Distribution: General

Final report

Evaluation of PROBA-V C2 products Final Report

Carolien Toté, Else Swinnen, Lieve Van den Heuvel, Dennis Clarijs

Study accomplished under the authority of BELSPO

2023/TAP/R/

v1.01, March 2024

VITO NV
Boeretang 200 - 2400 MOL - BELGIE
Tel. + 32 14 33 55 11 - Fax + 32 14 33 55 99
vito@vito.be - www.vito.be

BTW BE-0244.195.916 RPR (Turnhout)
Bank 375-1117354-90 ING
BE34 3751 1173 5490 - BBRUBEBB

SUMMARY

This report presents the results of the 'Phase 3 complete validation' of the reprocessing campaign on the Proba-V archive. The second reprocessing campaign aimed at improving the time series and harmonizing its content. The resulting archive is Proba-V Collection 2 (C2). The main algorithm modifications in the Proba-V processing chain from Collection 1 (C1) to C2 are related to: (1) updated radiometric correction, (2) a new and better cloud detection method and improved cloud shadow detections, (3) an improved atmospheric correction, (4) harmonisation of the compositing among the resolutions, and (5) updated geomodelling. A detailed description of the algorithm updates is provided in the PV C2 Algorithm Change Document.

In the Phase 3 validation we analysed almost 7 years of daily (S1), 5-daily (S5) and 10-daily composites (S10) data at global scale and included consistency checks with SPOT/VGT and comparison with external data sets from Sentinel-3/SYN, METOP/AVHRR (LSA-SAF/ENDVI10) and Terra/MODIS (MOD13A3). Data was subsampled by taking a systematic spatial subsample and by extracting information over the LANDVAL sites.

The Phase 3 validation of Proba-V C2 products is based on three pillars: (1) Comparison of the reprocessed archive (C2) with the previous archive (C1) over the entire operational phase of Proba-V; (2) Comparison of Proba-V C2 with related datasets derived from SPOT/Vegetation and Sentinel-3/SYN-VGT; and (3) Comparison to reference time series from external datasets derived from MetOp/AVHRR and Terra/MODIS.

In comparison to Proba-V C1, C2 shows more clear observations, and a general increase of 'good' pixels, i.e. clear pixels with good radiometric quality, at all resolutions. Differences between both collections at Top-of-Atmosphere (TOA) level are very small, and in line with radiometric calibration updates. Larger bias is observed at Top-of-Canopy (TOC) level for Blue and Red bands, related to adaptations to the atmospheric correction. Especially over densely vegetated areas, in C2 lower NDVI is observed.

Overall, there is high correspondence between Proba-V and SPOT/VGT. Nevertheless, C1 shows slightly better consistency with SPOT/VGT, especially for the Blue band. In contrast, there is high systematic bias between Proba-V and S3/SYN-VGT, related to a number of important quality issues in the current S3/SYN-VGT products.

Proba-V C2 NDVI shows strong correspondence with LSA-SAF ENDVI10 – albeit with a small systematic bias – and MOD13A3 NDVI.

The combined NDVI series of SPOT/VGT-C3 (2009-2013), Proba-V C2 (2014-June/2020) and S3A/SYN-VGT (July/2020-2022) shows a strong discontinuity at the switch to S3/SYN-VGT, leading to high bias with LSA-SAF ENDVI10 and MOD13A3 NDVI for the S3/SYN-VGT period. The temporal profiles over LANDVAL sites show strong temporal consistency between Proba-V C2 and C1, SPOT/VGT-C3, ENDVI10 and MOD13A3 NDVI, and large inconsistencies with S3/SYN-VGT NDVI.

In summary, the Proba-V reprocessing campaign was successful, yielding the expected impacts in terms of product completeness, and differences with the previous products. Proba-V C2 products show large consistency with the SPOT/VGT-C3 data archive and external datasets (except S3/SYN-VGT). Since the current S3/SYN-VGT products still suffer from important quality issues, users are advised not to use these products in combination with Proba-V or SPOT/Vegetation products.

Users are strongly recommended to update their Proba-V archive with Collection 2 (.V2 in the file naming).

TABLE OF CONTENTS

proba-v

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

A	Accuracy
Acc	Accuracy
AOT	Aerosol Optical Thickness
AVHRR	Advanced Very High Resolution Radiometer
BA	Bare areas
BRDF	Bidirectional Reflectance Distribution Function
C0	Collection 0
C1	Collection 1
CCI	Climate Change Initiative
CF	Climate and Forecasting
CUL	Cultivated areas and cropland
DBF	Deciduous broadleaf forests
DN	Digital Number
EBF	Evergreen broadleaf forests
ENDVI	EPS Normalized Difference Vegetation Index
EPS	EUMETSAT Polar System
ESA	European Space Agency
EUMETSAT	EUropean organization for the exploitation of METeorological SATellites
GMR	Geometric Mean Regression
HER	Herbaceous cover
ICP	Instrument Calibration Parameters
LSA-SAF	Land Surface Analysis – Satellite Applications Facility
LTS	Long Term Statistics
MBE	Mean Bias Error
MODIS	MODerate resolution Imaging Spectroradiometer
MVC	Maximum Value Composite
NDVI	Normalized Difference Vegetation Index
NIR	Near-infrared
NLF	Needleleaf forests
OTH	Other (land cover)
P	Precision
Prec	Precision
PROBA	PRoject for OnBoard Autonomy
PUM	Product User Manual
PV	Proba-V
RMSD	Root Mean Squared Error
S3	Sentinel-3
SHR	Shrubland
SM	Status map
SPOT	Satellite Pour l'Observation de la Terre
SWIR	Shortwave infrared
TOA	Top-of-Atmosphere
TOC	Top-of-Canopy
U	Uncertainty
Unc	Uncertainty
VAA	Viewing Azimuth Angle
VGT	Vegetation instrument on-board SPOT satellite
VNIR	Visual and near-infrared
VZA	Viewing Zenith Angle

CHAPTER 1 INTRODUCTION

1.1. INTRODUCTION

In 2020-2022, the second reprocessing campaign of the Proba-V (PV) archive was performed, aiming at improving the time series and harmonizing its content. The resulting archive is Proba-V Collection 2 (C2). The main modifications in the Proba-V processing chain from Collection 1 (C1) to C2 are related to:

- updates on the radiometric instrument calibration parameter (ICP) files;
- an improved cloud detection algorithm;
- an improved atmospheric correction scheme;

A detailed description of the algorithm updates is provided in the PV C2 Algorithm Change Document. This report focuses on the evaluation of the effect of the reprocessing on Top-of-Atmosphere (TOA) reflectances, Top-of-Canopy (TOC) reflectances and the Normalized Difference Vegetation Index (NDVI) over the entire time series (October/2013 – June/2020).

Important note:

The Proba-V Collection 2 archive is affected by artefacts where very high AOT values occur. When this occurs, the Top-Of-Canopy reflectance values are high. The impact is the largest in the Blue (B0), lower in Red and NIR (B2 and B3) and almost negligible in the SWIR. The Status Map (SM) of the products does not allow to remove these areas. This was not identified during the validation as the status map of PV C1 already masked out these values and intercomparison is done on those pixels where both collections have good observations.

An additional mask, the AOT Mask (AM), is provided to enable filtering out these high AOT areas. The AOT Mask contains several values to allow users to define their own filtering thresholds. The impact of the different thresholds was analysed and is summarized in a technical note (TN AOT masks). Recommended thresholds are formulated in this document as well.

1.2. SCOPE AND OBJECTIVES

The findings presented in this report are the result of the so-called 'Phase 3 - Complete validation'. The objective is to have a complete validation of the updates to the processing chain compared to Phase 1 (internal validation, in which the implementation of the changes were verified based on qualitative and visual checks of segments and S1 composites) and Phase 2 (validation over 12 months).

The Phase 3 validation is based on analysis almost 7 years of daily (S1), 5-daily (S5) and 10-daily composites (S10) data at global scale and includes consistency checks with SPOT/VGT and comparison with external data sets derived from Sentinel-3/SYN, Terra/MODIS and MetOp/AVHRR.

Separate validation was performed for the atmospheric correction and the pixel classification scheme. Validation reports are available, see §1.4.

1.3. Content of the document

This document is structured as follows:

- Chapter 2 describes the expected impact of changes in the Proba-V processing chain.
- Chapter 3 summarizes the evaluation methods.
- Chapter 4 lists the data sets used in the evaluation.
- Chapter 5 focuses on the comparison between Proba-V C2 and C1.
- Chapter 6 focuses on the consistency with SPOT/VGT and Sentinel-3 SYN-VGT.
- Chapter 7 gives the comparison between Proba-V C2 NDVI with external data.
- Chapter 8 lists the conclusions.

1.4. Related sources of documentation

Table 1: Reference documentation for Proba-V collection 2

Document ID	Document and link
PV C2 Algorithm Change Document	Proba-V C2 Algorithm Change Document https://proba-v.vgt.vito.be/sites/probavvgt/files/downloads/PROBA-V_C2_Algorithm_Change_Document.pdf
PV C2 Product User Manual	Proba-V C2 Product User Manual https://proba-v.vgt.vito.be/sites/probavvgt/files/downloads/PROBA-V_C2_Products_User_Manual.pdf
VR AC	Validation report of the Atmospheric Correction of Proba-V C2 https://proba-v.vgt.vito.be/sites/probavvgt/files/downloads/PROBA-V_C2_Atmospheric_Correction_Validation_Report.pdf
VR PC 1 km	Validation report of the Pixel Classification of Proba-V C2 at 1 km https://proba-v.vgt.vito.be/sites/probavvgt/files/downloads/PROBA-V_C2_Cloud_Mask_1km_Validation_Report.pdf
VR PC 300 m	Validation report of the Pixel Classification of Proba-V C2 at 300 m https://proba-v.vgt.vito.be/sites/probavvgt/files/downloads/PROBA-V_C2_Cloud_Mask_300m_Validation_Report.pdf
VR PC 100 m	Validation report of the Pixel Classification of Proba-V C2 at 100 m https://proba-v.vgt.vito.be/sites/probavvgt/files/downloads/PROBA-V_C2_Cloud_Mask_100m_Validation_Report.pdf
TN AOT masks	Technical note: PROBA-V C2 AOT masks https://proba-v.vgt.vito.be/probavvgt/files/downloads/PROBA-V_C2_TN_AOT_masks.pdf

CHAPTER 2 EXPECTED IMPACT OF COLLECTION 2 ALGORITHM UPDATES

2.1. INTRODUCTION

This chapter describes the expected impact of changes that were implemented in the Proba-V reprocessing campaign. These differences are described more in detail in the PV C2 Algorithm Change Document.

Compared to Collection 1, the following changes have an impact on the product content: (1) updated radiometric correction, (2) a new and better cloud detection method and improved cloud shadow detections, (3) an improved atmospheric correction, (4) harmonisation of the compositing among the resolutions, and (5) updated geomodelling.

2.2. UPDATED RADIOMETRIC CORRECTION

For the reprocessing of Proba-V C2, some changes were made to the radiometric Instrument Calibration Parameter (ICP) files.

The aim of the adjusted absolute calibration is to better characterize the conversion of digital counts measured by the instruments into reflectance. In C2 the change in radiometric responsivity is modelled for all bands and strips using a 2nd order degradation model. In addition, a small bias correction was applied on BLUE (left camera) and SWIR (right camera). When the absolute calibration is higher in C2, TOA reflectances will be lower, and vice versa. On average, differences of absolute calibration range between -1.3% and +1.8%. These changes in the absolute calibration will have small systematic impacts on the TOA reflectances, that vary over time.

In order to correct for brightness variations over the field-of-view in the SWIR strips (3 per camera), in C2 improved equalization coefficients were implemented for all cameras. This will have small unsystematic impacts on the SWIR TOA reflectances, leading to more homogenized images in the SWIR channel.

In C1, minor non-linearity corrections for VNIR were applied for the C1 reprocessing (from the start of the mission until November 2016), but not for the C1 operational processing (from November 2016 till end of mission). This is solved in C2.

2.3. IMPROVED CLOUD DETECTION ALGORITHM

The Proba-V C2 cloud detection algorithm uses a Multi-Layer Perceptron (MLP) neural network algorithm, without dependency on auxiliary input data. A single global model per resolution was established and validated. Final performance is greatly improved compared to both Collection 0 and Collection 1.

Validation showed that the new cloud detection method has a very good performance (see Validation Reports). The issues found in C1 (a.o. the quality issue with incorrect land cover ancillary data in the northern hemisphere) are largely solved and a good separation between cloud and snow/ice is found. Some trade-offs were made, resulting in some overestimation of clouds over bright surfaces, such as salt lakes, urban areas and turbid waters (but less than in C1), 50% of the thin semi-transparent clouds are detected, and sparse snow or melting ice is often not detected.

The cloud shadow detection was improved by removing the 1-pixel border between cloud and cloud shadow, which was present in the previous collections.

2.4. IMPROVED ATMOSPHERIC CORRECTION

As in Collection 1, the new atmospheric correction is also based on the Simplified Model for Atmospheric Correction (SMAC) (Rahman and Dedieu, 1994). C2 uses an external dataset, namely the MERRA-2 (Modern-Era Retrospective analysis for Research and Applications, version 2) for the inputs of the atmospheric correction (Gelaro et al., 2017). Top-Of-Canopy (TOC) reflectances are therefore better characterized in C2 than in C1. In addition, validation showed that artefacts due to the image based Aerosol Optical Thickness (AOT) retrieval in Collection 1 – which was done every 8x8 pixels and then interpolated – are removed.

The impact of the improved atmospheric correction will be mostly unsystematic, mainly caused by a more reliable AOT estimate compared to the image retrieved AOT in C1, although there might be systematic effects with seasonal and spatial patterns. Systematic impacts will be more pronounced in areas with high AOT, i.e. the semi-arid regions of North Africa, the Arabian Peninsula and the central-south Asian deserts, the Indian subcontinent, the African and South American tropics (Korras-Carraca et al., 2021). Seasonal variations of the impact are caused by seasonal variations in AOT concentrations, e.g. due to biomass burning in the tropics.

Occasionally, the MERRA-2 dataset contains very high AOT values, leading to unreliable atmospheric correction; SMAC is inaccurate for AOT above 0.8 (Rahman and Dedieu, 1994). Where the atmospheric correction results in out of range TOC reflectances (<0 or >1) due to the very high AOT inputs, these pixels are flagged in the status map as 'bad radiometric quality'. However, edge effects around these areas might occur, with subsequent spatial artefacts in downstream products.

2.5. UPDATE AND HARMONIZATION OF THE COMPOSITING METHOD

The compositing method is harmonized between the different resolutions in Collection 2. Previously, for 100 m and 300 m the radiometric quality of all 4 bands were checked prior to compositing. Since the SWIR band has quite a number of defect detectors, this resulted in composites with a striping effect: cloudy observations with good SWIR quality were preferred over clear pixels with bad SWIR quality. For the 1 km, the SWIR radiometric quality was not checked in the compositing process. This method is now applied to all resolutions. This adaptation will result in less cloudy (but more bad SWIR quality) observations in the 300 m and 100 m composite products.

In addition, in the previous collections, if all bands had a radiometric quality 'bad', then the pixels were set to 'undefined' in the compositing step. This rule is omitted in the Proba-V C2 processing. If this situation occurs, the pixels will now have radiometric quality 'bad'. The 'undefined' flag is reserved for situations when one of the processing steps (atmospheric correction or cloud detection) could not succeed, e.g. because one of the input bands is missing.

