



TERRASCOPE SENTINEL-2

ALGORITHM THEORETICAL BASE DOCUMENT (ATBD)

S2 – NDVI & BIOPAR – V220

Reference: *Terrascope Sentinel-2 Algorithm Theoretical Base Document S2 – NDVI & BIOPAR – V220*

Author(s): Isabelle Piccard, Liesbeth De Keukelaere, Else Swinnen, Dieter Wens, Stijn Caerts, Lieve Van den Heuvel, Simon Van Baelen

Version: 2.2

Date: 6/02/2026

DOCUMENT CONTROL

Signatures

Authors	Isabelle Piccard, Liesbeth De Keukelaere, Else Swinnen, Dieter Wens, Stijn Caerts, Lieve Van den Heuvel, Simon Van Baelen
Reviewers	Carolien Toté
Approvers	Dennis Clarijs, Jurgen Everaerts
Issuing authority	VITO

Change record

Release	Date	Updates	Approved by
0.1	01/02/2019	Initial external version	
1.0	15/10/2019	Updated	J. Everaerts
2.0	05/05/2020	TERRASCOPE Sentinel-2 Version 200	J. Everaerts
2.2	06/02/2026	TERRASCOPE Sentinel-2 Version 220	J. Everaerts

© VITO N.V. 2019

The copyright in this document is vested in VITO N.V.

This document may only be reproduced in whole or in part, or stored in a retrieval system, or transmitted, or copied, in any form, with the prior permission of VITO NV.

TABLE OF CONTENTS

1. INTRODUCTION.....	10
1.1. Terrascope explained	10
1.2. Scope of Document	10
1.3. Description.....	11
1.4. Feature added value/use case.....	12
1.5. Related documents	13
1.6. Definitions.....	14
1.6.1. LAI	14
1.6.2. FAPAR	14
1.6.3. FCOVER	15
1.6.4. CCC.....	15
1.6.5. CWC	15
2. INPUT	16
2.1. Top-Of-Canopy reflectance.....	16
2.2. Ancillary data and models	17
2.2.1. Neural Networks.....	17
2.2.2. Angle information.....	17
3. OUTPUT	18
3.1. Product layers	18
3.2. Product version	19
3.3. Product data access.....	20
4. METHODOLOGY	21
4.1. NDVI (10 m resolution).....	21
4.1.1. Justification.....	21
4.1.2. Implementation	21
4.2. BIOPAR (10 m and 20 m resolution)	21
4.2.1. Justification.....	21
4.2.2. Method.....	22
4.2.2.1. BIOPARS at 20 m resolution (NN-R8)	22
4.2.2.2. BIOPARS at 10 m resolution (NN-R3)	23
4.2.3. Implementation	23
5. LIMITATIONS	25
5.1. Algorithm limitations	25
5.2. Implementation limitations.....	25
5.3. Validation of the products.....	25
6. QUALITY ASSESSMENT	26
6.1. Terrascope Sentinel-2 v200 vs V102	26
6.1.1. SCENECLASSIFICATION layer.....	26
6.1.2. BIOPAR and NDVI comparison.....	28
6.1.2.1. FAPAR	28
6.1.2.2. FCOVER.....	29
6.1.2.3. LAI.....	30
6.1.2.4. NDVI	31

7. OTHER REFERENCE DOCUMENTS 33

LIST OF FIGURES

FIGURE 6.1 FRACTION [%] OF EACH CATEGORY IN THE YEAR 2018 FOR TILE 31UFS. THE BLUE BARS REPRESENT THE RESULTS FOR V102, GREEN BARS FOR V200.	27
FIGURE 6.2 FREQUENCY PLOT (LEFT) AND BIAS HISTOGRAMS (RIGHT) OF FAPAR 20 M PRODUCTS FOR V102 AND V200 FOR THE FULL TIME SERIES OF TILE 31UFS (I.E. STARTING IN 2015).	29
FIGURE 6.3 FREQUENCY PLOT (LEFT) AND BIAS HISTOGRAMS (RIGHT) OF FAPAR 10 M PRODUCTS FOR V102 AND V200 FOR THE FULL TIME SERIES OF TILE 31UFS (I.E. STARTING IN 2015).	29
FIGURE 6.4 FREQUENCY PLOT (LEFT) AND BIAS HISTOGRAMS (RIGHT) OF FCOVER 20 M PRODUCTS FOR V102 AND V200 FOR THE FULL TIME SERIES OF TILE 31UFS (I.E. STARTING IN 2015).	30
FIGURE 6.5 FREQUENCY PLOT (LEFT) AND BIAS HISTOGRAMS (RIGHT) OF FCOVER 10 M PRODUCTS FOR V102 AND V200 FOR THE FULL TIME SERIES OF TILE 31UFS (I.E. STARTING IN 2015).	30
FIGURE 6.6 FREQUENCY PLOT (LEFT) AND BIAS HISTOGRAMS (RIGHT) OF LAI 20 M PRODUCTS FOR V102 AND V200 FOR THE FULL TIME SERIES OF TILE 31UFS (I.E. STARTING IN 2015).	31
FIGURE 6.7 FREQUENCY PLOT (LEFT) AND BIAS HISTOGRAMS (RIGHT) OF LAI 10 M PRODUCTS FOR V102 AND V200 FOR THE FULL TIME SERIES OF TILE 31UFS (I.E. STARTING IN 2015).	31
FIGURE 6.8 FREQUENCY PLOT (LEFT) AND BIAS HISTOGRAMS (RIGHT) OF NDVI 10 M PRODUCTS FOR V102 AND V200 FOR THE FULL TIME SERIES OF TILE 31UFS (I.E. STARTING IN 2015).	32

LIST OF TABLES

TABLE 1.1: SUMMARY OF MAIN CHARACTERISTICS BETWEEN DIFFERENT TERRASCOPE VERSIONS UNTIL V200.....	12
TABLE 1.2: LIST OF CHANGES BETWEEN TERRASCOPE SENTINEL-2 V220 AND V200.....	12
TABLE 1.3: LIST OF RELATED DOCUMENTS.....	13
TABLE 2.1: SPATIAL AND SPECTRAL CHARACTERISTICS OF THE S2 TOC PRODUCTS. BANDS USED FOR NN-R8 IN ORANGE SHADING AND FOR NN-R3 IN BLUE CHARACTERS.	16
TABLE 3.1: CHARACTERISTICS OF THE NDVI AND BIOPAR IMAGES AND RESCALING INFORMATION. PHYSICAL MIN AND MAX ARE THE PHYSICAL RANGE THAT IS RETAINED IN THE OUTPUT, THE DIGITAL NUMBERS (DN) ARE THE VALUE OF THE PHYSICAL MIN AND MAX AFTER RESCALING TO BYTE. THE SLOPE AND OFFSET ARE THE COEFFICIENTS TO USE TO RECOMPUTE THE PHYSICAL VALUES FROM THE BYTE OUTPUT IMAGES USING EQUATION EQ 1.	18
TABLE 3.2: MEANING OF THE VALUES IN THE SCENE CLASSIFICATION LAYER.	18
TABLE 4.1: VALID RANGE OF OUTPUT VALUES.....	24
TABLE 6.1 CONFUSION MATRIX OF THE SCENECLASSIFICATION LAYER BETWEEN V102 AND V200. THE NUMBERS ARE EXPRESSED AS PERCENTAGES.	27

