J. . Privette *et al.*, “The EOS Prototype Validation Exercise (PROVE) at Jornada,” *Remote Sens. Environ.*, vol. 74, no. 1, pp. 1–12, Oct. 2000.

[23] J. L. Campbell, S. Burrows, S. T. Gower, and W. B. Cohen, *BigFoot: Characterizing Land Cover, LAI and NPP at the Landscape Scale for EOS/MODIS Validation - Field Manual*, 2.1. Oak Ridge, Tennessee, United States: Oak Ridge National Laboratory, 1999.

[24] R. J. Swap and J. L. Privette, “Overview of the Southern African Regional Science Initiative - SAFARI 2000,” in *Proceedings of the 1999 IEEE International Geoscience and Remote Sensing Symposium*, 1999.

[25] Y. Tian *et al.*, “Multiscale analysis and validation of the MODIS LAI product II. Sampling strategy,” *Remote Sens. Environ.*, vol. 83, no. 3, pp. 431–441, Dec. 2002.

[26] S. Garrigues *et al.*, “Validation and intercomparison of global Leaf Area Index products derived from remote sensing data,” *J. Geophys. Res.*, vol. 113, p. G02028, Jun. 2008.

[27] F. Camacho, J. Cernicharo, R. Lacaze, F. Baret, and M. Weiss, “GEOV1: LAI, FAPAR essential climate variables and FCOVER global time series capitalizing over existing products. Part 2: Validation and intercomparison with reference products,” *Remote Sens. Environ.*, vol. 137, pp. 310–329, Oct. 2013.

[28] F. Baret *et al.*, “Evaluation of the representativeness of networks of sites for the global validation and intercomparison of land biophysical products: proposition of the CEOS-BELMANIP,” *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 7, pp. 1794–1803, Jul. 2006.

[29] M. Weiss, F. Baret, S. Garrigues, and R. Lacaze, “LAI and fAPAR CYCLOPES global products derived from VEGETATION. Part 2: validation and comparison with MODIS collection 4 products,” *Remote Sens. Environ.*, vol. 110, no. 3, pp. 317–331, Oct. 2007.

[30] M. Weiss *et al.*, “On Line Validation Exercise (OLIVE): A Web Based Service for the Validation of Medium Resolution Land Products. Application to FAPAR Products,” *Remote Sens.*, vol. 6, no. 5, pp. 4190–4216, May 2014.

[31] K. Yan *et al.*, “Evaluation of MODIS LAI/FPAR Product Collection 6. Part 1: Consistency and Improvements,” *Remote Sens.*, vol. 8, no. 5, p. 359, Apr. 2016.

[32] S. Garrigues *et al.*, “Validation and intercomparison of global Leaf Area Index products derived from remote sensing data,” *J. Geophys. Res. Biogeosciences*, vol. 113, no. 2, 2008.

[33] M. Weiss, F. Baret, S. Garrigues, and R. Lacaze, “LAI and fAPAR CYCLOPES global products derived from VEGETATION. Part 2: validation and comparison with MODIS collection 4 products,” *Remote Sens. Environ.*, vol. 110, no. 3, pp. 317–331, 2007.

[34] J. T. Morisette *et al.*, “Validation of global moderate-resolution LAI products: a framework proposed within the CEOS land product validation subgroup,” *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 7, pp. 1804–1817, Jul. 2006.

[35] K. Fitzpatrick-Lins, “Comparison of sampling procedures and data analysis for a land-use and land-cover map,” *Photogramm. Eng. Remote Sensing*, vol. 47, no. 3, pp. 343–351, 1981.

[36] F. Camacho, F. Baret, and R. Lacaze, *Guidelines for a Field Campaign*, 1.10. Valencia, Spain: EOLAB.

[37] C. O. Justice and J. R. G. Townshend, “Integrating ground data with remote sensing,” in *Terrain Analysis and Remote Sensing*, J. R. G. Townshend, Ed. London, United Kingdom: Allen and Unwin, 1981, pp. 35–58.

[38] LI-COR, *LAI-2200C Plant Canopy Analyser Instruction Manual*. Lincoln, Nebraska, United States: LI-COR, 2013.

[39] T. Majasalmi, M. Rautiainen, P. Stenberg, and H. Rita, “Optimizing the sampling scheme for LAI-2000 measurements in a boreal forest,” *Agric. For. Meteorol.*, vol. 154–155, pp. 38–43, Mar. 2012.

[40] M. Weiss, F. Baret, G. J. Smith, I. Jonckheere, and P. Coppin, “Review of methods for in situ leaf area index (LAI) determination Part II: Estimation of LAI, errors and sampling,” *Agric. For. Meteorol.*, vol. 121, no. 1–2, pp. 37–53, Jan. 2004.