2.6. UPDATED GEOMODELLING

In C2 the geometric quality checks were slightly relaxed. This will have a positive effect on product completeness, since less lines will be marked as 'bad'. Another consequence is that, if the first or last scanlines in a segment are included in the C2 processing where lines were omitted in C1, a slightly different intermediate projection will be defined, and slight differences in geolocation can be expected. The overall geolocation accuracy is not impacted.

CHAPTER 3 EVALUATION METHODS

3.1. GENERAL APPROACH

The Phase 3 validation of Proba-V C2 is based on three pillars:

1. **Comparison of the reprocessed archive (C2) with the previous archive (C1) over the entire operational phase of Proba-V**: October/2013 – June/2020. Analyses comprise TOA surface reflectances, TOC reflectances and Normalized Difference Vegetation Index (NDVI), at 1 km, 300 m and 100 m spatial resolution. The evaluation is done at global scale, on a systematic subsample (see §3.3).

   The key questions that are answered are:

   Q1.1. What is the completeness of the data, in terms of spatial pattern and temporal evolution? What is the difference in flag occurrences between Proba-V C1 and C2?

   Q1.2. What is the magnitude of the difference between Proba-V C2 and C1? What is the spatial and temporal pattern of the difference?

2. **Comparison of Proba-V C2 with related datasets derived from SPOT/Vegetation and Sentinel-3/SYN-VGT**. The intercomparison of Proba-V with SPOT/Vegetation is done for the overlap period (November/2013 – May/2014). In addition, intercomparison is done between the Long Term Statistics (LTS) of Proba-V (2014-2018) and SPOT/Vegetation (2009-2013). The LTS of Proba-V is compared with Sentinel-3 (S3) SYN-VGT products of 2022. Analyses focus on S10-TOC at 1 km resolution. The evaluation is done at global scale, on a systematic subsample (see §3.3).The key questions that are answered are:

   Q2.1. What is the statistical consistency of Proba-V C2 with SPOT/VGT C3 and Sentinel-3 SYN-VGT? What is the magnitude of the difference

   - *between Proba-V and SPOT/VGT C3 for the overlapping period?*
   - *between the LTS of Proba-V and SPOT/VGT C3?*
   - *between the LTS of Proba-V and recent data of S3/SYN-VGT?*

   Q2.2. What is the spatial consistency of Proba-V C2 with SPOT/VGT C3 and S3/SYN-VGT?

   Q2.3. How do the results of Proba-V C2 compare with those of Proba-V C1?

3. **Comparison to reference time series from external datasets** derived from MetOp/AVHRR (LSA-SAF/ENDVI10) and Terra/MODIS (MOD13A3). The evaluation is done for the Proba-V C2 S10-TOC 1 km NDVI at global scale on a systematic subsample (see §3.3.1), except for the temporal plots generated over LANDVAL sites (see §3.3.2).

   The key questions that are answered are:

   Q3.1. What is the statistical consistency between Proba-V C2 NDVI and external data?

   Q3.2. What is the spatial pattern of the differences?

   Q3.3. What is the spatio-temporal evolution of the differences between a combined NDVI series from SPOT/VGT C3, Proba-V C2 and S3/SYN-VGT and external data?

   Q3.4. What is the temporal consistency between a combined NDVI series from SPOT/VGT C3, Proba-V C2 and S3/SYN-VGT and external data over LANDVAL sites?

The following paragraphs describe the criteria under evaluation and the different methods that are used. An overview of the analysis methods is given in §3.8. For a more detailed description of the data sets used, see Chapter 4.

3.2. Criteria

Table 2 provides a description of the criteria under evaluation.

Table 2: Description of the criteria under evaluation

Criterium	Description
Product completeness	Product completeness is linked with the absence of gaps and the occurrences of quality flags in the product, both in space and in time. Gaps are caused by cloud or snow contamination, bad illumination conditions (e.g. in winter), poor atmospheric conditions or technical problems during image acquisition, and are generally considered as a severe limitation. We focus on the temporal evolution and spatial distribution of product completeness, and occurrences of quality flags (see §3.5).
Spatial consistency	Spatial consistency refers to the realism and repeatability of the spatial distribution of retrievals over the globe, including the absence of artefacts (e.g., missing data, stripes, unrealistic values, etc.), based on expert knowledge. The analysis is based on the spatial distribution of validation metrics at global level (on a systematic subsample, see §3.3.1), and on the statistical consistency stratified per biome (see §3.3.3).
Statistical consistency	The evaluation of statistical consistency or magnitude of the difference between datasets is based on (i) the geometric mean regression between datasets (see §3.4.1); (ii) comparison of histograms between datasets; and (iii) histograms of overall bias between datasets.
Temporal consistency	Temporal consistency involves evaluation of the temporal evolution of validation metrics computed per scene (i.e. 10-daily or monthly period). Spatio-temporal evolution of validation metrics is assessed through Hovmöller plots (see §3.7). The realism of temporal variations of the product are qualitatively assessed for validation sites, well distributed over the globe (see §3.3.2).

3.3. Sampling

3.3.1. Global spatial subsample

The global S1 and S10 images are systematically spatially subsampled over the whole globe taking one pixel every 20 (for 1 km products), 60 (for 300 m products) or 180 (for 100 m products) pixels in both X and Y direction. This (arbitrary) subsample is representative for the global patterns of vegetation and considerably reduces processing time, while retaining the original resolution and the relation between the observation and its viewing and illumination geometry.

3.3.2. LANDVAL sites

The LANDVAL network of around 700 sites (Figure 1) is used to evaluate the temporal consistency between Proba-V C2 NDVI and other datasets.

This network is composed of 521 sites coming from Surface Albedo Validation Sites (SAVS) 1.0 network (Loew et al., 2016), available at http://savs.eumetsat.int. SAVS 1.0 was created during the Surface Albedo Validation 2 (ALBEDOVAL-2) study (Fell et al., 2015), in the framework of Quality Assurance for Essential Climate Variable (QA4ECV) project. Note that this SAVS 1.0 network contains 256 sites from BELMANIP2.1 network. In addition, 20 sites (*'calibration sites'*) in the Sahara Desert and Arabia desert are included in order to increase the sampling over desertic areas and African region. These reference sites, well known for their high temporal stability, are used by Centre National d'Études Spatiales (CNES) for the absolute calibration of remote sensing sensors. Finally, 184 sites coming from existing (e.g. ImagineS (http://fp7-imagines.eu/), AsiaFlux, NARMA or OzFlux) networks or Geo-Wiki platform (http://www.geo-wiki.org/) were included in order to cover under sampled regions (Asia, Africa, Oceania) and biome types (shrubs, deciduous broadleaf forest (DBF), needle leaf forest (NLF)). The methodology for the selection of sites is described in Fuster et al. (2020).

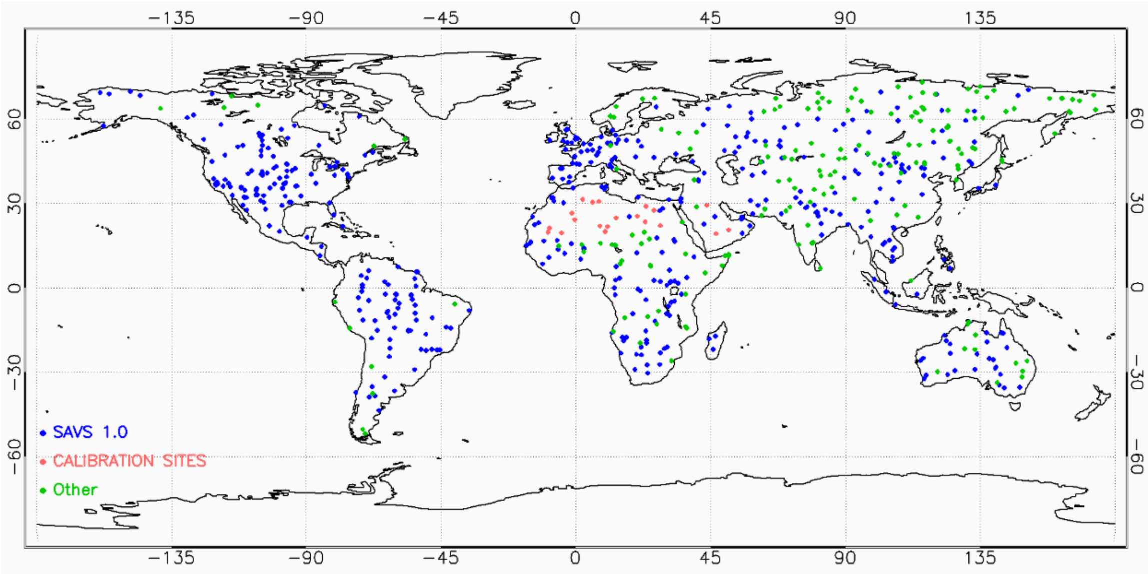


Figure 1: Global distribution of the selected LANDVAL sites (Fuster et al., 2020).

The LANDVAL sites are used for the generation of temporal plots. Biome information per site is based on the aggregated version of the CGLS Global Land Cover at 100 m (see §3.3.3).

3.3.3. STRATIFICATION PER BIOME

An aggregated version of the Copernicus Global Land Service (CGLS) – Land Cover at 100 m (LC100), epoch 2015 (Buchhorn et al., 2019) is used to distinguish between major land cover classes at the global scale (Figure 2). The classes were aggregated according to the scheme in Table 3. A global map at 300 m spatial resolution was generated from the CGLS-LC100 discrete classification map: for each of the 60 global UTM zones, downscaling from 100 m to 300 m was performed by aggregating following the mode of the discrete classes. In case of equal occurrence of discrete classes, a set of expert rules based on cover fractions is applied. A new global tiling grid at 300 m resolution in EPSG:4326 (WGS84) was created to which each of the UTM zones was transformed following the best-available-pixel approach.

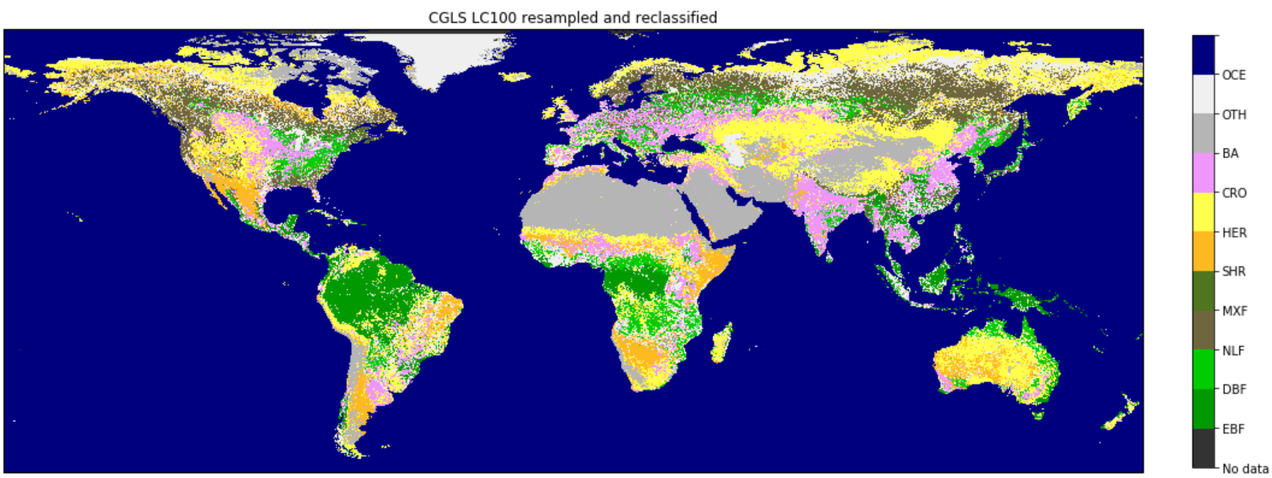


Figure 2: The CGLS LC100 classification aggregated into 8 classes

Table 3: Aggregation scheme for CGLS Land Cover 100 m classes into 8 major biomes and proportion of each biome at global scale

Abbreviation	Name	CGLS-LC100 classes	Proportion at global scale (%)
EBF	Evergreen broadleaf forest	112, 122	7.6
DBF	Deciduous broadleaf forest	114, 124	6.2
NLF	Needleleaf forest	111, 113, 121, 123	10.4
MXF	Mixed forest	115, 125	1.4
SHR	Shrubland	20	7.5
HER	Herbaceous	30	21.2
CRO	Crop	40	10.9
BA	Bare/sparse vegetation	60, 100	15.3
OTH	Other (not considered in the analyses)	0, 50, 70, 80, 90, 116, 126, 200	19.4

3.3.4. PIXEL SELECTION

For the pairwise comparison between PV-C1 and PV-C2 TOA or TOC reflectances and NDVI, pixels identified in both status maps (SM) as 'clear' and with good radiometric quality in all bands (i.e. 'good' quality, see §3.5) are selected. A complementary constraint is based on the identical time of observation. As a result, C1 and C2 reflectance and NDVI that are derived from identical observations are compared. This means that the same observation was selected in the compositing step for C1 and C2 data.

In the period under consideration, changes (variable in time) were applied to the VNIR and SWIR absolute calibration (see §2.2). In order to discriminate between the three different Proba-V cameras (see instrument layout in Figure 3) to evaluate differences between cameras, additional sampling is further used. This is done based on thresholds on the Viewing Zenith Angle (VZA) and Viewing Azimuth Angle (VAA) of each VNIR observation (Table 4).

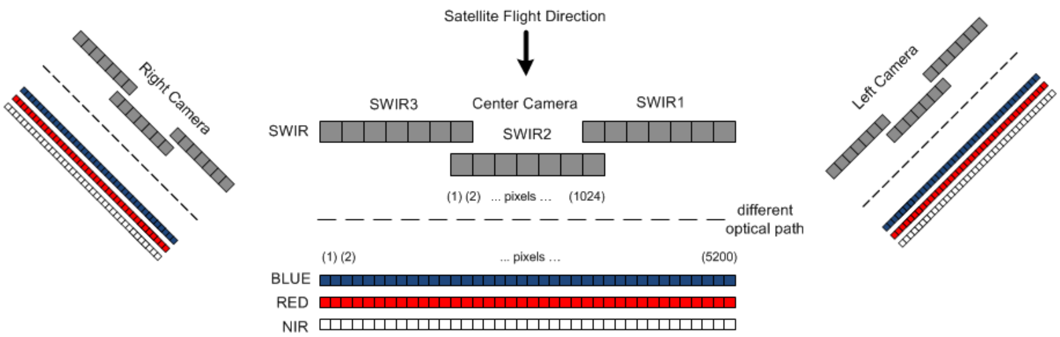


Figure 3 Proba-V instrument layout

Table 4 Thresholds on VNIR VZA and VAA to discriminate between LEFT, CENTER and RIGHT camera

	LEFT	CENTER	RIGHT
VZA (VNIR)	> 20°	< 18°	> 20°
VAA (VNIR)	< 90° OR > 270°		between 90° and 270°

For the consistency analysis between Proba-V and SPOT/VGT or S3/SYN-VGT (CHAPTER 6), pixels identified in both status maps (SM) as 'clear' are selected. No additional constraints were applied. Also for the comparison to external datasets (CHAPTER 7), pixels identified in both status maps (SM) as 'clear' (or 'unflagged') are selected.