LIST OF ACRONYMS

ACRONYM	EXPLANATION
ATBD	Algorithm Theoretical Base Document
BIOPAR	Biophysical Parameter
CCC	Canopy Chlorophyll Content
CGS	Collaborative Ground Segment
CWC	Canopy Water Content
DEM	Digital Elevation Model
ECV	Essential Climate Variable
ESA	European Space Agency
fAPAR	Fraction of Absorbed Photosynthetically Active Radiation
fCOVER	Fraction of green vegetation Cover
FTP	File Transfer Protocol
GCOS	Global Climate Observing System
GeoTIFF	Geospatial Tagged Image File Format
GLOBE	Global Land One-km Base Elevation Project
iCOR	Image Correction for atmospheric effects
IDL	Interactive Data Language
L1C	Radiometrically and geometrically calibrated Level-1 data
LAI	Leaf Area Index
LUT	Look-Up-Table
MEP	Mission Exploitation Platform
MSI	Multi-Spectral Instrument
MODTRAN5	MODerate resolution atmospheric TRANsmission
MSI	Multispectral Instrument
NDVI	Normal Difference Vegetation Index
NIR	Near-Infrared
NRT	Near Real Time
OGC	Open Geospatial Consortium
PDGS	Payload Data Ground Segment
PDP	Product Distribution Portal
PROBA-V	Project for On-Board Autonomy – Vegetation
PUM	Product User Manual
RAA	Relative Azimuth Angle
RD	Related Document
ROI	Region Of Interest
S2	Sentinel-2
SAA	Sun Azimuth Angle
SAR	Synthetic Aperture
Sen2Cor	Sentinel-2 Correction
SNAP	Sentinel Application Platform
SOLSPEC	SOLar SPECTrum

Terrascope Sentinel-2

List of acronyms

SPOT-VGT	Satellite Pour l'Observation de la Terre - Végétation
SRF	Spectral Response Function
STEP	Science Toolbox Exploitation Platform
SWIR	Short Wave InfraRed
SYN	Synergy
SZA	Solar Zenith Angle
TOA	Top-Of-Atmosphere
TOC	Top-Of-Canopy
UTC	Universal Time Coordinate
UTM	Universal Transverse Mercator
VAA	View Azimuth Angle
VITO	Vlaamse Instelling voor Technologisch Onderzoek
VM	Virtual Machine
VNIR	Visible and Near InfraRed
VZA	View Zenith Angle
WGS84	World Geodetic System 1984
XML	Extensible Markup Language

1. Introduction

1.1. Terrascope explained

Terrascope is the Belgian platform for accessing Copernicus, PROBA-V, and SPOT-VEGETATION satellite data, products, and services. It offers easy, full, free and open access to all users without restrictions. This enables both specialist and non-specialist users to explore the wealth of available remote sensing information and to develop value-added products and services.

Terrascope currently provides access to the following datasets:

- The full SPOT-VEGETATION archive
- The full PROBA-V archive
- Sentinel-1 SAR data over Belgium and its surroundings
- Sentinel-2 optical data over Europe and Africa
- Sentinel-3 optical Synergy (SYN) – Vegetation (VGT) and thermal Land Surface Temperature data

In addition to Sentinel-2 surface reflectance (Top-Of-Canopy, TOC) data, Terrascope provides a standard suite of biophysical indicators derived from Sentinel-2—such as NDVI, FAPAR, FCOVER, LAI, CCC, and CWC—as well as water quality products, including aquatic reflectance (RHOW), Suspended Particulate Matter (SPM), turbidity (TUR), and chlorophyll-a (CHL) concentrations.

Users can generate their own derived information products tailored to specific needs by leveraging the Terrascope processing cluster—available through virtual machines or interactive Notebooks. This approach eliminates the need for local data downloads and storage, as all data are directly accessible within the cluster environment. Integration of Terrascope data and products into external applications is supported through Open Geospatial Consortium (OGC) web services.

Terrascope is user-centered, and we welcome all suggestions for new or improved functionality. For feedback or inquiries, please contact us at info@terrascope.be.

1.2. Scope of Document

This Algorithm Theoretical Base Document (ATBD) describes the processing steps to go from the Sentinel-2 (S2) Level2 Top-Of-Canopy (TOC) reflectance products to NDVI and Biophysical Parameters (BIOPARs), embedded in the Terrascope Sentinel-2 v220 processing chain.

The document is organised as following:

- Section 2 provides an overview of all input data needed to feed the processing workflow.
- Section 3 explains the data available to users.
- Section 4 provides a detailed description of the different algorithms that compose the L2 to NDVI & BIOPAR workflow.
- Section 5 discusses the limitations of the implemented algorithms.
- Section 6 justifies the overall workflow with a quality assessment.

1.3. Description

In the first step of the Terrascope S2 processing chain, S2 Level-2A (L2) Top-Of-Canopy (TOC) products [RD1] were downloaded from COPERNICUS data hub.

From the TOC reflectance products the NDVI and BIOPARs are derived. The BIOPARs are:

- Leaf Area Index (LAI),
- Fraction of Absorbed Photosynthetically Active Radiation (FAPAR),
- Fraction of Vegetation Cover (FCOVER),
- Canopy Chlorophyll Content (CCC)
- Canopy Water Content (CWC).

LAI and FAPAR are Essential Climate Variables (ECVs) as defined by the Global Climate Observing System (GCOS) in support of the work of the UNFCCC and the IPCC.

The methodology used to derive the biophysical parameters from Sentinel-2 is developed by INRA-EMMAH. The methodology was initially developed to generate biophysical products from SPOT-VEGETATION, ENVISAT-MERIS, SPOT-HRVIR, and LANDSAT-OLI sensors and was later adapted for Sentinel-2. It mainly consists in simulating a comprehensive data base of canopy (TOC) reflectances based on vegetation characteristics and observation and illumination geometry. Neural networks are then trained to estimate a number of these canopy characteristics (BIOPARs) from the simulated TOC reflectances along with set corresponding angles defining the observational configuration.

For Sentinel-2 two neural networks (NN) were developed by INRA-EMMAH and implemented in Terrascope:

- one based on 10m input bands only (R3-NN)
- one based on a combination of 10m and 20m input bands (R8-NN). This NN is also implemented in the Sentinel-2 Toolbox[RD3].

The document is applicable for the Terrascope S2 v220 processing chain. Table 1.1 summarizes the main characteristics of the different Terrascope versions until v220. The changes between v102 and v200 are mainly related to the changes in the L2 TOC products and are listed in . Validation results between both versions are included in Section 6.1.

Table 1.1: Summary of main characteristics between different Terrascope versions until v200.

Version	Main characteristics
V101	Input: Terrascope TOC v 101 Applied masks on NDVI & BIOPAR products: cloud and cloud shadow
V102	Input: Terrascope TOC v102 Applied masks on NDVI & BIOPAR products: cloud and cloud shadow
V200	Input: Terrascope TOC v200 L2A data (NRT) or L1C (historic) Applied masks on NDVI & BIOPAR products: cloud, cloud shadow, snow, cirrus or saturated pixels
V220	Input: Terrascope TOC v220 Applied masks on NDVI & BIOPAR products: cloud cover mask

Table 1.2: List of changes between TERRASCOPE Sentinel-2 v220 and v200

Adaptations between v220 and v200	Clarification
Input data	
L2A supported from ESA Processing Baseline V5.0	Ensures consistent time series: v220 includes a full reprocessing of the TERRASCOPE-S2 archive starting from ESA Processing Baseline 5.0

1.4. Feature added value/use case

Terrascope provides easy access not only to the basic S2 data, but also the derived products that are generated in a standardized and automated way. In addition, the products are validated. Terrascope is currently the only source of such data sets for the Belgian users.

The service allows users to derive directly information from the S2 data on vegetation and crop condition.