[41] R. Fernandes *et al.*, “Global Leaf Area Index Product Validation Good Practices,” in *Best Practice for Satellite-Derived Land Product Validation*, 2.0., R. Fernandes, S. Plummer, and J. Nightingale, Eds. Land Product Validation Subgroup (Committee on Earth Observation Satellites Working Group on Calibration and Validation), 2014.

[42] A. Süß, M. Danner, C. Obster, M. Locherer, T. Hank, and K. Richter, *EnMAP Field Guides: Technical Report: Measuring Leaf Chlorophyll Content with the Konica Minolta SPAD-502 Plus*. Potsdam, Germany: GeoForschungsZentrum Data Services, 2015.

[43] L. A. Brown, J. Dash, A. L. Lidon, E. Lopez-Baeza, and S. Dransfeld, “Synergetic Exploitation of the Sentinel-2 Missions for Validating the Sentinel-3 Ocean and Land Color Instrument Terrestrial Chlorophyll Index Over a Vineyard Dominated Mediterranean Environment,” *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 12, no. 7, pp. 2244–2251, Jul. 2019.

[44] J. Dash, P. J. Curran, M. J. Tallis, G. M. Llewellyn, G. Taylor, and P. Snoeij, “Validating the MERIS Terrestrial Chlorophyll Index (MTCI) with ground chlorophyll content data at MERIS spatial resolution,” *Int. J. Remote Sens.*, vol. 31, no. 20, pp. 5513–5532, Oct. 2010.

[45] G. Watmough, J. Dash, F. Vuolo, and W. Frampton, *Technical Note 2 - Validation Protocols*, 1st ed. Southampton, United Kingdom: University of Southampton, 2011.

[46] R. Fernandes and S. Leblanc, “Parametric (modified least squares) and non-parametric (Theil–Sen) linear regressions for predicting biophysical parameters in the presence of measurement errors,” *Remote Sens. Environ.*, vol. 95, no. 3, pp. 303–316, Apr. 2005.

[47] B. Martínez, F. J. García-Haro, and F. Camacho-de Coca, “Derivation of high-resolution leaf area index maps in support of validation activities: Application to the cropland Barrax site,” *Agric. For. Meteorol.*, vol. 149, no. 1, pp. 130–145, Jan. 2009.

[48] A. Verger, F. Baret, and F. Camacho, “Optimal modalities for radiative transfer-neural network estimation of canopy biophysical characteristics: Evaluation over an agricultural area with CHRIS/PROBA observations,” *Remote Sens. Environ.*, vol. 115, no. 2, pp. 415–426, Feb. 2011.

[49] J. Heiskanen, M. Rautiainen, L. Korhonen, M. Mõttus, and P. Stenberg, “Retrieval of boreal forest LAI using a forest reflectance model and empirical regressions,” *Int. J. Appl. Earth Obs. Geoinf.*, vol. 13, no. 4, pp. 595–606, Aug. 2011.

[50] M. Claverie, E. F. Vermote, M. Weiss, F. Baret, O. Hagolle, and V. Demarez, “Validation of coarse spatial resolution LAI and FAPAR time series over cropland in southwest France,” *Remote Sens. Environ.*, vol. 139, pp. 216–230, Dec. 2013.

[51] K. Richter, C. Atzberger, T. B. Hank, and W. Mauser, “Derivation of biophysical variables from Earth observation data: validation and statistical measures,” *J. Appl. Remote Sens.*, vol. 6, no. 1, pp. 063557–1, Sep. 2012.

[52] ESA, “S3 Product Notice – OLCI,” 2019. [Online]. Available: https://sentinel.esa.int/documents/247904/2700530/Sentinel-3-Product-Notice-OLCI.pdf.

[53] ESA, “Data Product Quality Reports.” [Online]. Available: https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-3-olci/data-quality-reports. [Accessed: 18-Mar-2019].

[54] M. G. De Kauwe, M. I. Disney, T. Quaife, P. Lewis, and M. Williams, “An assessment of the MODIS collection 5 leaf area index product for a region of mixed coniferous forest,” *Remote Sens. Environ.*, vol. 115, no. 2, pp. 767–780, Feb. 2011.

[55] J. Gorroño *et al.*, “A Radiometric Uncertainty Tool for the Sentinel 2 Mission,” *Remote Sens.*, vol. 9, no. 2, p. 178, Feb. 2017.

[56] J. Gorroño *et al.*, “Providing uncertainty estimates of the Sentinel-2 top-of-atmosphere measurements for radiometric validation activities,” *Eur. J. Remote Sens.*, vol. 51, no. 1, pp. 650–666, Jan. 2018.

###### Appendix

*This section is pending further development during the course of FRM4VEG Phase 2*