3.4. Inter-comparison metrics

3.4.1. Geometric Mean Regression

The geometric mean (GM) regression model is used to identify the relationship between two data sets of remote sensing measurements. Because both data sets are subject to noise, it is most appropriate to use an orthogonal (model II) regression. The GM regression model minimizes the sum of the products of the vertical and horizontal distances (errors on Y and X) and is of the form:

$$Y = a + b \cdot X$$

By applying an eigen decomposition to the covariance metrics of X and Y, two eigenvectors v1 and v2 are obtained that describe the principal axes of the point cloud (Duveiller et al., 2016)

$$b = \frac{\lambda_1 - {\sigma_X}^2}{\sigma_{XY}} \qquad \text{(GMR slope)}$$

$$a = \bar{Y} - b \cdot \bar{X} \qquad \text{(GMR intercept)}$$

with

λ_1: the eigenvalue associated to the first eigenvector defining the principal axis

σ_X: the standard deviation of X

σ_{XY}: the covariance between X and Y

$\bar{X}$: the mean value of X

$\bar{Y}$: the mean value of Y

3.4.2. Coefficient of determination (R^2)

The coefficient of determination (R^2) indicates agreement or covariation between two data sets with respect to a linear regression model. It summarizes the total data variation explained by this linear regression model. The result varies between 0 and 1 and higher R^2 values indicate higher covariation between the data sets.

$$R^2 = \left(\frac{\sigma(X,Y)}{\sigma(X) \cdot \sigma(Y)} \right)^2$$

with σ(X) and σ(Y) the standard deviation of X and Y and σ(X, Y) the co-variation of X and Y. R^2 cannot be used for the comparison of the time series per pixel, due to temporal autocorrelation.

The GM regression slope and intercept and R^2 are added as quantitative information related to the scatterplots.

3.4.3. APU METRICS

The differences between two datasets are evaluated through assessment of the Accuracy, Precision and Uncertainty (APU) metrics.

→ **Mean Bias Error (MBE) or Accuracy (*Acc*)**

The Mean Bias Error (MBE) measures the average actual difference between two data sets and positive and negative differences between observations. It is defined as

$$MBE = \mathrm{Acc} = \frac{1}{n}\sum_{i=1}^{n}(X_i - Y_i) = \bar{X} - \bar{Y}$$

The MBE retains the sign of the difference between the data sets and is a measure for systematic errors. MBE is therefore used to evaluate overall bias or Accuracy (*Acc* or A). The bias distribution is represented in bias histograms.

→ **Standard deviation of the bias (STD) or Precision (*Prec*)**

Precision (*Prec* or P) represents the dispersion of product retrievals around their expected value and can be estimated by the standard deviation (STD) of the bias between retrieved satellite products:

$$STD = \ Prec = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - Y_i - A)^2}$$

→ **The Root Mean Squared Difference (RMSD) or Uncertainty (*Unc*)**

The Root Mean Squared Difference (RMSD) measures how far the difference between the two data sets X and Y deviates from 0 and is defined as

$$RMSD = Unc = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - Y_i)^2}$$

The RMSD is an expression of the overall difference, including random and systematic differences, and is as such also a measure for uncertainty (*Unc* or U).

3.5. STATUS MAP LABELLING

In order to evaluate product completeness, the missing values or pixels flagged as invalid over land are quantified over all land pixels. Pixels labelled as ‘Sea’ are excluded from the analyses. Pixels are considered ‘Good’ if the Cloud/Ice/Snow/Shadow Flag indicates ‘Clear’ and the radiometric quality in all 4 bands is good.

For the assessment of the differences in the Status Map (SM) between C1 and C2, the Status Maps are converted into a single label for unclear observations, based on the SM bit combinations (Table 5).

Table 5 Interpretation of the Status Map and labelling of unclear observations

S10 SM bit	Description	Value	Key	Label used in the analysis
0-2	Cloud/Ice/Snow/Shadow Flag	000	Clear	
		001	Shadow	Shadow
		010	Undefined	Undefined
		011	Cloud	Cloud
		100	Ice	Snow/ice
3	Land/Sea	0	Sea	(excluded)
		1	Land	
4	Radiometric quality SWIR flag	0	Bad	bad SWIR
		1	Good	
5	Radiometric quality NIR flag	0	Bad	bad NIR
		1	Good	
6	Radiometric quality RED flag	0	Bad	bad RED
		1	Good	
7	Radiometric quality BLUE flag	0	Bad	bad BLUE
		1	Good	

3.6. LONG TERM STATISTICS (LTS)

Long term statistics (LTS) are calculated on the extracted sample (see §3.3): for each period in the year, an average value is calculated based on the 5 years of input products. For Proba-V, the input period is 2014-2018; for SPOT/VGT this is 2009-2013.

Since there is no overlap between Proba-V C2 and S3/SYN-VGT products with sufficient maturity, intercomparison with 10-daily composite S3/SYN-VGT products of 2022 is done with the long term statistics (LTS) of Proba-V. Also the LTS of Proba-V (both C1 and C2) and SPOT/VGT C3 are intercompared.

3.7. SPATIO-TEMPORAL EVOLUTION OF THE DIFFERENCE

Hovmöller diagrams are used to perform a combined assessment of the spatial and temporal variability of the inter-comparison metrics. The metrics are derived for each time step (10-day period or month) and for each spatial subset, defined as latitude bands of 6° wide (Figure 4), thereby depicting the temporal evolution of the spatial agreement (Meroni et al., 2013). The resulting time-latitude Hovmöller diagram allows summarizing the space-time features of the time series evaluation.

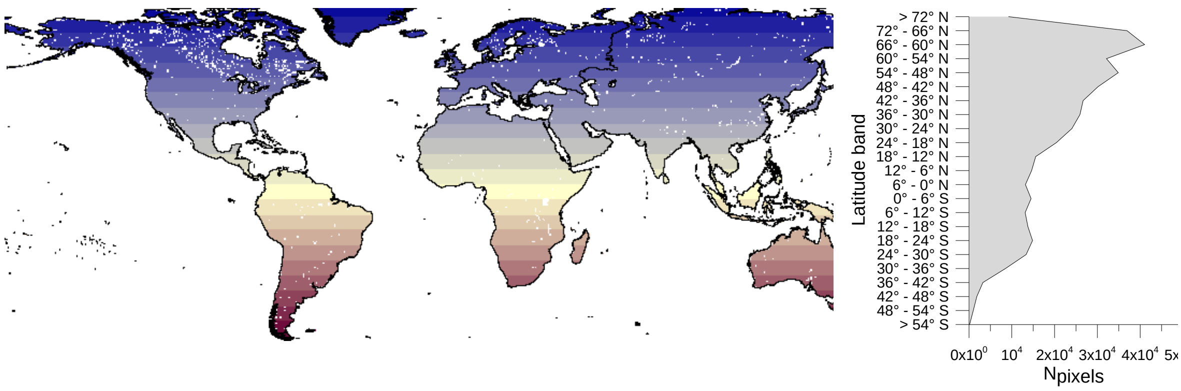


Figure 4 Latitude bands used for spatio-temporal analysis (left) and number of pixels per band (right)

3.8. OVERVIEW OF THE ANALYSIS METHODS

Table 6 provides an overview of the analysis methods. The table lists the method or validation metric (see §3.4) used to evaluate a certain criterium (see §3.2), thereby answering the key questions listed in §3.2. The table also mentions the section where the results of the specific analysis can be found. The datasets that are used for intercomparison are described in detail in CHAPTER 4.

Table 6: Overview of the analysis methods

Section	Key question	Criterium	Method and/or Validation metric	Data / Comparison	Spatial coverage	Temporal coverage
§Q1.1	Q1.1	Product completeness	Overall, spatial and temporal occurrences of product completeness and quality flags (see §3.5)	For both PV-C2 and PV-C1: S10-TOC 1 km S10-TOC 300 m S5-TOC 100 m S1-TOA 1 km	Global subsample	Oct/2013 – Jun/2020
§5.3	Q1.2	Statistical consistency	Geometric mean regression, comparison of histograms, bias histograms and APU metrics.	PV-C2 vs. PV-C1 S1-TOA 1 km S10-TOC 1 km	Global subsample	Oct/2013 – Jun/2020
§5.4	Q1.2	Spatial consistency	Bias histograms stratified per biome. Spatial distribution of APU metrics.	PV-C2 vs. PV-C1 S10-TOC 1 km	Global subsample	Oct/2013 – Jun/2020
§5.5	Q1.2	Temporal consistency	Temporal evolution of APU metrics. Spatio-temporal evolution of APU metrics (Hovmöller plots, see §3.7) (S10 only).	PV-C2 vs. PV C1 S1-TOA 1 km (per camera) S10-TOC 1 km	Global subsample	Oct/2013 – Jun/2020
§6.2.1	Q2.1 Q2.3	Statistical consistency	Comparison of PV-C2 resp. PV C1 with VGT C3 for the overlapping period. Geometric mean regression, comparison of histograms, bias histograms and APU metrics.	PV-C2 (C1) S10-TOC 1 km vs. VGT-C3 S10-TOC 1 km	Global subsample	Oct/2013 – Jun/2014
			Comparison of PV-C2 resp. PV C1 LTS (2014-2018) with VGT C3 LTS (2009-2013). Geometric mean regression, comparison of histograms, bias histograms and APU metrics.	LTS PV-C2 (C1) S10-TOC 1 km vs. LTS VGT-C3 S10-TOC 1 km	Global subsample	LTS (2009-2013) vs. LTS (2014-2018)
§6.3.1	Q2.2 Q2.3	Spatial consistency	APU metrics stratified per biome.	PV-C2 (C1) S10-TOC 1 km vs. VGT-C3 S10-TOC 1 km	Global subsample	Oct/2013 – Jun/2014
				LTS PV-C2 (C1) S10-TOC 1 km vs. LTS VGT-C3 S10-TOC 1 km	Global subsample	LTS (2009-2013) vs. LTS (2014-2018)
§6.2.2	Q2.1 Q2.3	Statistical consistency	Comparison of PV-C2 resp. PV-C1 LTS (2014-2018) with SYN V10 (2022). Geometric mean regression, comparison of histograms, bias histograms and APU metrics.	LTS PV-C2 (C1) S10-TOC 1 km vs. S3/SYN V10	Global subsample	LTS (2014-2018) vs 2022
§6.3.2	Q2.2 Q2.3	Spatial consistency	APU metrics stratified per biome.	LTS PV C2 (C1) S10-TOC 1 km vs. S3/SYN V10	Global subsample	LTS (2014-2018) vs 2022
§7.2	Q3.1	Statistical consistency	Geometric mean regression, comparison of histograms, bias histograms and APU metrics.	PV C2 S10-TOC 1 km NDVI vs. LSA-SAF ENDVI10 v2 PV C2 S10-TOC 1 km NDVI vs. MOD13A3 C6.1 NDVI	Global subsample	Oct/2013 – Jun/2020

§7.3	Q3.2	Spatial consistency	Geometric mean regression, comparison of histograms, stratified per biome.	PV C2 S10-TOC 1 km NDVI vs. LSA-SAF ENDVI10 v2 PV C2 S10-TOC 1 km NDVI vs. MOD13A3 C6.1 NDVI	Global subsample	Oct/2013 – Jun/2020
§7.4	Q3.3	Temporal consistency	Spatio-temporal evolution of APU metrics (Hovmöller plots, see §3.7)	PV C2 S10-TOC 1 km NDVI vs. LSA-SAF ENDVI10 v2 PV C2 S10-TOC 1 km NDVI vs. MOD13A3 C6.1 NDVI	Global subsample	Oct/2013 – Jun/2020
§7.5	Q3.4	Temporal consistency	Temporal plots over sites	PV C2 S10-TOC 1 km NDVI PV C1 S10-TOC 1 km NDVI VGT C3 S10-TOC 1 km NDVI S3/SYN V10 NDVI LSA-SAF ENDVI10 v2 MOD13A3 C6.1 NDVI	LANDVAL	2009-2022

CHAPTER 4 DATA SETS USED IN THE EVALUATION

4.1. INTRODUCTION

The datasets that are used in this report are

- Proba-V (PV) C2 and the former C1 (Oct/2013-Jun/2020):
  - S1-TOA reflectance and status map at 1 km resolution
  - S5-TOC status map at 100 m resolution
  - S10-TOC surface reflectance, NDVI and status map at 1 km and 300 m resolution
  - Monthly maximum value composite (MVC) NDVI at 1 km resolution
  - LTS based on S10 surface reflectance and NDVI at 1 km resolution (2014-2018)
- SPOT/VGT Collection 3 (C3) (2009-Jun/2014):
  - S10-TOC surface reflectance, NDVI and status map at 1 km resolution
  - Monthly MVC NDVI at 1 km resolution
  - LTS based on S10 surface reflectance and NDVI at 1 km (2009-2013)
- Sentinel-3 SYN-VGT (Jul/2020-2022):
  - SYN V10 surface reflectance, NDVI and status map at 1 km resolution
- LSA-SAF MetOp/AVHRR (2009-2022):
  - ENDVI10 v2 NDVI and status map at 1 km resolution
- Terra/MODIS (2009-2022):
  - MOD13A3 monthly NDVI and quality layer at 1 km resolution

All datasets are described in more detail in the following paragraphs.

4.2. PROBA-V COLLECTION 1 AND COLLECTION 2

Proba-V was designed to bridge the gap in space-borne vegetation measurements between the SPOT/VGT mission (March/1998 – May/2014) and the Sentinel-3 satellites (from February/2016 onwards) and was launched in May/2013 (Francois et al., 2014). Detailed descriptions of the Proba-V mission and processing chains are provided in Dierckx et al. (2014) and Sterckx et al. (2014). More information on the processing of Proba-V C2 data can be found in the PV C2 Algorithm Change Document and the PV C2 Product User Manual. The expected impacts of algorithm updates in C2 are described in CHAPTER 2. Details about Proba-V C1 and the evaluation of the Proba-V C1 products are described in Toté et al. (2018).

For the comparison of product completeness and status map labelling between C2 and C1, we looked at S10 composites (at 1 km and 300 m resolution), S5 composites (at 100 m resolution) and S1 composites (at 1 km resolution). S1-TOA reflectances were compared at 1 km resolution. Finally, S10 surface reflectances and NDVI were compared at 1 km resolution.

All analyses were performed over the entire reprocessing period (15/10/2013 – 30/06/2020).

For the comparison with S3/SYN-VGT and SPOT/VGT, long term statistics (LTS) were calculated over the period 2014-2018 (see §3.6). For the comparison with MODIS NDVI, maximum value compositing was done to convert the S10s to monthly data.

In the combined time series of SPOT/VGT (see §4.3), Proba-V and Sentinel-3/SYN-VGT (see §4.4), the switch from SPOT/VGT to Proba-V is made on 01/01/2014. The switch to Sentinel-3/SYN-VGT is made from 01/07/2020 onwards.

4.3. SPOT/VEGETATION COLLECTION 3

SPOT4 (launched March/1998) and SPOT5 (launched May/2002) carried the Vegetation 1 (VGT1) and Vegetation 2 (VGT2) multispectral instruments. The switch from VGT1 to VGT2 was made in February/2003, with higher geometric performances for VGT2 in comparison to VGT1 thanks to the on-board star tracker of SPOT5 (Deronde et al., 2014).

After the end of the SPOT/VGT mission in May/2014, the complete VGT2 Collection 2 (VGT-C2) data archive was reprocessed, resulting in the Collection 3 archive (VGT-C3) (Toté et al., 2017). The VGT2 S10-TOC reflectance and NDVI from 2009 till June/2014 of VGT-C3 were used in the evaluation. For more details on the processing of SPOT/VGT data, we refer to the SPOT/VGT Products User Manual (Wolters et al., 2016).

Comparison of VGT-C3 with Proba-V C1 and C2 is done over the overlapping period (15/10/2013 – 01/06/2014, i.e. 24 S10 products). In the comparison of the combined series of SPOT/VGT and Proba-V with external datasets, the switch from SPOT/VGT to Proba-V is made from 01/01/2014 onwards. For the comparison with MODIS NDVI, maximum value compositing was done to convert the S10s to monthly data. For the comparison with Proba-V, long term statistics (LTS) were calculated over the period 2009-2013 (see §3.6).