Terrascope products are used for instance by WatchITgrow, a web application for potato monitoring in Belgium, which has been developed by VITO, CRA-W and ULg, in collaboration with Belgapom, the federation of the Belgian potato trade and processing industry.

Via WatchITgrow potato farmers, traders and processing companies have access to Sentinel-2A and 2B derived FAPAR maps and graphs processed by Terrascope.

- FAPAR maps are used to detect variability within the field or between fields. The causes may be diverse and can range from (natural) soil heterogeneity to climate induced problems such as drought or water logging, or local damages due to pests or diseases, emergence problems of seed potatoes, etc. Knowledge of variability within a field is helpful both for farmers, e.g. as input for variable rate applications of fertilizers, and for industry representatives, e.g. when taking yield samples. Industry users are also interested in variability between fields, especially towards the end of the season, as the fAPAR maps provide information on the maturity of the crop and whether haulm killing has been applied or not. This information is helpful for planning field visits and ultimately for harvest planning.
- The FAPAR graph of the field, showing the FAPAR evolution throughout the season, or the “crop growth curve”, provides useful information on crop development and phenology. Comparison of the growth curve of a field with growth curves of surrounding fields allows the farmer to benchmark his field.

1.5. Related documents

Table 1.3 lists the related documents (RD) that are complementary to this ATBD. Other Reference Documents (ORD) are listed in Section 7.

Table 1.3: List of related documents

[RD1]	Gatti, A., Galoppo, A. Castellani, C., Carriero, F. Sentinel-2 Products Specification Document, REF: S2-PDGS-TAS-DI-PSD issue 14.5,20/03/2018 https://sentinel.esa.int/documents/247904/685211/Sentinel-2-Products-Specification-Document
[RD2]	De Keukelaere, L., Gebruers, S., Swinnen, E., Wens, D., Caerts, S., Van den Heuvel, L., Van Baelen, S.. (2026). Terrascope Sentinel-2 Algorithm Theoretical Base Document (ATBD) S2 – TOC – V220.
[RD3]	Weiss, M., Baret, F. (2016). S2ToolBox Level 2 products: LAI, FAPAR, FCOVER, version 1.1, 02/05/2016. http://step.esa.int/docs/extra/ATBD_S2ToolBox_L2B_V1.1.pdf
[RD4]	Paepen, M., Wens, D., De Keukelaere, L., Swinnen, E., Clarijs, D. (2020) Terrascope Sentinel-2 Product User Manual V200.
[RD5]	Piccard, I., Gobin, A., Wellens, J., Goffart, J.-P., Curnel, Y., Planchon, V., Leclef, A., Cools, R., Cattoor, N. (2017). Potato monitoring in Belgium with “WatchITgrow”. 2017 9 th International Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp). https://doi.org/10.1109/Multi-Temp.2017.8035229

1.6. Definitions

The definitions of the biophysical variables are given below. These are derived from [RD3].

1.6.1. LAI

LAI is defined as half the developed area of photosynthetically active elements of the vegetation per unit horizontal ground area. It determines the size of the interface for exchange of energy (including radiation) and mass between the canopy and the atmosphere. This is an intrinsic canopy primary variable that should not depend on observation conditions. LAI is strongly non-linear related to reflectance. Therefore, its estimation from remote sensing observations will be strongly scale dependent (Garrigues et al. 2006a; Weiss et al. 2000). Note that vegetation LAI as estimated from remote sensing will include all the green contributors, i.e. including understory when existing under forests canopies. However, except when using directional observations (i.e. observed with a specific sun and viewing geometry) (Chen et al. 2005), LAI is not directly accessible from remote sensing observations due to the possible heterogeneity in leaf distribution within the canopy volume. Therefore, remote sensing observations are rather sensitive to the 'effective' leaf area index, i.e. the value that would produce the same remote sensing signal as that actually recorded, while assuming a random distribution of leaves. The difference between the actual LAI and effective LAI may be quantified by the clumping index (Chen et al. 2005) that roughly varies between 0.5 (very clumped canopies) and 1.0 (randomly distributed leaves). The LAI provided by Terrascope is actual LAI. Moreover, it is related to all green vegetation (not only leaves or needles), but excluding non-green parts (e.g. stems). This is sometimes referred to as GAI (Green Area Index).

1.6.2. FAPAR

FAPAR corresponds to the fraction of photosynthetically active radiation absorbed by the canopy. The FAPAR value results directly from the radiative transfer model in the canopy which is computed instantaneously. It depends on canopy structure, vegetation element optical properties and illumination conditions. FAPAR is very useful as input to a number of primary productivity models based on simple efficiency considerations (Prince 1991). Most of the primary productivity models using this efficiency concept are running at the daily time step. Consequently, the product definition should correspond to the daily integrated FAPAR value that can be approached by computation of the clear sky daily integrated FAPAR values as well as the FAPAR value computed for diffuse conditions. To improve the consistency with other FAPAR products that are sometimes considering the instantaneous FAPAR value at the time of the satellite overpass under clear sky conditions (e.g. MODIS), a study was proposed to investigate the differences between alternative FAPAR definitions (Baret et al. 2003). Results show that the instantaneous FAPAR value at 10:00 (or 14:00) local solar time is very close to the daily integrated value under clear sky conditions. FAPAR is relatively linearly related to reflectance values, and will be little sensitive to scaling issues. Note also that the FAPAR refers only to the green parts (leaf chlorophyll content higher than $15\mu\text{g}\cdot\text{cm}^{-2}$) of the canopy.

1.6.3. FCOVER

It corresponds to the gap fraction for nadir direction. FCOVER is used to separate vegetation and soil in energy balance processes, including temperature and evapotranspiration. It is computed from the leaf area index and other canopy structural variables and does not depend on variables such as the geometry of illumination as compared to FAPAR. For this reason, it is a very good candidate for the replacement of classical vegetation indices for the monitoring of green vegetation. Because of its quasi-linear relationship with reflectances, FCOVER will be only marginally scale dependent (Weiss et al. 2000). Note that similarly to LAI and FAPAR, only the green elements (leaf chlorophyll content higher than $15\mu\text{g}\cdot\text{cm}^{-2}\text{P}$) are considered.

1.6.4. CCC

The chlorophyll content is a very good indicator of stresses including nitrogen deficiencies. It is strongly related to leaf nitrogen content (Houlès et al. 2001). This quantity can be calculated both at the leaf level and at the canopy level by multiplication of the leaf level chlorophyll content by the leaf area index. In this case it is obviously an intrinsic secondary variable. Recent studies tend to prove that this product could be of very high interest in primary production models because it partly determines the photosynthetic efficiency (Green et al. 2003). In addition, studies have demonstrated that a direct estimation of CCC is more robust and accurate than an estimation based on the product of the individual estimation of LAI and C_{ab} (Weiss et al. 2000). Therefore, the estimation of CCC has been preferred to that of the leaf chlorophyll content.

1.6.5. CWC

Since radiation is absorbed significantly by water in the near and middle infrared, the spectral configuration of S2 allows accessing this variable. Water represents between 60 % and 80% of the living plant mass. The variable that is the best related to the remote sensing signal is defined as the mass of water per unit ground area ($\text{g}\cdot\text{m}^{-2}$). One of the difficulties in retrieving this variable is the possible confusion with soil moisture effects.

2. Input

2.1. Top-Of-Canopy reflectance

The NDVI and BIOPAR workflow starts from Level 2A Top-Of-Canopy (TOC) reflectance data. These data are directly downloaded from the Copernicus Data Space Ecosystem for near-real time processing. Archive data have been reprocessed to TOC products using the SEN2COR atmospheric correction algorithm, to be in line with ESA products [RD2].

The **S2 TOC Spectral Bands** span the range from the visible and Near Infra-Red to the Short Wave Infra-Red in different resolutions. The spatial and spectral characteristics are listed in Table 2.1.

Table 2.1: Spatial and spectral characteristics of the S2 TOC products. Bands used for NN-R8 in orange shading and for NN-R3 in blue characters.

Layer	Spatial resolution (m)	S2A		S2B	
		Central wavelength (nm)	Bandwidth (nm)	Central wavelength (nm)	Bandwidth (nm)
TOC-B01_60M	60	443.9	27	442.3	45
TOC-B02_10M	10	496.6	98	492.1	98
TOC-B03_10M	10	560.0	45	559.0	46
TOC-B04_10M	10	664.5	38	665.0	39
TOC-B05_20M	20	703.9	19	703.8	20
TOC-B06_20M	20	740.2	18	739.1	18
TOC-B07_20M	20	782.5	28	779.7	28
TOC-B08_10M	10	835.1	145	833.0	133
TOC-B8a_20M	20	864.8	33	864.0	32
TOC-B11_20M	20	1613.7	143	1610.4	141
TOC-B12_20M	20	2202.4	242	2185.7	238

The two Neural Networks addressed in the BIOPAR chain make use of following TOC band combinations:

- For the **NN-R8** eight bands are used (indicated in orange shading in Table 2.1): TOC-B03_10M, TOC-B04_10M, TOC-B05_20M, TOC-B06_20M, TOC-B07_20M, TOC-B8a_20M, TOC-B11_20M and TOC-B12_20M
- For the **NN-R3** three bands are used (indicated in blue characters in Table 2.1): TOC-B03_10M, TOC-B04_10M, TOC-B08_10M

2.2. Ancillary data and models

While the NDVI calculation only requires TOC reflectances as input, ancillary data and models are required for the calculation of the BIOPARs.

2.2.1. Neural Networks

The methodology to estimate the biophysical parameters is based on neural networks. Two sets of networks were trained, one set based on the input of 8 reflectance bands with output at 20 m resolution (NN-R8) and another set based on 3 reflectance bands with output at 10 m resolution (NN-R3). For each of the sets, there is a neural network trained independently for every biophysical variable.

The neural network coefficients are ancillary data in the Terrascope processing chain.

2.2.2. Angle information

The angle information, i.e. Solar Zenith Angle (SZA), View Zenith Angles (VZA) and Relative Azimuth Angles (RAA), are obtained from the S2 L1C metadata.

In the S2 L1C products, the angles are provided at 5000 m resolution by detector. The solar angles (SZA and Sun Azimuth Angle (SAA)) given in the metadata eXtensible Markup Language (XML) file are resampled to 10, 20 and 60m. However, the detector dependency hampers such a simple resampling for the viewing angles (VZA and View Azimuth Angle (VAA)): the detector footprints overlap at their edges, and in older formats of S2 it is not possible to determine exactly from which detector a pixel of the overlap area originates. On 06/11/2018 ESA deployed a new Production Baseline (02.07) which includes an accurate detector footprint. With this information it is possible to link one pixel to one detector..

3. Output

3.1. Product layers

For the BIOPAR and NDVI products, the following layers are generated:

- The actual parameter (NDVI or BIOPAR)
- The scene classification

The files are delivered together with an XML-file containing the metadata of the parameter. In addition, a quicklook is provided.

The NDVI is only delivered at 10 m resolution, whereas CCC and CWC are only available at 20 m resolution. LAI, FAPAR and FCOVER are available at both 10 m and 20 m resolution.

Table 3.1 provides the technical information of the NDVI and BIOPARs, like their physical range and the rescaling coefficients. These latter should be applied to the data to translate them to physical units, as the data are stored in BYTE. To rescale the BYTE output layers, the following formula has to be used:

$$\text{Physical value} = \text{DN} \cdot \text{slope} + \text{offset} \quad \text{Eq 1}$$

Table 3.1: Characteristics of the NDVI and BIOPAR images and rescaling information. Physical min and max are the physical range that is retained in the output, the Digital Numbers (DN) are the value of the physical min and max after rescaling to BYTE. The slope and offset are the coefficients to use to recompute the physical values from the BYTE output images using equation Eq 1.

	units	Physical min	Physical max	DN min	DN max	offset	slope	No data
NDVI	-	-0.08	0.92	0	250	-0.08	0.004	255
FAPAR	-	0	1	0	200	0	0.005	255
LAI	m ² /m ²	0	8	0	250	0	0.05	255
FCOVER	-	0	1	0	200	0	0.005	255
CCC	g/cm ²	0	600	0	250	0	0.05	255
CWC	µg/cm ²	0	0.55	0	250	0	0.026	255

The scene classification is copied from the TOC reflectance product as described in [RD2]. As mentioned before, the scene classification layer is always outputted in 20 m resolution Table 3.2 specifies the meaning of the pixel values in this layer.

Table 3.2: Meaning of the values in the scene classification layer.

Layer	Value	Meaning
Scene classification	0	No data
	1	Saturated or defected
	2	Cast shadows
	3	Cloud shadow
	4	Vegetation
	5	Bare soil
	6	Water
	7	Unclassified
	8	Cloud medium probability
	9	Cloud high probability
	10	Thin cirrus
	11	Snow

The folder structure used on the Terrascope platform is:

1. Product and version - *e.g. NDVI_V2*
2. Year - *e.g. 2019*
3. Month - *e.g. April*
4. Day - *e.g. 04*
5. Tile_ID - *e.g. S2B_20190504T105629_31UDS_FCOVER_V200*
6. Resolution (only if applicable) - *e.g. 10M*

3.2. Product version

Terrascope products are produced in a controlled way. Every product has a version indicator, consistent with the Semantic Versioning 2.0.0 protocols (<https://semver.org/>). The version indicator has three digits: XYZ.

- X is 0 during prototyping and pre-operational use. X is 1 for the first operational setup, and increments when if its results are no longer backward compatible (i.e. any further processing will have to be adapted to deal with e.g. format changes, value scaling, ...).
- Y is reset to 0 with an X increment. Y increments when functionality is added, but backward compatibility is guaranteed (e.g. when a different approach is taken for atmospheric or geometric correction).

- Z is reset to 0 when Y increments. Z increments when the software is patched (bug fixed) without any functional changes.

The current version of the Terrascope Sentinel2 workflow is v220.

Whenever X or Y changes, the impact of the updates will be reported and the new and previous versions of the workflow will be run in parallel, for a 3-4 month period. This allows users to implement changes to their subsequent processing. Users are informed about version changes through the Terrascope newsletter (to subscribe: <https://terrascope.be/en/stay-informed>).

3.3. Product data access

The Terrascope S2 data products can be accessed through:

- Terrascope viewer: <https://viewer.terrascope.be/en>
For viewing, discovery and data access. The viewer provides fast access to satellite data including Sentinel. You can easily search, view and compare various data layers. Via the 'Export' tab you can even download png images, GIF timelapses or the original data in just a few clicks. Would you like to implement your own processing? In that case, you can also directly retrieve the satellite data. You can do so through our data portal.
- Web services: Web Map Service (WMS) and Web Map Tile Service (WMTS): https://bit.ly/TerrascopeFAQ_WMTS
Protocols for downloading images and integrating them into GIS software
- Notebooks (login required): <https://notebooks.terrascope.be/hub/login>
Programming environment to quickly access and edit data
- Virtual Machines (VM) (login required): <https://forum.terrascope.be/en/request-vm>
External computer used to view data and process it in the cloud

The details of each of these access points are described on <https://terrascope.be/en/services>.