Several aspects related to the satellite/sensor and processing definitions impact the consistency between products derived from SPOT/VGT and Proba-V. The relative spectral response of the Proba-V sensor was defined to be as similar as possible to SPOT/VGT (see also Annex A), and differences are indeed very small, except for the SWIR band (Dierckx et al., 2015).

There is an important difference in equator local overpass time between SPOT5 and Proba-V (see Annex B). From 2009 onwards, SPOT5/VGT2 experienced orbital drift, causing the satellite overpass time to gradually shift over time. This was also the case for Proba-V, since this satellite had no onboard propulsion. This evolution causes systematic changes in illumination conditions and related BRDF effects; e.g. NDVI tends to increase with higher solar zenith angles (Galvao et al., 2004; Sellers, 1985; Swinnen et al., 2014).

The atmospheric correction scheme of SPOT/VGT C3 is more similar to Proba-V C1 (e.g. with image based AOT retrieval) than to Proba-V C2.

Another important aspect is the radiometric calibration accuracy. Updated absolute calibration parameters were generated for the reprocessing of the entire VGT archive (resulting in Collection 3), aiming to improve the stability of the instrument responses (Toté et al., 2017). It is to be noted that SPOT/VGT and Proba-V are independently calibrated and no inter-calibration between the sensors is performed.

The nominal (sub-nadir) resolution of the SPOT/VGT sensors is approximately 1 km, whereas this is 100 m at nadir and up to 350m at the extremity of the swath for Proba-V VNIR channels (Francois et al., 2014).

4.4. SENTINEL-3 SYN-VGT

Sentinel-3 (S3) Level-2 synergy (SYN) products rely on the combination of input products of the Ocean and Land Colour Instrument (OLCI) and Sea and Land Surface Temperature Radiometer (SLSTR). The SYN-VGT products are designed to provide surface vegetation products similar to those obtained from the Vegetation (VGT) instrument onboard the SPOT satellites, with complete Earth coverage every 1-2 days.

The SYN-VGT consists of the TOA observations (VGT-P or VGP), daily TOC synthesis (VGT S1 or VG1) and 10 day synthesis (VGT S10 or V10). The VG1 and V10 products are maximum NDVI value composites of ground reflectance measurements and NDVI of all segments received during 1 day and 10 days, respectively. The VG1 and V10 products provide surface reflectance for all spectral bands, the NDVI and ancillary data on image acquisition parameters. All products are provided on the same regular latitude-longitude grid as SPOT/VGT and Proba-V products.

S3/SYN-VGT V10 products from July/2020 onwards are used for the intercomparison with Proba-V C2. S3/SYN-VGT products available before this date show important inconsistencies in the spatial (gridding) and temporal (compositing scheme) domain. Products from July/2020 onwards are made available via the Belgian Collaborative Ground Segment[1]. It is to be noted that these S3/SYN-VGT products still suffer from a number of quality issues that have important impacts on the consistency with Proba-V, such as:

- The use of SPOT4/VGT1 spectral response functions in the spectral resampling procedure, with large differences with Proba-V relative spectral responses, especially in the Red band (see Annex A)
- Absolute radiometric calibration issues of the SLSTR sensor (Smith, 2020)
- Limited global coverage of daily composite products because S3A and S3B observations are not combined in one synthesis product. Since the swath width of Sentinel-3/OLCI is smaller, the combination of OLCI onboard Sentinel-3 A and B is needed to obtain similar coverage.
- Overestimation of AOT in the atmospheric correction (Wolters et al., 2021)
- No cloud shadow detection
- Occasionally, there are problems with availability of S3/SYN VGT products, with missing products (e.g. June/2021), empty products (e.g. July/2021), or products having an incorrect date (e.g. December/2020).
- Until May 2021, the S3/SYN-VGT V10 NDVI product was erroneously based on TOA reflectances.
- Recently, an issue was discovered in the spectral resampling procedure, leading to incorrect band re-mapping of OLCI and SLSTR bands to the VGT-like response. This error affects the NIR and SWIR bands.

S3/SYN-VGT V10 products of 2022 are compared to the Proba-V S10-TOC long term statistics (2014-2018) (see §3.6). In the comparison of the combined series of SPOT/VGT, Proba-V and S3/SYN-VGT with external datasets, the switch from Proba-V to S3/SYN-VGT is made from July/2020 onwards.

For more information on the SYN-VGT product, e.g. product ATBD, product data format specification, metadata specification, Annual Performance Reports etc., we refer to the Sentinel-3 SYN Document Library[2].

4.5. LSA-SAF ENDVI10 v2

The European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) operates the MetOp-satellites, a series of polar orbiting meteorological satellites which form the space

[1] https://www.terrascope.be

[2] https://sentinel.esa.int/web/sentinel/user-guides/sentinel-3-synergy/document-library

segment component of the overall EUMETSAT Polar System (EPS). MetOp carries the AVHRR-instrument and has a stable equator local overpass time at around 8:20.

The 10-daily NDVI composite product from MetOp /AVHRR (ENDVI10) is provided in near-real time by the LSA-SAF[3] from 2008 onwards. ENDVI10 v2 data for the period 2009-2022 is used in the analyses.

The ENDVI10 v2 is processed in a similar way to those of SPOT/VGT and Proba-V, with the same algorithm for atmospheric correction and a similar compositing method (Wolters et al., 2020), which makes the inter-comparison with Proba-V useful. Nevertheless, there are differences in spectral band definition (see Annex A), radiometric calibration, atmospheric inputs, cloud detection and overpass time stability (see Annex B).

Global ENDVI10 v2 data is derived from MetOp-A (launched in 2006) until April/2013. From May/2013 onwards, the nominal sensor is MetOp-B (launched in 2012). It is to be noted that MetOp-B geometric accuracy analyses revealed longitudinal deviations fluctuating between 1.5 and 2 km on the ENDVI10 product (Toté et al., 2019). Independent analysis on the AVHRR Global Area Coverage (GAC) data showed an average value of west shift in the across-track direction of 2.6 km (Wu et al., 2020). For the intercomparison with Proba-V C2, therefore a 2-pixel longitudinal shift was applied on the ENDVI10 v2 data for the MetOp-B period.

4.6. TERRA/MODIS MOD13A3 NDVI

The Moderate Resolution Imaging Spectroradiometer (MODIS) is mounted on two platforms, EOS-TERRA and AQUA. TERRA/MODIS has a stable morning equator overpass at around 10:30 local time (see Annex B), and captures data in 36 bands with a spatial resolution of 250 m (2 bands), 500 m (5 bands) and 1 km (29 bands).

The MODIS sensors are across-track scanners that scan the Earth in a series of lines, perpendicular oriented to the direction of the orbit. Each line is scanned from one side to the other, using a rotating mirror placed in front of a sensor. The mirror sweeps with a constant angular velocity, resulting in the same angular resolution for every measurement. The sensors instantaneous FOV remains the same, and when sweeping away from the nadir position, the distance to the Earth increases and so does the ground surface resolved by the satellite. Because of the large swath width, the Earth's curvature adds an additional panoramic distortion to the off-nadir pixels. This leads to large off-nadir spatial deformations and the bow tie effect (Meyer, 1996).

The NDVI is provided to the user in the form of 16-day or monthly composites at various resolutions; no 10-daily composites are distributed. The data that are used in this study are the MOD13A3 Collection 6.1 data set, which is the TERRA/MODIS global monthly gridded NDVI at 1 km resolution, available from February/2000. These data are weighted temporal averages and are normalized for viewing and illumination angles using the simple Walthall model (Walthall et al., 1985) to normalize the reflectance data to nadir and compute nadir-based VIS reflectances. MOD13A3 NDVI data for the period 2009-2022 is used. In order to prepare this data for intercomparison with the Proba-V archive, the data is reprojected to the 1 km Proba-V grid. The quality indicators were used to exclude water, and pixels contaminated by clouds, shadow and/or snow/ice.

The dataset differs in many ways with those from Proba-V: spectral band definition (see Annex A), radiometric calibration, atmospheric correction, viewing/illumination dependency, cloud detection overpass time stability (see Annex B), and compositing methodology. To allow pairwise comparison, MVC compositing is applied to adapt the temporal resolution of 10-daily composites of SPOT/VGT, Proba-V and S3/SYN-VGT to monthly MVC NDVI syntheses.

[3] https://landsaf.ipma.pt/en/products/vegetation/endviv2/

CHAPTER 5 COMPARISON BETWEEN COLLECTION 2 AND COLLECTION 1

5.1. INTRODUCTION

This chapter focuses on the comparison of PV-C2 with PV-C1 at global level, based on a systematic spatial subsample (see §3.3.1). Product completeness is intercompared for S10-TOC at 1 km, S10-TOC at 300 m, S5-TOC at 100 m and S1-TOA at 1 km. Statistical intercomparison is done for S1-TOA and S10-TOC at 1 km spatial resolution. In this case, pixel selection is done based on clear observations for the same observation day, and if relevant per Proba-V camera (see §3.3.4). For an overview of the analysis methods used, see §3.8. Similar results were observed at 300 m and 100 m resolution, but were excluded from this report for the sake of brevity.

The key questions that are answered in this chapter are:

Q1.1. What is the completeness of the data, in terms of spatial pattern and temporal evolution? What is the difference in flag occurrences between Proba-V C1 and C2?

Q1.2. What is the magnitude of the difference between Proba-V C2 and C1? What is the spatial and temporal pattern of the difference?

5.2. PRODUCT COMPLETENESS AND STATUS MAP LABELLING

5.2.1. GLOBAL AVERAGES

Global averages of the amount of clear and unclear observations over land in the S10-TOC 1 km, S10-TOC 300 m, S5-TOC 100 m and S1-TOC 1 km C1 and C2 archives for the period October/2013 – June/2020 are summarized in Table 7. The methodology for interpretation of the status map is described in §3.5.

In the C2 S10 products, there are more clear observations (+4.4% at 1 km, +11.2% at 300 m resolution), especially due to the lower amount of pixels that are labeled as clouds (resp. -5.3% and -11.3%) and less overdetection of clouds (compared to C1). In C2, slightly more pixels are detected as snow/ice (resp. +1.1% and +1.0%). More clouds and cloud shadows are detected in the 300 m compared to the 1 km: sub-pixel clouds (i.e. smaller than 1 km) are less likely to be identified at 1 km resolution by the cloud detection algorithm.

Similar evolutions are observed for the S5-TOC product at 100 m resolution, with +12.3% clear observations. In general, there are less clear observations at 100 m because only acquisitions of the centre camera are used for this product. Also the S1-TOA at 1 km shows a similar evolution towards more clear observations in C2 (+7%). The decrease in occurrences of ‘undefined’ is related to a change in the compositing scheme (see §2.5).

The S10-TOC product at 300 m resolution shows a slight increase in occurrences of bad SWIR quality. This is related to the change in the compositing rules, that are now harmonized between resolutions (see §2.5). As a result, clear pixels with bad SWIR quality are preferred over cloudy pixels with good SWIR quality. The S5-TOC product at 100 m resolution also shows slight increase of bad radiometric quality.

Overall, this leads to an general increase of ‘good’ pixels occurrences, i.e. clear pixels with good radiometric quality, between 4.6% and 11.5%, depending on the product.

Table 7 Overall difference in status map labelling between C2 and C1 of different products (% land pixels, October/2013 – June/2020)

%	S10-TOC 1 km			S10-TOC 300 m			S5-TOC 100 m			S1-TOA 1 km		
	C1	**C2**		**C1**	**C2**		**C1**	**C2**		**C1**	**C2**	
Good*	72.4	77.1	+4.6	67.8	78.2	+10.4	37.3	48.8	+11.5	33.6	41.3	+7.7
Clear	75.5	80.0	+4.4	69.0	80.2	+11.2	38.4	50.7	+12.3	35.1	42.1	+7.0
Not clear												
cloud	15.2	9.9	-5.3	21.5	10.2	-11.3	51.5	39.2	-12.2	56.6	50.4	-6.2
shadow	1.0	1.2	+0.1	1.9	1.4	-0.4	2.6	2.8	+0.1	1.8	2.2	+0.4
snow/ice	7.7	8.8	+1.1	7.1	8.1	+1.0	6.7	7.2	+0.5	6.3	5.0	-1.3
undefined	0.6	0.1	-0.4	0.5	0.1	-0.4	0.7	0.1	-0.7	0.3	0.3	0.0
Bad rad. Quality												
bad BLUE	7.5	7.5	0.0	6.5	6.5	0.0	17.4	17.6	+0.2	24.1	24.1	0.0
bad RED	3.4	3.4	0.0	3.2	3.2	0.0	9.3	9.5	+0.1	7.5	7.5	0.0
bad NIR	3.3	3.4	0.0	3.2	3.2	0.0	5.3	5.7	+0.5	1.9	1.9	0.0
bad SWIR	5.2	5.2	0.0	3.2	4.1	+0.9	5.3	5.7	+0.4	1.9	1.9	0.0

* 'Good' refers to 'clear' status with good radiometric quality

5.2.2. SPATIAL DISTRIBUTION

When comparing the spatial distribution of the good observations (i.e. observations not labelled as cloud, shadow, snow/ice or undefined, with good radiometric quality) between PV-C1 and PV-C2, overall similar patterns can be discerned (see Figure 5 for S10-TOC 1 km; other results in Annex C). Slightly higher amounts of good observations are visible in C2, especially at high latitudes and in tropical regions. This is especially related to a lower amount of pixels labeled as cloud (see Figure 6; Annex C). The latitudinal bands in the PV-C1 good observations, due to a quality issue in PV-C1, disappear in PV-C2. The 'undefined' quality occurrences at high latitudes largely disappear in PV-C2 due to a different handling of bad radiometric quality in the compositing step (see §2.5): the 'undefined' flag is now reserved for pixels where one of the processing steps (cloud detection or atmospheric correction) was unsuccessful due to missing inputs.