4. Methodology

4.1. NDVI (10 m resolution)

4.1.1. Justification

The NDVI is a dimensionless vegetation index, which provides information on the greenness of the surface. It is closely related to the Fraction of Absorbed Photosynthetically Active Radiation (FAPAR), and is little scale-dependent. Even though it is not a physical property of the vegetation cover, its simple formulation makes it widely used for ecosystem monitoring.

4.1.2. Implementation

The NDVI is calculated from two individual TOC reflectance measurements as follows:

$$NDVI = \frac{NIR - RED}{NIR + RED}$$

where NIR and RED are the spectral reflectances measured in the near infrared and red wavebands respectively. In Terrascope two 10 m bands are used, which results in following formula:

$$NDVI = \frac{B08_{TOC} - B04_{TOC}}{B08_{TOC} + B04_{TOC}}$$

The NDVI provided in Terrascope is a directional NDVI, as it is based on directional reflectances that are not normalized for viewing and illumination geometry.

The cloud and shadow masks are integrated in the images: if either of the masks has value 1 (i.e. cloud or shadow), then the NDVI value is replaced with the value 255, which is the 'No data' flag.

4.2. BIOPAR (10 m and 20 m resolution)

4.2.1. Justification

The biophysical parameters provide information of the land surface that is quantifiable with measurements, and independent of the sensor characteristics (spectral bands, illumination and observation geometry). It is therefore a more direct estimate of true land surface conditions.

As already mentioned before, the BIOPARs are generated at two spatial resolutions. The 20 m resolution BIOPARs use the method available in the S2 Toolbox which was designed by INRA-

EMMAH. This neural network is based on 8 spectral bands. The neural network (NN) to generate 10 m BIOPARs has been designed specifically for the WatchItGrow service (www.watchitgrow.be). The NN only uses three bands at 10 m resolution.

4.2.2. Method

4.2.2.1. BIOPARS at 20 m resolution (NN-R8)

The methodology for deriving Sentinel-2 BIOPAR products at 20 m resolution using the NN-R8 network is described in detail in [RD3]. This method is the same as is implemented in the SNAP Toolbox (*SNAP - ESA Sentinel Application Platform v2.0.2*, <http://step.esa.int>). A short summary is provided below.

The neural network approach to derive BIOPARS is established in two phases: (1) the definition of the neural network and its training (described here), and (2) the application of the neural network on the data (described in section 4.2.3). An independent network was established for each of the BIOPAR variables.

To test and train a neural network, a large database of input and output data is required. The input data are the reflectances of the 8 spectral bands (see Table 2.1), and the angular configuration. The output data are the BIOPAR variables LAI, FAPAR, FCOVER, CCC and CWC. A large database of input and corresponding output data is simulated using Radiative Transfer Models (RTMs). These RTMs are used to simulate the spectra (surface reflectance) given a set of input parameters describing the leaf optical properties (e.g. chlorophyll), the canopy properties (e.g. LAI), the background properties (e.g. background soil reflectance), and the sun and observational geometry. The simulation database was created using the RTM PROSAIL (Jacquemoud et al., 2009) that consists of PROSPECT v5 (Jacquemoud et al., 1996) and 4SAIL (Verhoef and Bach, 2003). In order to generate a globally representative database of all possible vegetation conditions, the parameterisation of the PROSAIL is based on literature review of all in the input parameters used. The details are described in [RD3].

The neural network is defined by the type of neurons used (the transfer function), the way they are organised and connected (the architecture) and the learning rule. The optimal architecture has been defined by testing several neural networks and compare their outputs with the simulated 'true' values. In addition, networks with less coefficients were preferred.

After the training of NN-R8, its theoretical performance was tested on a test dataset, which is one third of the total simulation database that is not used for training the network. [RD3] reports RMSE values of 0.89 for LAI, 0.05 for FAPAR, 0.4 for FCOVER, 56 $\mu\text{g}/\text{cm}^2$ for CCC and 0.03 g/cm^2 for CWC which demonstrate a good performance of the network. FAPAR and FCOVER show the best performance, with higher RMSE values for mid-range values of the product. LAI is well estimated up to values of LAI=6, and increasing uncertainties with LAI, and thus also CCC and CWC because of their dependency on LAI, are observed. Furthermore, the networks are unbiased between the BIOPAR variables as expected.

An exhaustive description of the methodology, the set up and evaluation is given in [RD3].

4.2.2.2. BIOPARS at 10 m resolution (NN-R3)

In order to fulfil the needs of having the same BIOPAR variables at 10 m spatial resolution, INRA-EMMAH provided also the coefficients of the neural networks based on the three 10 m spectral bands. In order to establish these neural networks, the same method as described in the previous section was used. However, no separate documentation was delivered for this network. Hence, the theoretical performance is unknown.

4.2.3. Implementation

The implementation of both the NN-R8 and NN-R3 networks is done in the same way. The processing consists of the steps explained below.

For each pixel in each image the following procedure is followed

- (1) The status map is used to check if the pixel is of good quality. If not, the pixel is labelled in the output product as 255. If yes, the pixel is further processed
- (2) The input reflectance values are read and normalized according to the specifications of [RD3]. The angular information is read and the cosine of the view and sun zenith angle and the relative azimuth is calculated. Note that only one set of angles are used per scene (see section 2.2.2 for explanation).
- (3) The neural network is run.
- (4) The output is denormalized according to the specifications of [RD3].

Unlike the description in [RD3], the quality indicator layer is not generated, hence there is no information available on the quality of the retrieval process.

This concerns mainly two tests that are omitted: (1) check if the input values are within a specified range (min-max), and (2) check if the output values are within the definition domain. The valid range is both based on the values present in the training database and the experience and knowledge of the authors of the neural networks.

The tests were not implemented, because many pixels were then flagged. A user can evaluate if the output values are outside the definition domain by checking if the output values are within the range specified in Table 4.1.

Table 4.1: Valid range of output values

	min	max
LAI	0	8
FAPAR	0	0.94
FVC	0	1
CCC	0	600
CWC	0	0.55

5. Limitations

5.1. Algorithm limitations

The algorithm is 'generic', i.e. it should apply to any type of vegetation with reasonable performances. However, to better match the specifications of given canopies, either a simple correction could be calibrated, or a more specific algorithm could be developed.

One strong assumption embedded in any single pixel retrieval algorithm as this one, is that the pixel targeted belongs to a landscape patch presenting enough homogeneity (at the pixel scale) preventing unexpected loss or gain of radiation fluxes. Therefore, it can be applied for larger resolution than 20 m. For forests with large crowns, or any pixel showing strong heterogeneity such as pixels at the intersection between two different vegetation patches, results may be uncertain. This extends also to pixels where the neighbouring ones are very different. Specific algorithms should be developed to detect such situations and possibly propose alternative retrieval methods [RD3].

The technical performance of the NN-R8 described in [RD3] reports that the LAI retrievals are accurately retrieved up to values of 6. The LAI retrievals above 6 have larger uncertainty.

The technical performance of the NN-R3 is not documented, nor the results of the establishment of the neural network. Although this was done in line with the NN-R8 approach, any differences are currently unknown.

5.2. Implementation limitations

Pixels classified in the SCENECLASSIFICATION layer as cloud, cloud shadow, snow, cirrus or saturated pixels are masked in the NDVI and BIOPAR images, hence these pixels are set to 'NoData' and the NDVI or BIOPAR value is not available. This hampers the use of another user-defined cloud or cloud shadow screening.

5.3. Validation of the products

At present, validation of the products has been performed at different levels of the processing chain. These were however more ad hoc analyses and are not yet available in a validation report. The products are currently being validated in a more comprehensive analysis.

Due to the lack of in situ validation data for CCC and CWC, these products have not been validated with in situ data.

6. Quality assessment

This section reports on the quality assessment of the products prior to the distribution to Terrascope. It consists of ad hoc analysis to verify the quality of the products, and a comparison between the previous (v102) and the current (v200) Terrascope product version.

The implementation of the processing chain with a test data set from INRA to verify the processing chain's output was done and the same results were obtained. This is not documented here.

6.1. Terrascope Sentinel-2 v200 vs V102

6.1.1. SCENECLASSIFICATION layer

A new version of Sen2COR (v2.8) is implemented with support of the CCI package. Figure 6.1 and Table 6.1 show the differences in identified classes between v102 and v200. Figure 6.1 show the fraction of each category for a full year of tile 31UFS. v200 detects more 'Cirrus', 'Snow' and 'Clouds medium probability', while the total number of pixels identified as 'Cloud high probability' is lower. The total number of 'vegetated' pixels remains stable. Table 6.1 is a confusion matrix, showing the percentage of pixels of different V102 classes belonging to one V200 class. The percentages of the columns sums to 100%. Of all pixels classified in v200 as 'Cloud medium probability', only 61% belonged in v102 in the same category, while 38% of these pixels were classified as 'Cloud high probability' in v102. Most changes occurred in the category 'Snow': from the pixels classified as 'Snow' in V200, 70% were in V102 classified as 'Cloud high probability', 15% as 'Cloud medium probability', only 9% also as 'Snow', and another 4% as 'Unclassified'. Also the increase in 'Cirrus' percentage in v200 is not only linked to one specific category shift. It comprises of pixels classified in v102 as 'Water', 'Cloud high probability', 'Vegetation' and 'Unclassified'.

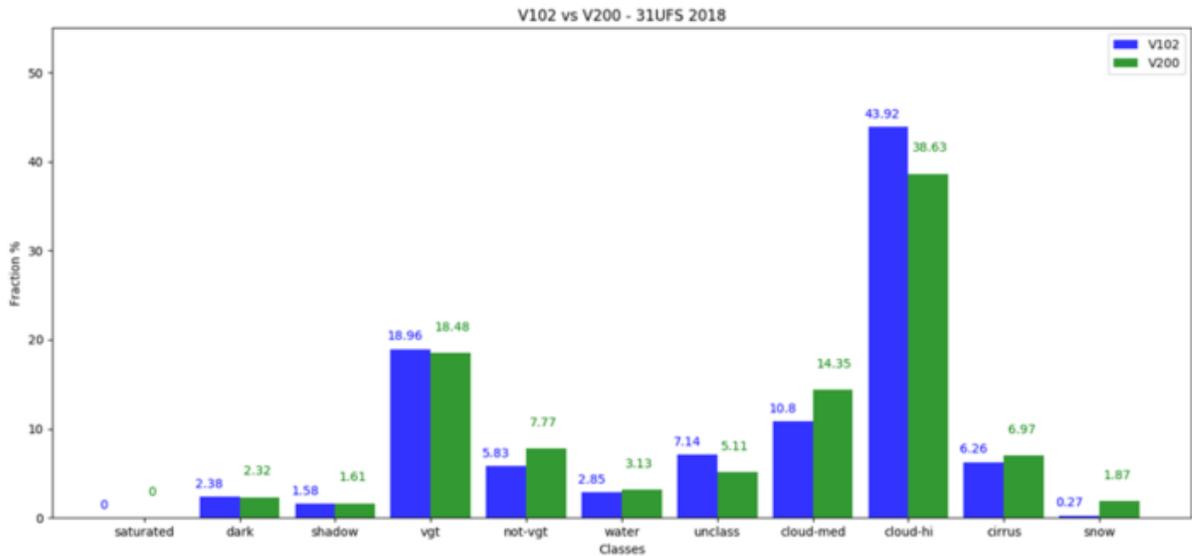


Figure 6.1 Fraction [%] of each category in the year 2018 for tile 31UFS. The blue bars represent the results for v102, green bars for v200.

Table 6.1 Confusion matrix of the Sceneclassification layer between V102 and V200. The numbers are expressed as percentages.

		V200										
		No data	Dark area pixels	Cloud shadow	Vegetation	Bare soil	Water	Unclassified	Cloud medium probability	Cloud high probability	Thin cirrus	Snow
V102	No data	100.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Dark area pixels	0.00	84.99	0.01	0.00	0.00	15.63	0.00	0.00	0.00	0.17	0.00
	Cloud shadow	0.00	0.21	99.98	0.00	0.00	13.17	0.00	0.00	0.00	0.14	0.00
	Vegetation	0.00	0.00	0.00	99.98	0.00	0.00	0.00	0.00	0.00	2.37	0.00
	Bare soil	0.00	0.00	0.00	0.00	82.21	0.00	0.00	0.00	0.00	1.32	0.00

Water	0.00	0.07	0.00	0.00	0.00	62.06	0.00	0.00	0.00	0.32	0.30
Unclassified	0.00	3.41	0.00	0.00	13.75	8.32	99.88	0.00	0.00	0.30	4.23
Cloud medium probability	0.00	8.77	0.00	0.00	3.35	0.16	0.01	61.43	0.00	0.56	15.48
Cloud high probability	0.00	0.92	0.00	0.00	0.66	0.56	0.04	38.55	100.00	4.91	70.13
Thin cirrus	0.00	0.02	0.01	0.01	0.02	0.00	0.07	0.02	0.00	89.89	0.01
Snow	0.00	1.62	0.00	0.00	0.00	0.09	0.00	0.00	0.00	0.02	9.86

6.1.2. BIOPAR and NDVI comparison

The impact of another atmospheric correction model, as well as inclusion of sun and viewing angle has an impact of the retrieved BIOPAR and NDVI products. This section covers a high level intercomparison. More in-depth results will be provided in a separate validation report.

6.1.2.1. FAPAR

Figure 6.2 and Figure 6.3 show the frequency plots and bias histograms for the FAPARS 20 m and 10 m products respectively between v102 and v200. In the frequency plots, v200 is marked as green, while v102 is marked as red. For both the 20 m and 10m products, more low FCOVER values are retrieved, but the general shape remains the same. The bias histograms show that a majority of the bias between two versions is located between -0.1 and 0.05. While the 20m products show mainly a negative bias, i.e. small shift towards lower values, the 10m products shift towards both directions.

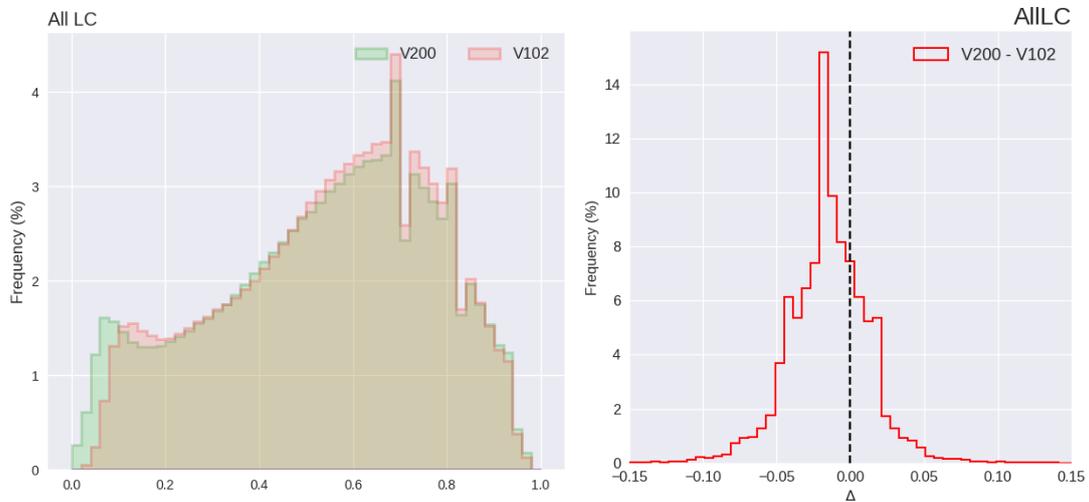


Figure 6.2 Frequency plot (left) and bias histograms (right) of FAPAR 20 m products for v102 and v200 for the full time series of tile 31UFS (i.e. starting in 2015).