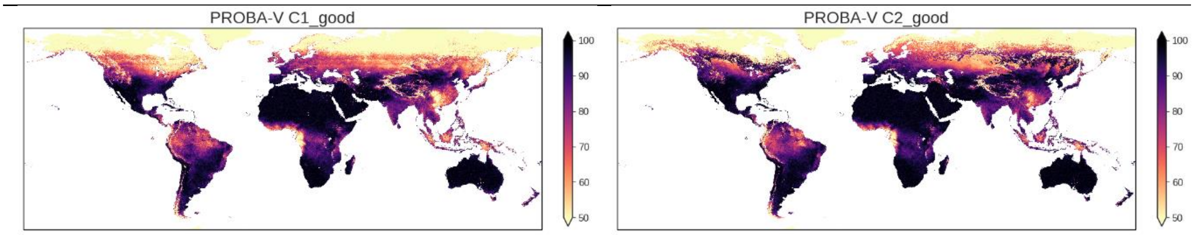


Figure 5: Spatial distribution of the occurrences (%) of good pixel status (clear and good radiometric quality) in PV-C1 (left) and PV-C2 (right) for S10-TOC 1 km (October/2013 – June/2020)

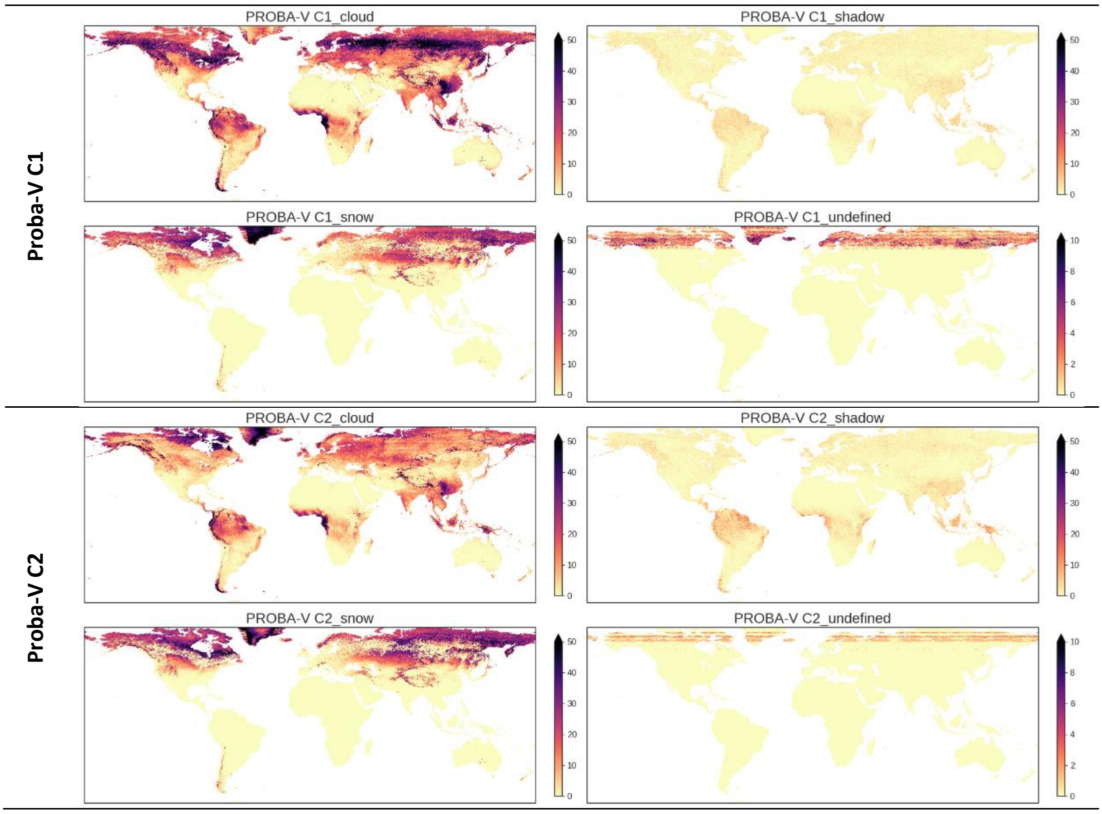


Figure 6: Spatial distribution of the occurrences (%) of cloud, shadow, snow/ice and undefined flags in PV-C1 (top) and PV-C1 (bottom) for S10-TOC 1 km (October/2013 – June/2020)

5.2.3. TEMPORAL EVOLUTION

The temporal evolution of the proportion of land pixels with clear pixel status, good pixel status (i.e. clear with good radiometric quality) or flagged due to cloud, shadow, snow/ice or undefined in both the PV-C1 and PV-C2 archives is illustrated in Figure 7. The seasonal variation in the amount of observations masked as 'good' is resulting from bad illumination conditions in the Northern hemisphere winter, and a cyclic pattern in snow/ice and (to a lesser extent) cloud occurrences. Largest differences between PV-C1 and PV-C2 are visible for the cloud occurrences, although the temporal pattern remains the same. In agreement with the previous results, the differences between PV-C1 and PV-C2 are larger for the 300 m and 100 m resolution products, related to the change in compositing rules (see §2.5). The temporal evolution of flag occurrences shows no trend or deviating patterns.

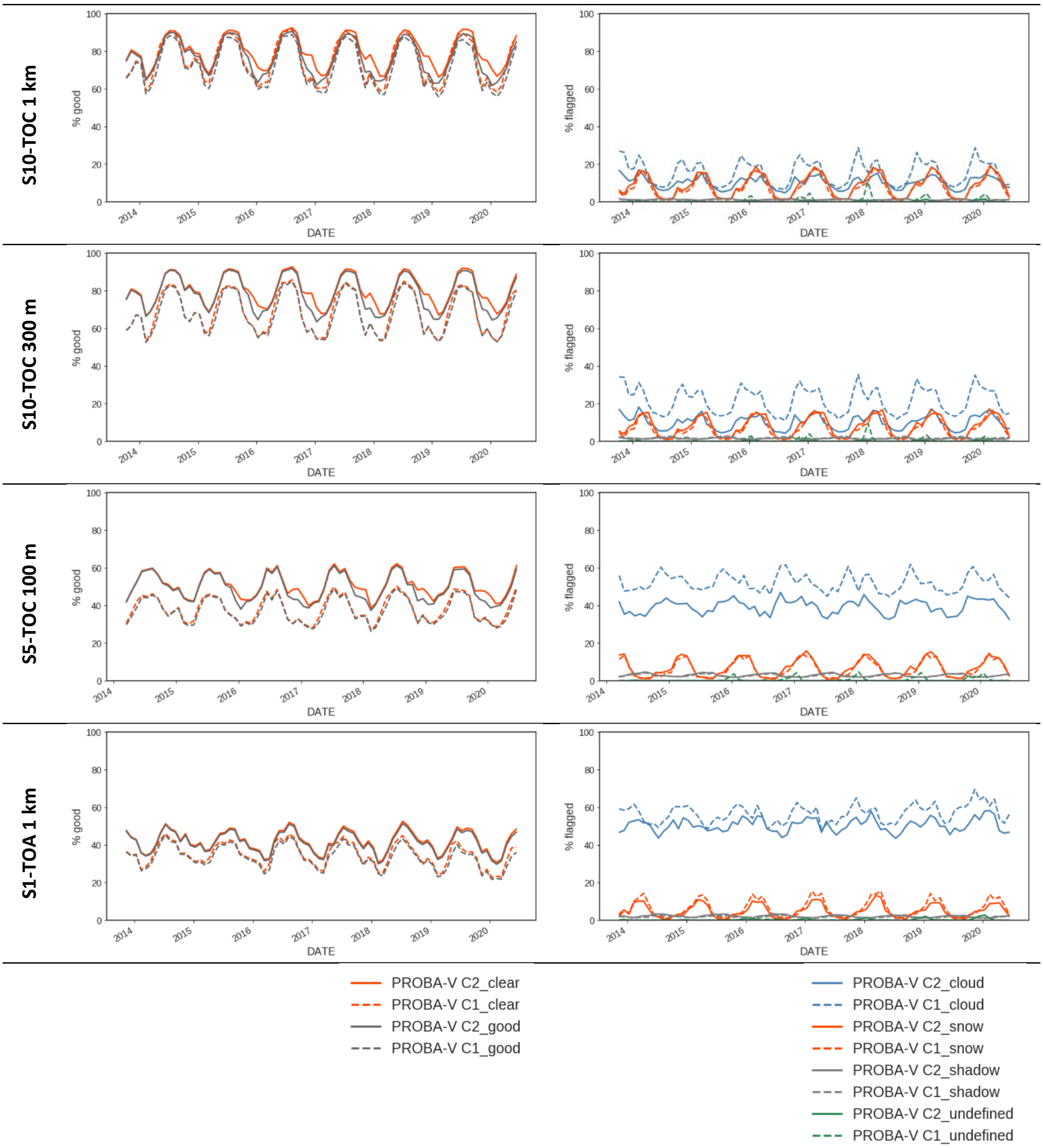


Figure 7: Temporal evolution of percentages (% land pixels) of clear or good (clear with good radiometric quality) pixels (left), or flagged as cloud, snow, shadow or undefined (right) in PV-C2 (solid lines) and PV-C1 (dashed lines), for different products

5.3. MAGNITUDE OF THE DIFFERENCES

5.3.1. TOA REFLECTANCES

The results of the pairwise comparison of PV-C1 vs. PV-C2 TOA reflectances, based on S1-TOA at 1 km resolution, indicate that overall the differences at TOA level are very small (Figure 8). Almost 1:1 regression functions are found for the 4 bands, with R^2 values of 1.0. The histograms overlap almost completely, and the bias histograms show only very small deviations from 0. The average bias (*Acc*) is -0.2% for Blue, +0.1% for Red, and 0 for NIR and SWIR. These differences are in line with the average difference in absolute calibration, over the entire Proba-V lifetime and over all cameras (see

PV C2 Algorithm Change Document). Largest (but still very small) dispersion (*Prec*) is observed for the SWIR band, related to the implementation of improved equalization coefficients, causing small unsystematic bias between PV-C1 and PV-C2.

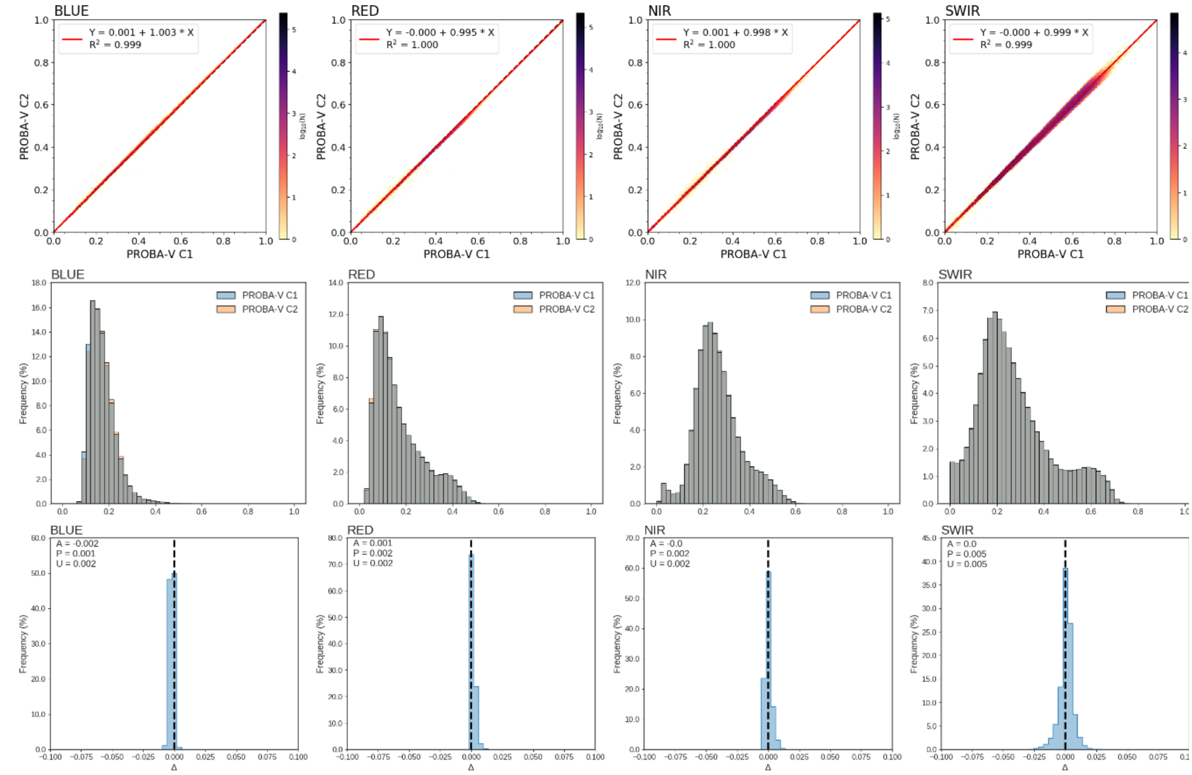


Figure 8: Scatter density plots and GM regression (top), histograms (middle, grey colour indicates where both histograms overlap) and bias histograms with APU metrics (bottom) between PV-C1 (X) and PV-C2 (Y) S1-TOA reflectances

5.3.2. SURFACE REFLECTANCES

At the level of surface reflectances, the overall differences between PV-C1 and PV-C2 are larger (Figure 9), related to the adaptations to the atmospheric correction (see §2.4). Nevertheless, the GMR equations are still very close to the 1:1 line, with very high R^2, histograms largely overlap, and average bias (*Acc*) remains very low. Largest *Acc* is observed for Blue (-0.9%) and Red (-0.3%), where the impact of using different atmospheric inputs in the atmospheric correction is largest, thus amplifying the differences observed at TOA level for the Blue band (see §5.3.1). The average bias is opposite for Red (-0.3%) and NIR (+0.2%), which will have an impact on the NDVI (see below).

The scatterplots and bias histograms also indicate some scatter or unsystematic bias exists between both datasets, especially for the Blue and Red bands. The bias histogram for the Blue band shows the widest distribution. Since the sample is based on pairwise comparison of good quality pixels with identical time of observation in PV-C1 and PV-C2 (see §3.3.4), this is not related to different pixel selection in the compositing process or undetected clouds. It is more likely related to occurrences with high disturbances in the atmosphere, where a more realistic AOT is applied in PV-C2, whereas in PV-C1 AOT was limited to a maximum value of 0.5.

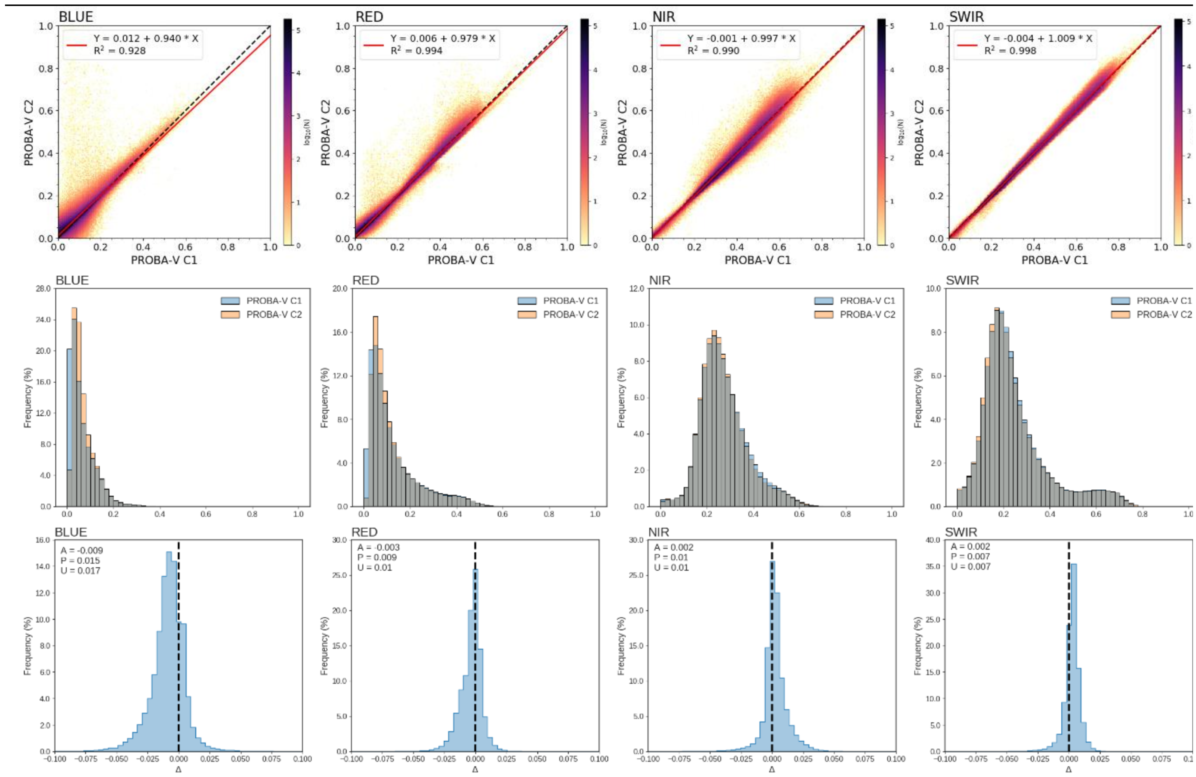


Figure 9: Scatter density plots and GM regression (top), histograms (middle, grey colour indicates where both histograms overlap) and bias histograms with APU metrics (bottom) between PV-C1 (X) and PV-C2 (Y) S10-TOC reflectances

5.3.3. NDVI

Average biases between PV-C1 and PV-C2 are opposite for Red (PV-C2 is 0.3% brighter) and NIR (PV-C2 is 0.2% darker). This leads to lower NDVI in PV-C2, especially for higher NDVI values (Figure 10), with an average bias of 0.02. Where PV-C1 NDVI tends to saturate, this is no longer the case for PV-C2 NDVI. The bias histogram is skewed towards positive values (PV-C1 > PV-C2), and large scatter (high *Prec*) is observed.