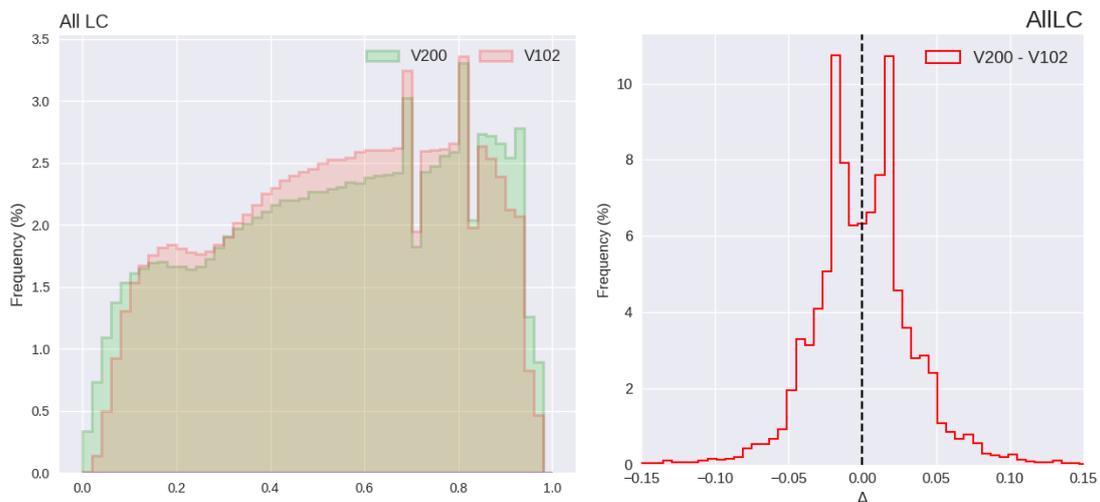


Figure 6.3 Frequency plot (left) and bias histograms (right) of FAPAR 10 m products for v102 and v200 for the full time series of tile 31UFS (i.e. starting in 2015).

6.1.2.2. FCOVER

The next set of figures (Figure 6.4 and Figure 6.5) show the frequency plots and bias histograms for the FCOVER products respectively between v102 and v200. Similar trends as for FAPAR are observed: v200 allows lower FCOVER values for both 10 and 20m products, which relates to a slightly negative

shift. For the 10M products, a second peak in the bias histogram is observed resulting in a small increase in high FCOVER values as well. The range in bias values is mainly located between -0.1 and 0.1.

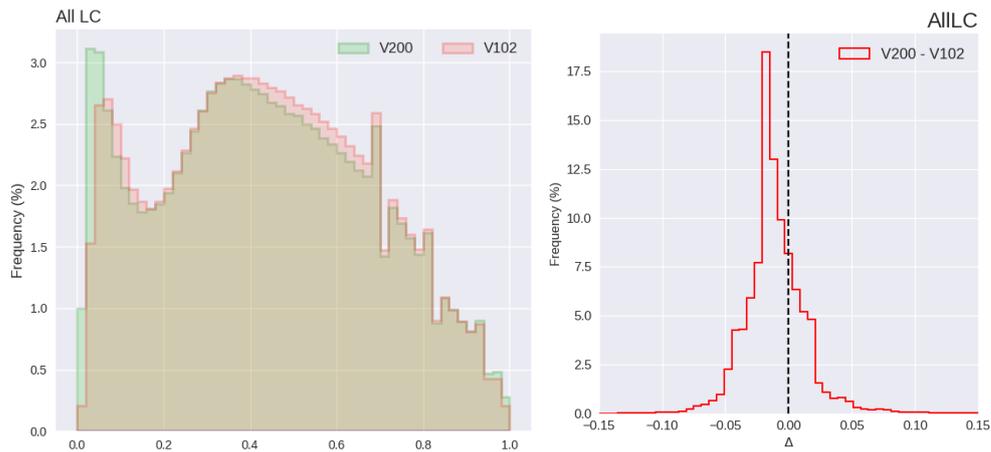


Figure 6.4 Frequency plot (left) and bias histograms (right) of FCOVER 20 m products for v102 and v200 for the full time series of tile 31UFS (i.e. starting in 2015).

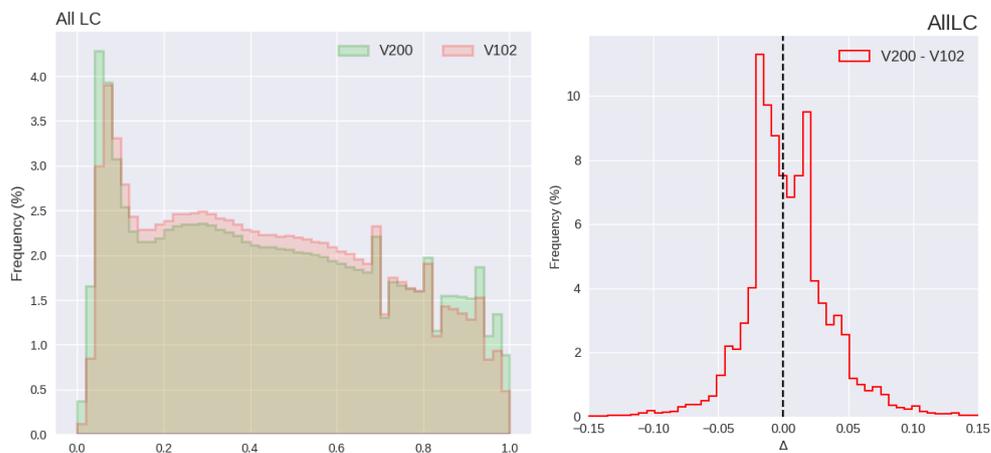


Figure 6.5 Frequency plot (left) and bias histograms (right) of FCOVER 10 m products for v102 and v200 for the full time series of tile 31UFS (i.e. starting in 2015).

6.1.2.3. LAI

The frequency plots and bias histograms between v102 and v200 for LAI are depicted in Figure 6.6 and Figure 6.7. The peak in bias values is centred around zero for 20 m as well as 10 m products. This is also visible in the frequency plots.