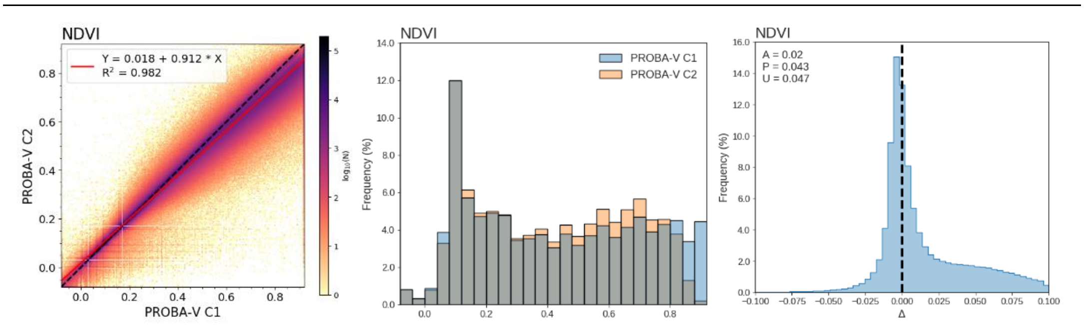


Figure 10: Scatter density plot and GM regression, histogram (grey colour indicates where both histograms overlap) and bias histogram with APU metrics between PV-C1 (X) and PV-C2 (Y) S10-TOC NDVI

5.4. Spatial pattern of the differences

In order to evaluate the spatial pattern of the differences between PV-C1 and PV-C2, bias histograms (see Figure 11 for NDVI; Annex D for surface reflectances) and APU metrics (Table 8) are calculated per biome (see §3.3.3).

While over bare areas (BA) the NDVI bias histogram is very narrow and APU metrics are close to zero, the forest biomes show large skewedness towards positive values and high APU metrics: PV-C2 NDVI is smaller than PV-C1 NDVI. This is related to opposite effects in the Red (negative bias) and NIR (positive bias) bands, which are more pronounced for areas with more dense vegetation cover.

The global spatial distribution of APU metrics shows large *Prec* and *Unc* for surface reflectance bands over the Tropics, where the influence of the atmosphere is largest, and the impact of the use of a more realistic AOT in the atmospheric correction (compared to PV-C1) is largest. The Red and NIR bands show opposite *Acc* in vegetated areas, while this is not the case for e.g. the Sahara region. This leads to strong positive *Acc* for the NDVI over densely vegetated areas.

Table 8: APU metrics between PV-C1 and PV-C2 surface reflectance and NDVI, stratified per biome

		EBF	DBF	NLF	MXF	SHR	HER	CRO	BA
Blue	A	-0.015	-0.010	-0.011	-0.014	-0.008	-0.008	-0.009	-0.003
	P	0.018	0.014	0.016	0.020	0.013	0.013	0.016	0.012
	U	0.024	0.017	0.020	0.024	0.015	0.015	0.018	0.012
Red	A	-0.011	-0.005	-0.005	-0.006	-0.001	-0.001	-0.003	-0.001
	P	0.009	0.008	0.009	0.010	0.006	0.007	0.007	0.012
	U	0.015	0.010	0.011	0.012	0.006	0.007	0.008	0.012
NIR	A	0.009	0.004	0.003	0.006	0.001	0.002	0.004	-0.004
	P	0.011	0.009	0.007	0.009	0.007	0.008	0.010	0.014
	U	0.014	0.010	0.007	0.011	0.007	0.008	0.010	0.014
SWIR	A	0.004	0.003	0.002	0.002	0.003	0.002	0.003	-0.001
	P	0.004	0.004	0.003	0.003	0.006	0.006	0.006	0.011
	U	0.005	0.005	0.004	0.004	0.007	0.007	0.007	0.011
NDVI	A	0.063	0.030	0.038	0.043	0.009	0.009	0.015	-0.003
	P	0.049	0.044	0.056	0.056	0.028	0.030	0.035	0.007
	U	0.079	0.053	0.068	0.070	0.029	0.031	0.038	0.008

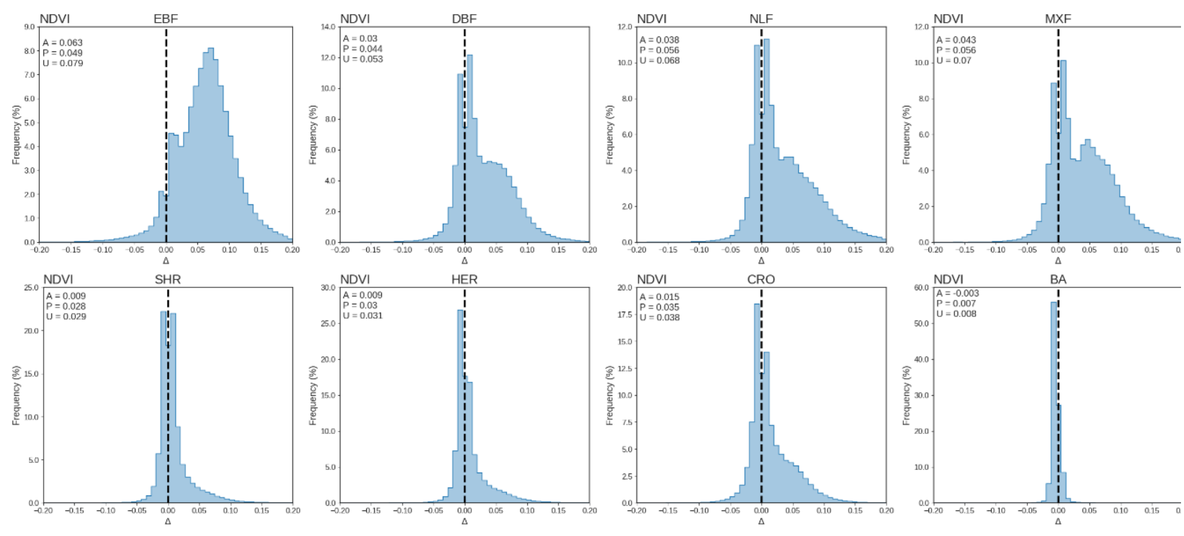


Figure 11: Bias histograms between PV-C1 and PV-C2 TOC NDVI stratified per biome

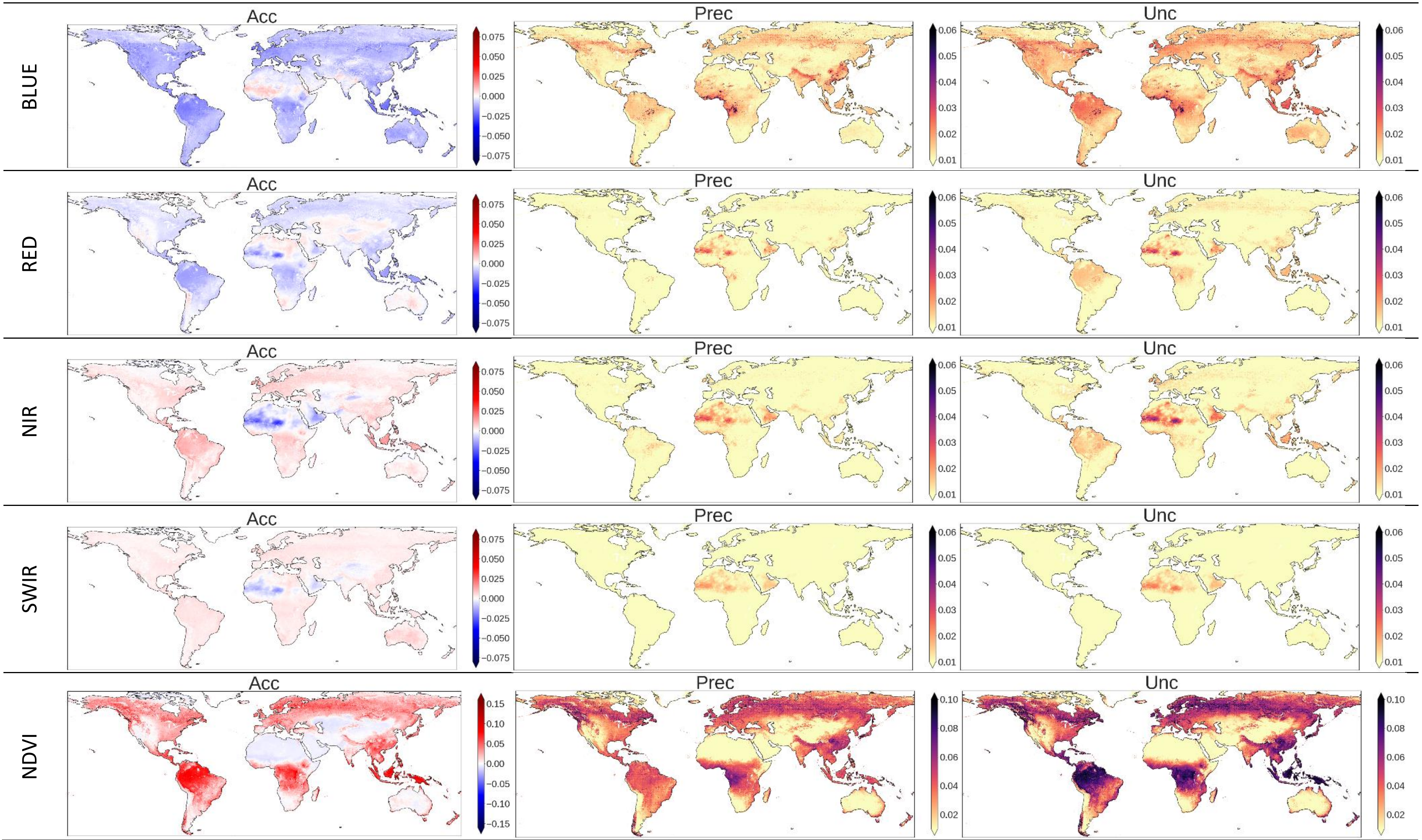


Figure 12: Spatial distribution of APU metrics between PV-C1 and PV-C2 surface reflectance and NDVI

5.5. TEMPORAL EVOLUTION OF THE DIFFERENCES

In the following analyses, we focus on the temporal evolution of the differences between PV-C1 and PV-C2.

Figure 14 (p.28) shows the APU metrics of the intercomparison between PV-C1 and PV-C2 at TOA level per band and per camera. The overall temporal patterns match with the temporal patterns of changes in the absolute calibration coefficients (see the PV C2 Algorithm Change Document). Small spikes appear in the *Prec* and *Unc*, related to small unsystematic differences caused by the difference in geomodelling between PV-C1 and PV-C2, where for some segments a different intermediate projection leads to small differences in geolocation (see §2.6). The discontinuity in November/2016 is caused by the fact that in C1 operations the non-linearity corrections were not correctly applied; this is fixed in C2. Finally, the high frequency fluctuations in the APU metrics for the SWIR, notorious for specific periods, are related to the improvements in the multi-angular calibration of the SWIR strips (see the PV C2 Algorithm Change Document)

The temporal evolution of APU metrics at surface reflectance level (overall and per camera) is summarized in Annex E. The overall effect on the TOC NDVI shows a seasonal pattern (Figure 13), with larger APU metrics during the northern hemisphere summer period (when vegetation cover is more dense). Bias (*Acc*) gradually declines, thanks to the decline in opposite *Acc* in the Red and NIR bands over time.

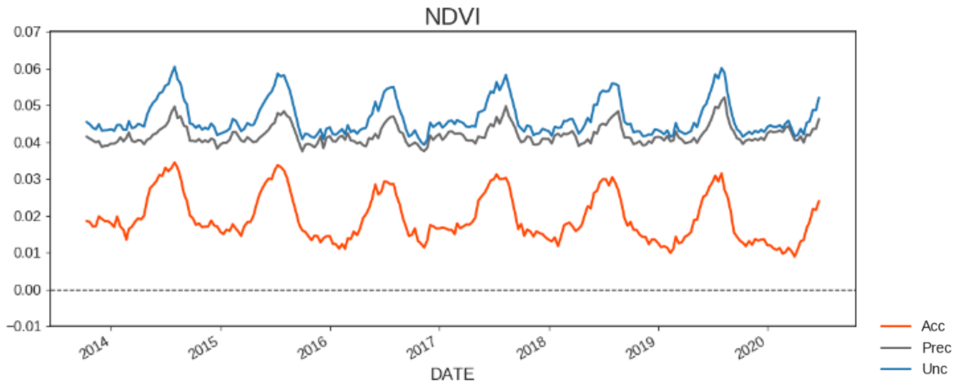


Figure 13: Temporal evolution of APU metrics between C2 and C1 S10-TOC NDVI at 1 km resolution calculated over October/2013 – June/2020, using all clear pixels that have the same observation day.

5.6. SPATIO-TEMPORAL EVOLUTION OF THE DIFFERENCES

The combined analysis of spatial and temporal patterns is performed by evaluation of Hovmöller plots (see §3.7). The Hovmöller plots of APU metrics for the surface reflectance bands and the NDVI all show a cyclic pattern. Standard deviations of the bias (*Prec*) is largest for the Blue band, related to unsystematic impacts of the improved atmospheric correction (see §2.4). The *Acc* for the Red and NIR bands show seasonal patterns with negative resp. positive values over densely vegetated areas in the North-South shifting Tropics (around the equator). This results in positive NDVI *Acc* and high *Unc* over the same region. Also the Northern hemisphere summer period at latitudes above 40°N shows high *Acc* and *Unc* for the NDVI.