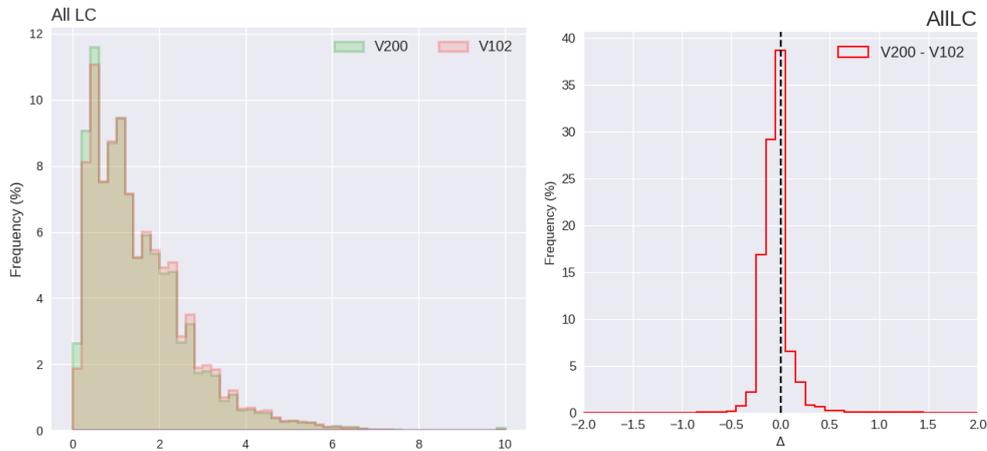


Figure 6.6 Frequency plot (left) and bias histograms (right) of LAI 20 m products for v102 and v200 for the full time series of tile 31UFS (i.e. starting in 2015).

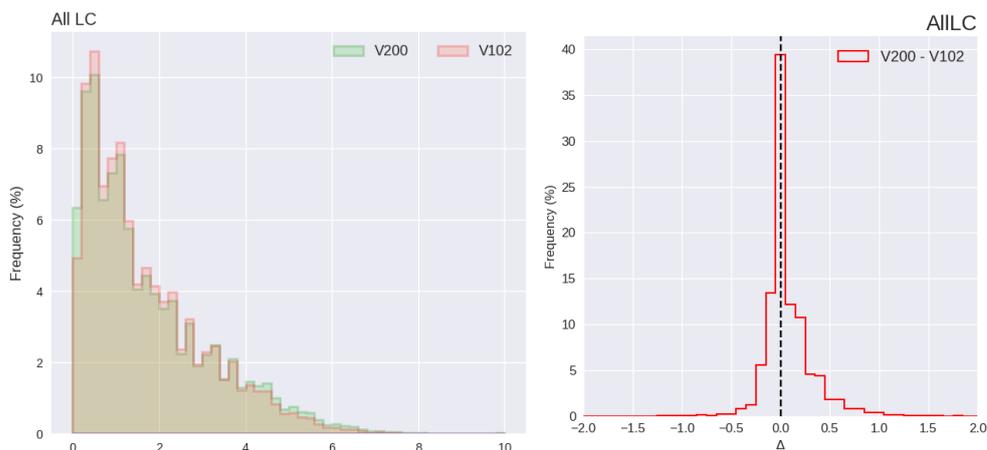


Figure 6.7 Frequency plot (left) and bias histograms (right) of LAI 10 m products for v102 and v200 for the full time series of tile 31UFS (i.e. starting in 2015).

6.1.2.4. NDVI

Figure 6.8 shows the frequency plot and bias histogram of the NDVI 10m products. The bias is centered around zero and ranges predominantly from -0.1 to 0.1.

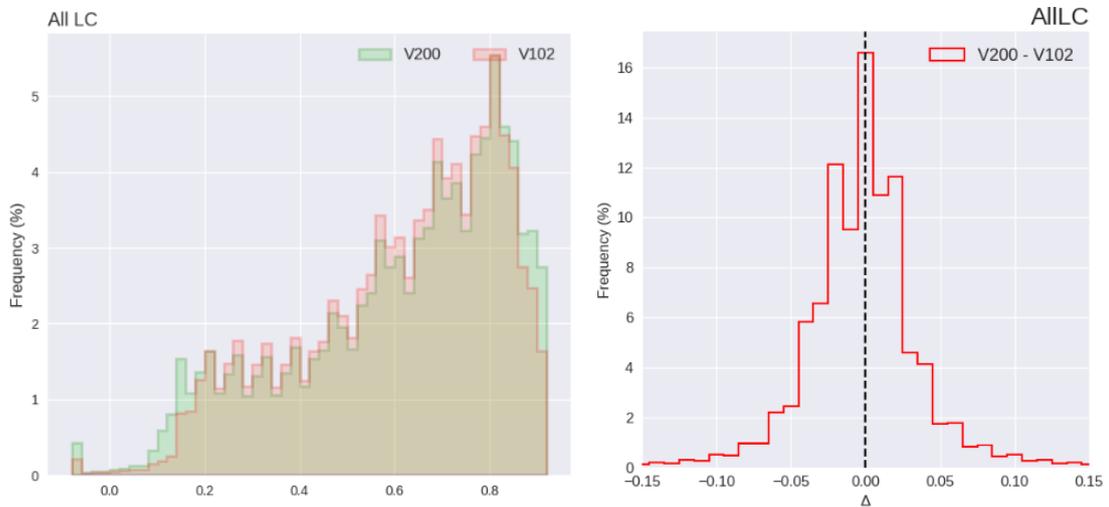


Figure 6.8 Frequency plot (left) and bias histograms (right) of NDVI 10 m products for v102 and v200 for the full time series of tile 31UFS (i.e. starting in 2015).

7. Other Reference documents

Baret, F., Leroy, M., Roujean, J.L., Knorr, W., Lambin, E., & Linderman, M. (2003). CYCLOPES User Requirement Document. In. Avignon: INRA-CSE

Chen, J.M., Menges, C.H., & Leblanc, S.G. (2005). Global mapping of foliage clumping index using multi-angular satellite data. *Remote Sensing of Environment*, 97, 447-457

Garrigues, S., Allard, D., Baret, F., & Weiss, M. (2006a). Influence landscape spatial heterogeneity on the non-linear estimation of leaf area index from moderate spatial resolution remote sensing data. *Remote Sensing of Environment*, 105, 286-298

Green, D.S., Erickson, J.E., & Kruger, E. (2003). Foliar morphology and canopy nitrogen as predictors of light use efficiency in terrestrial vegetation. *Agricultural and Forest Meteorology*, 115, 163-171

Houlès, V., Mary, B., Machet, J.M., Guérif, M., & Moulin, S. (2001). Do crop characteristics available from remote sensing allow to determine crop nitrogen status? In G. Grenier, & S. Blackmore (Eds.), 3rd European Conference on Precision Agriculture (pp. 917-922). Montpellier: Agro Montpellier
http://step.esa.int/thirdparties/sen2cor/2.5.5/docs/S2-PDGS-MPC-L2A-SUM-V2.5.5_V2.pdf

Jacquemoud S., Ustin S.L., Verdebout J., Schmuck G., Andreoli G., & Hosgood B. (1996), Estimating leaf biochemistry using the PROSPECT leaf optical properties model. *Remote Sensing of Environment*, 56, 194-202.

Jacquemoud S., Verhoef W., Baret F., Bacour C., Zarco-Tejada P.J., Asner G.P., François C., & Ustin S.L. (2009). PROSPECT + SAIL models: a review of use for vegetation characterization. *Remote Sensing of Environment*, 113, S56-S66.

Prince, S.D. (1991). A model of regional primary production for use with coarse resolution satellite data. *International Journal of Remote Sensing*, vol.21, issue 6, p. 1313-1330

Verhoef W., & Bach H. (2003), Simulation of hyperspectral and directional radiance images using coupled biophysical and atmospheric radiative transfer models. *Remote Sensing of Environment*, 87, 23-41.

Weiss, M., Baret, F., Myneni, R., Pragnère, A., & Knyazikhin, Y. (2000). Investigation of a model inversion technique for the estimation of crop characteristics from spectral and directional reflectance data. *Agronomie*, 20, 3-22