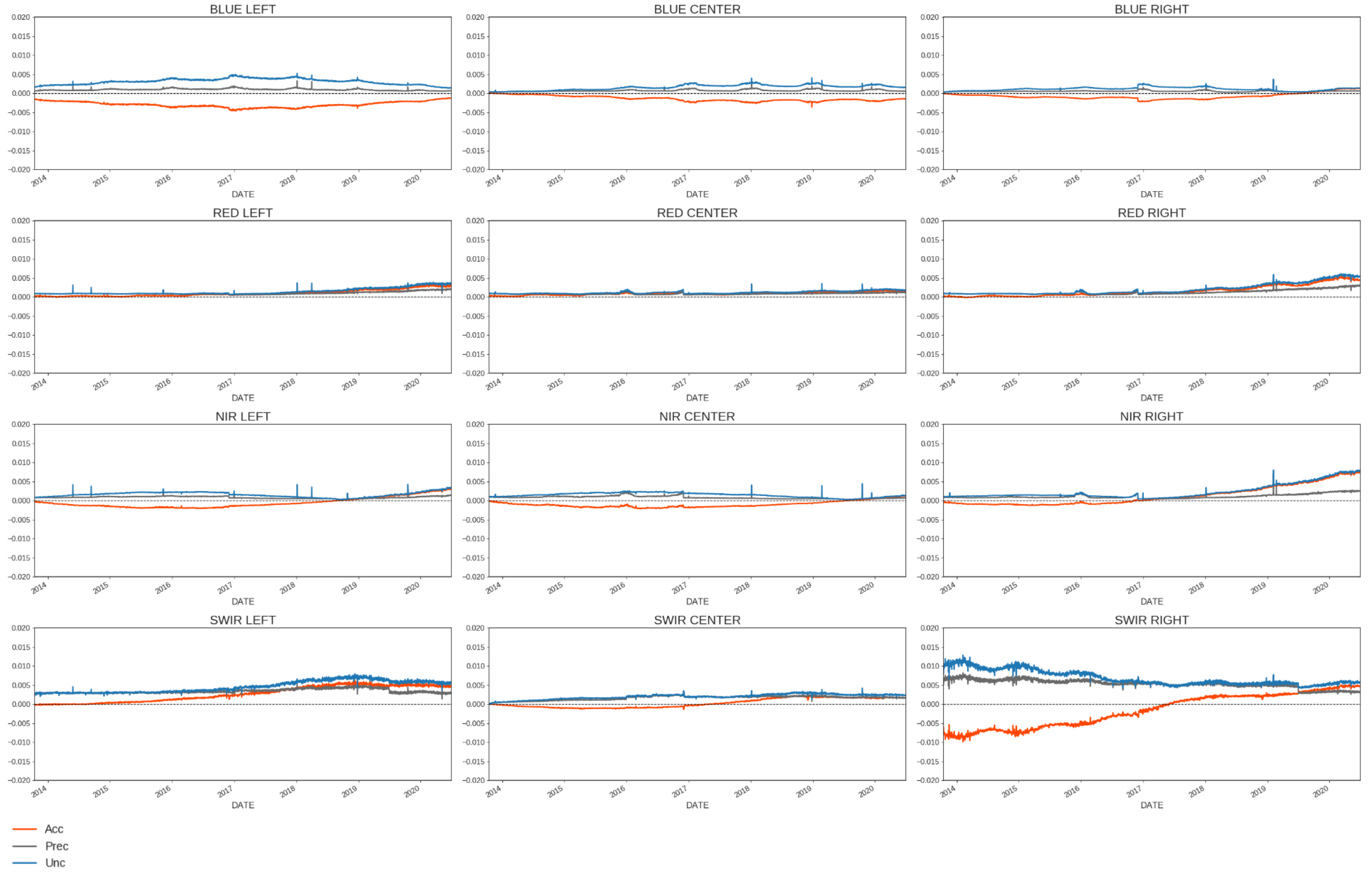


Figure 14: Temporal evolution of APU metrics between C2 and C1 S1-TOA reflectance per band and per camera

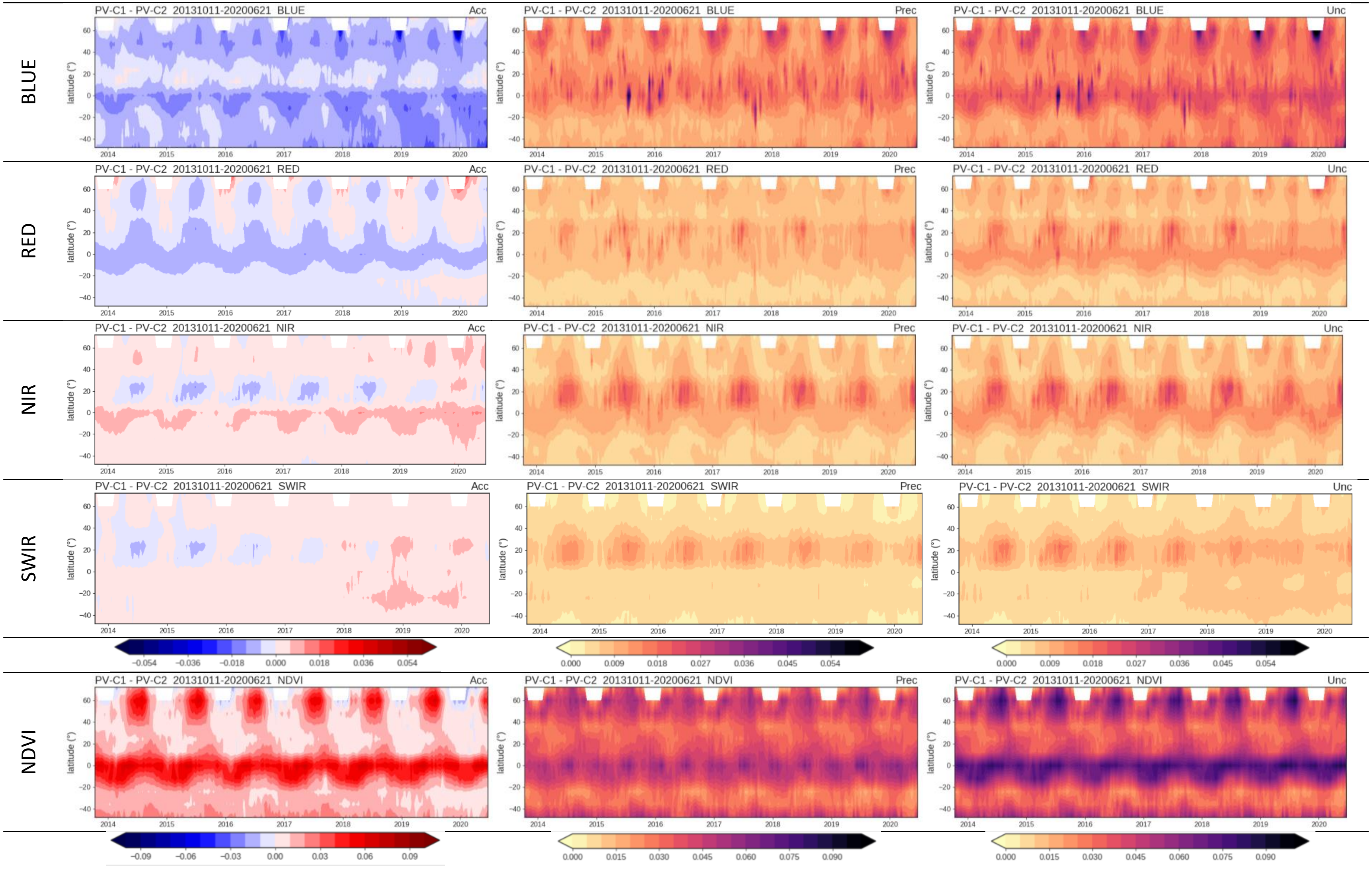


Figure 15: Spatio-temporal evolution of APU metrics between PV-C1 and PV-C2 surface reflectances and NDVI

CHAPTER 6 CONSISTENCY WITH RELATED DATASETS: SPOT/VEGETATION, SENTINEL-3 SYN-VGT

6.1. INTRODUCTION

The analyses in this chapter describe the consistency of Proba-V C2 (and the previous Proba-V C1) with similar datasets from SPOT/VGT C3 (see §4.3) and Sentinel-3/SYN-VGT (see §4.4). We focus on the magnitude of the difference (i.e. statistical consistency) and the spatial variation of the differences. In the next chapter (CHAPTER 7 Comparison to external datasets) spatio-temporal consistency of a combined series of SPOT/VGT-C3, Proba-V C2 and S3/SYN-VGT NDVI with external datasets is assessed (see §7.4). Also temporal plots over LANDVAL sites (see §3.3.2) are evaluated (see §7.5).

The analyses are based on spatially subsampled global S10-TOC reflectances and NDVI at 1 km spatial resolution. For pairwise comparison only good quality pixels in both datasets are considered.

The key questions that are answered in this chapter are:

Q2.1 *What is the statistical consistency of Proba-V C2 with SPOT/VGT C3 and Sentinel-3 SYN-VGT? What is the magnitude of the difference*
- *between Proba-V and SPOT/VGT C3 for the overlapping period?*
- *between the LTS of Proba-V and SPOT/VGT C3?*
- *between the LTS of Proba-V and recent data of S3/SYN-VGT?*

Q2.2 *What is the spatial consistency of Proba-V C2 with SPOT/VGT C3 and S3/SYN-VGT?*

Q2.3 *How do the results of Proba-V C2 compare with those of Proba-V C1?*

6.2. MAGNITUDE OF THE DIFFERENCES

6.2.1. BETWEEN PROBA-V AND SPOT/VGT

In order to evaluate the statistical consistency between PV-C2 (and PV-C1) with VGT-C3, geometric mean regression, histograms, bias histograms and APU metrics are calculated for the overlapping period (October/2013 – June/2014). The results are shown in Figure 16 for PV-C2 and Figure 17 for PV-C1. Overall, high correspondence is visible between Proba-V and SPOT/VGT, although there is quite some scatter, e.g. related to angular effects. The horizontal lines in the scatterplot for the Blue band originate from the cloud detection thresholds applied in the VGT re-processing (see also Toté et al., 2017). The results seem rather similar for PV-C1 and PV-C2, although the R^2 and APU metrics are slightly better for PV-C1. Especially for the blue band, the histograms show larger overlap, and the bias histogram peaks closer to zero, leading to lower *Acc*. This is related to the fact that the atmospheric correction of VGT-C3 is more similar to PV-C1 than to PV-C2.

Because the overlap period between Proba-V and SPOT/VGT is rather short, also the LTS were compared (see §3.6). The results are shown in Figure 18 for PV-C2 and Figure 19 for PV-C1. The patterns are very similar to the results for the overlap period, although different time periods are intercompared. The scatterplots of PV-C1 LTS versus VGT-C3 LTS show a strong scatter below the 1:1 line. This is probably related to omission of clouds or cloud shadows in PV-C1 (see §4.3).

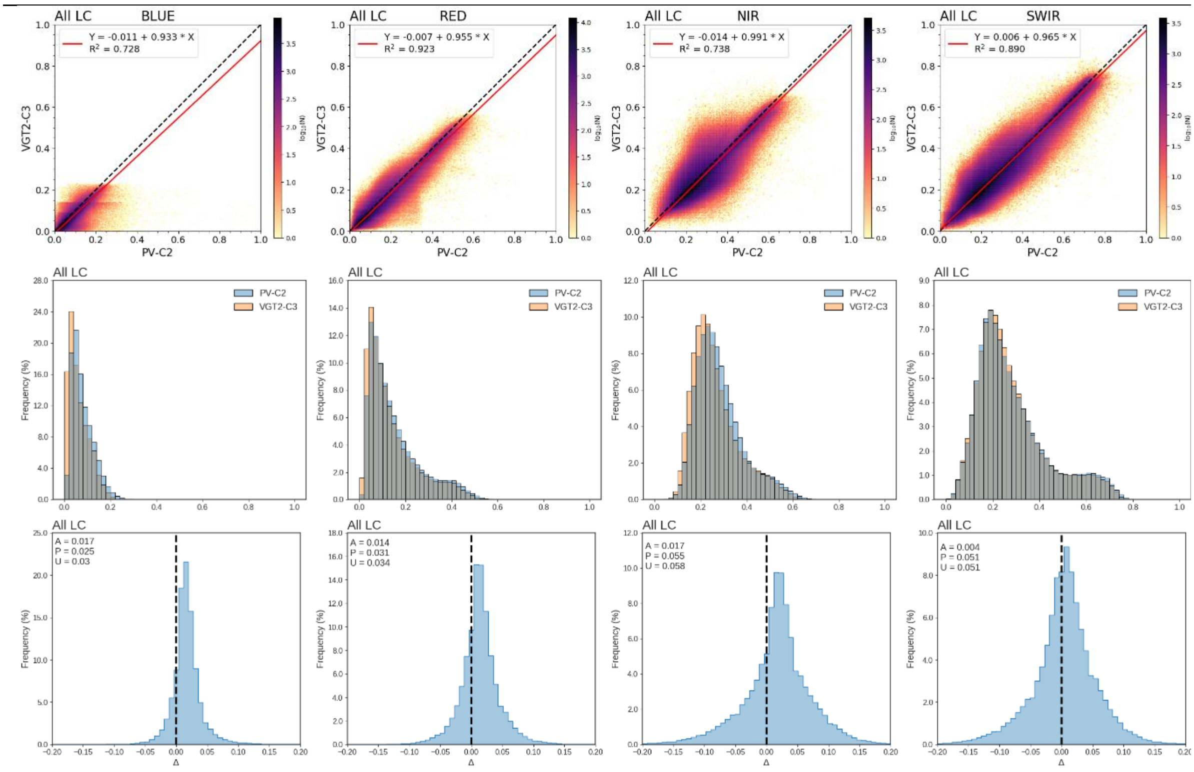


Figure 16: Scatter density plots and GM regression (top), histograms (middle, grey colour indicates where both histograms overlap) and bias histograms with APU metrics (bottom) between PV-C2 (X) and VGT-C3 (Y) S10-TOC reflectances for the overlap period

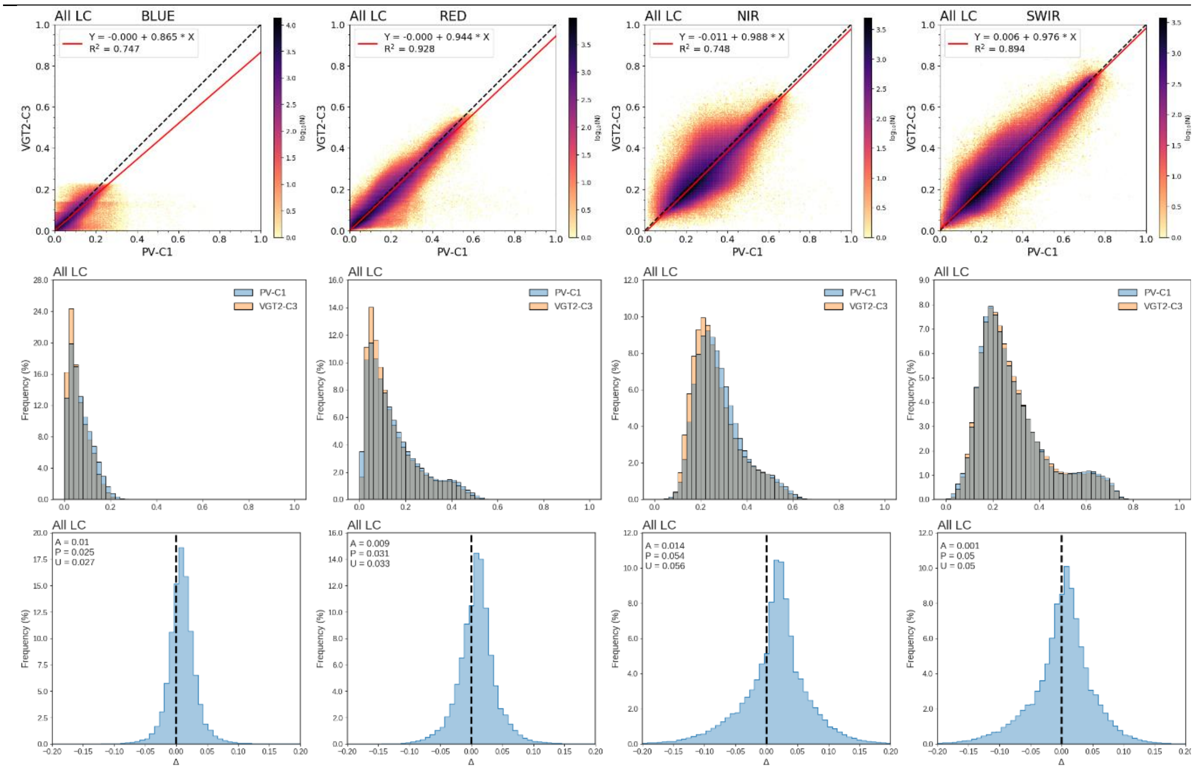


Figure 17: Scatter density plots and GM regression (top), histograms (middle, grey colour indicates where both histograms overlap) and bias histograms with APU metrics (bottom) between PV-C1 (X) and VGT-C3 (Y) S10-TOC reflectances for the overlap period

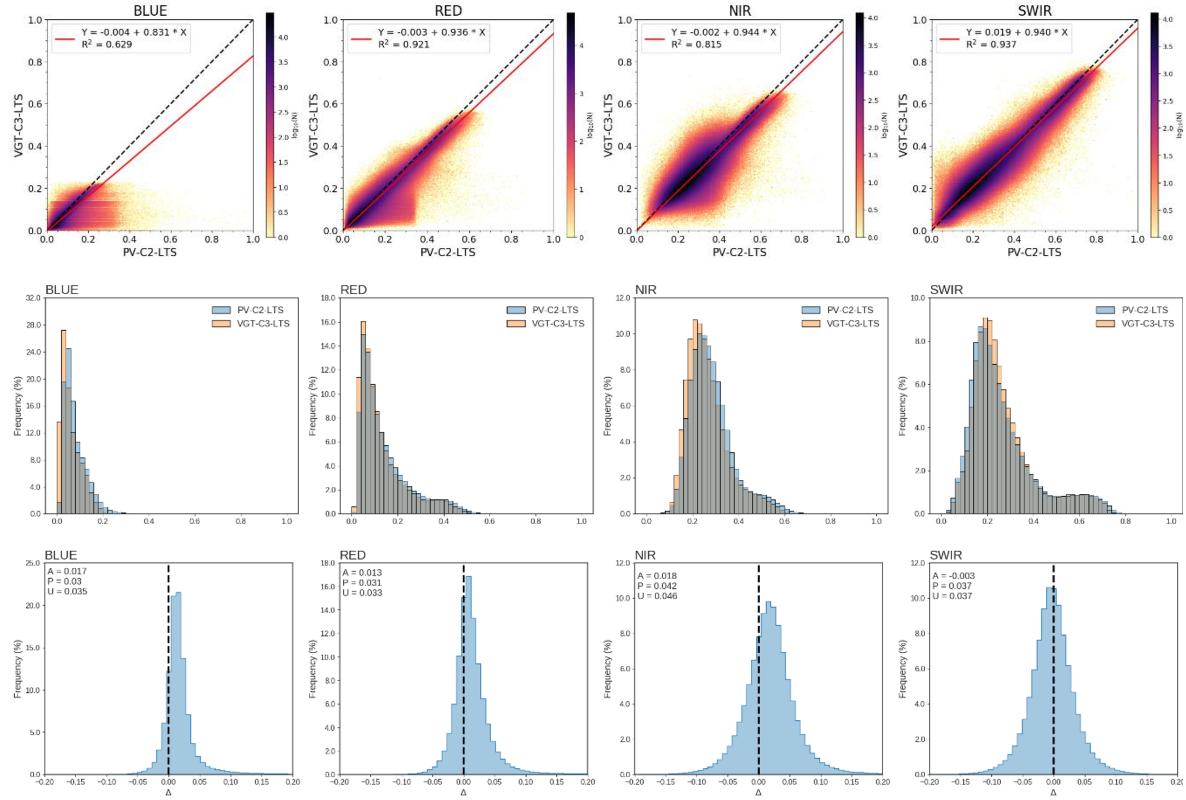


Figure 18: Scatter density plots and GM regression (top), histograms (middle, grey colour indicates where both histograms overlap) and bias histograms with APU metrics (bottom) between PV-C2 LTS (2014-2018) (X) and VGT-C3 LTS (2009-2013) (Y) S10-TOC reflectances

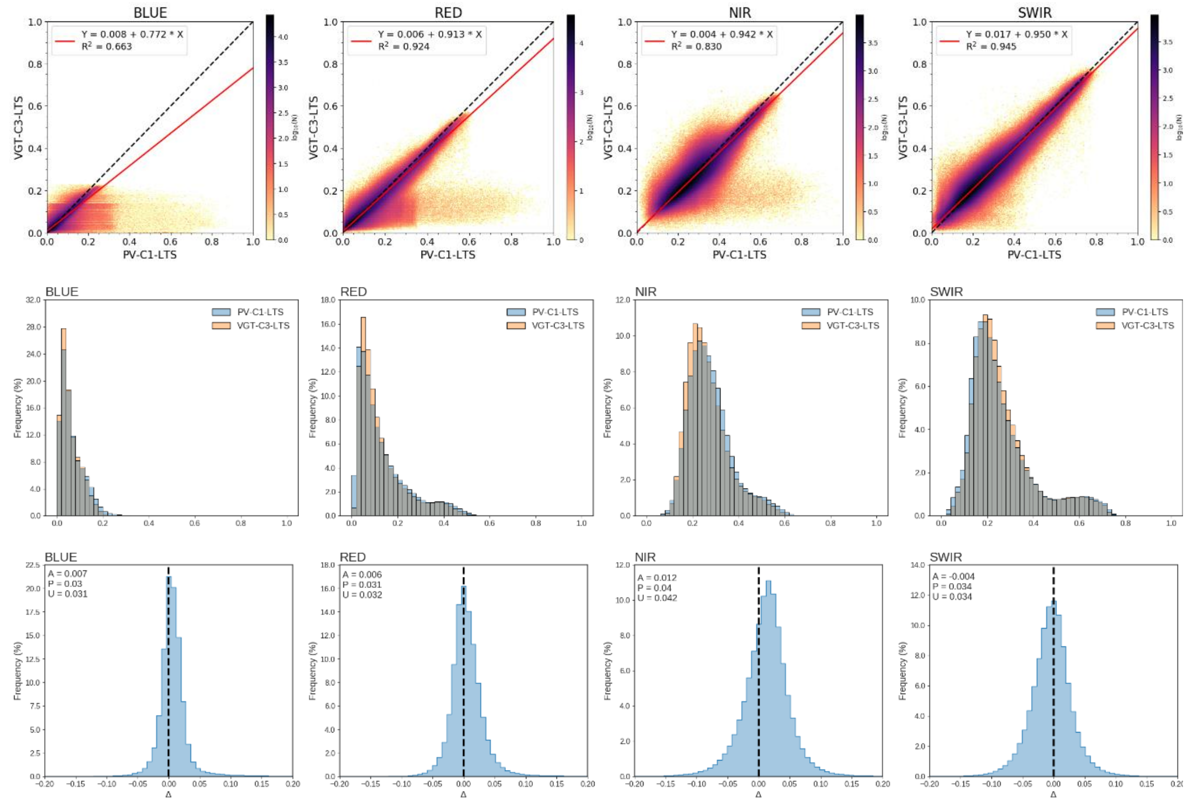


Figure 19: Scatter density plots and GM regression (top), histograms (middle, grey colour indicates where both histograms overlap) and bias histograms with APU metrics (bottom) between PV-C1 LTS (2014-2018) (X) and VGT-C3 LTS (2009-2013) (Y) S10-TOC reflectances

The results for the NDVI are shown in Figure 20. Overall, the correspondence between Proba-V and SPOT/VGT NDVI is quite large, although large scatter is observed. This is related to angular effects, imperfect cloud detection and to the fact that different periods are intercompared. The GM regression line of PV-C2 has intercepts close to zero and slopes above 1. As previously discussed in CHAPTER 5, especially over densely vegetated areas, the reprocessed PV-C2 displays lower NDVI values. This leads to less saturation, where this was the case for PV-C1. While the higher range of NDVI values tends to shift to lower values for PV-C2, the histograms match evenly well for low NDVIs. Where the bias histogram for PV-C1 is skewed towards positive values (PV-C1 NDVI > VGT-C3 NDVI), the opposite counts for PV-C2 (PV-C2 NDVI < VGT-C3 NDVI). The scatterplots of PV-C1 show larger scatter above the 1:1 line, probably related to undetected clouds or cloud shadows in PV-C1.

PV-C1 vs. VGT-C3 (overlap period)	PV-C1 LTS vs. VGT-C3 LTS	PV-C2 vs. VGT-C3 (overlap period)	PV-C2 LTS vs. VGT-C3 LTS

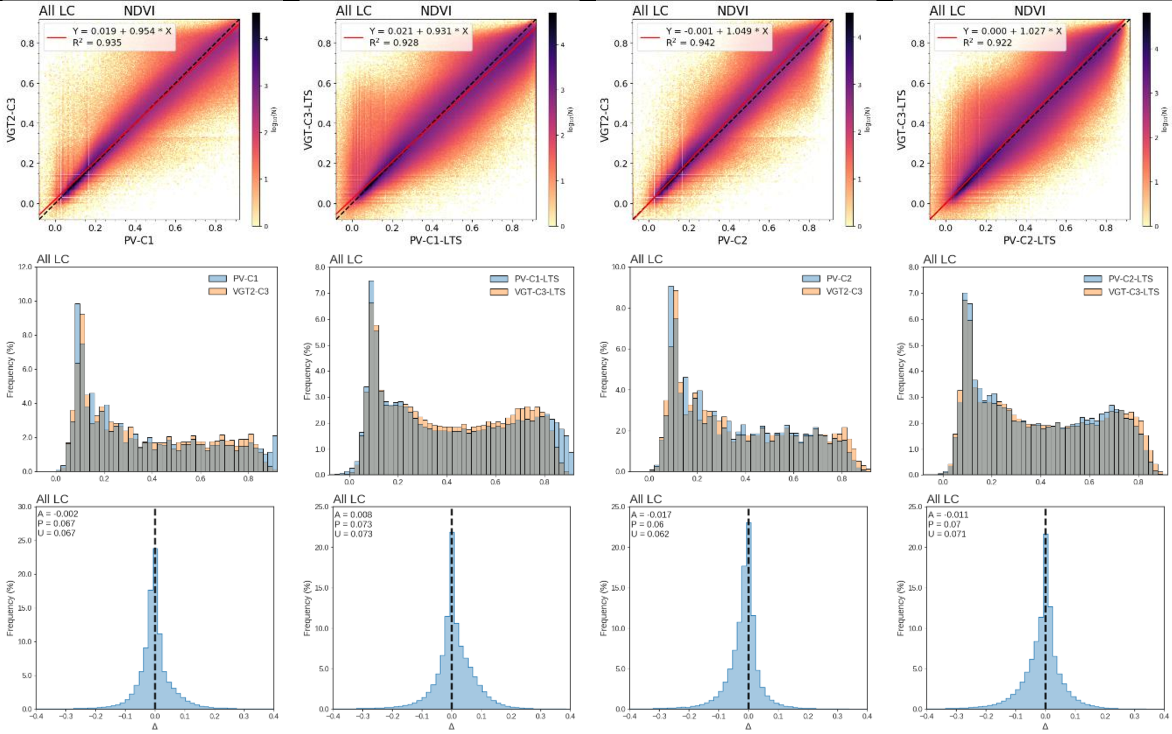


Figure 20: Scatter density plots and GM regression (top), histograms (middle, grey colour indicates where both histograms overlap) and bias histograms with APU metrics (bottom) between PV-C1 (X) and VGT-C3 (Y) (left) and between PV-C2 (X) and VGT-C3 (Y) (right) S10-TOC NDVI for the overlap period

6.2.2. Between Proba-V and S3/SYN-VGT

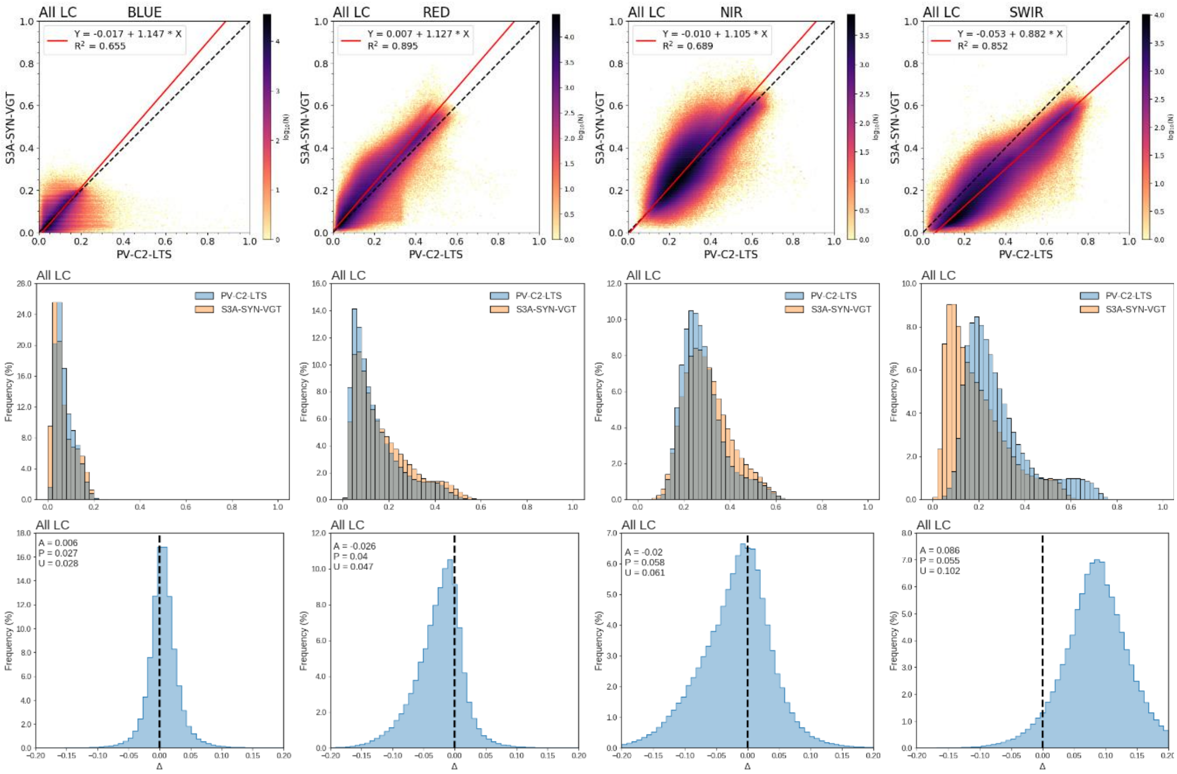
Since there is no overlap between the Proba-V operational phase and S3/SYN-VGT products of good quality, the statistical consistency is evaluated through intercomparison between the LTS of Proba-V based on 2014-2018 (see §3.6) and recent S3A and S3B SYN-VGT data (2022). It is to be noted that the current S3/SYN-VGT products still suffer from a number of quality issues, that have an important impact on product consistency (see §4.4).

The geometric mean regression, histograms, bias histograms and APU metrics are shown in Figure 21 for PV-C2 and Figure 22 for PV-C1. Systematic bias (*Acc*) is highest for the Red, NIR and especially the SWIR band. The latter is caused by SLSTR calibration issues (Smith, 2020). NIR and SWIR consistency is also strongly influenced by a (recently discovered) error in the band mapping procedure (see §4.4). Except for the Blue band, histograms show large deviations, and bias histograms are skewed. Similar results with low consistency for especially SWIR, but also Red and

NIR, are observed for PV-C1, except that more scatter is observed below the 1:1 line, due to erroneous cloud or cloud shadow masking in PV-C1.

The low consistency at surface reflectance level for the Red and NIR bands inevitably leads to limited consistency at the level of the NDVI (Figure 23), with an average bias or *Acc* of 0.04 (for PV-C2) and 0.06 (for PV-C1), *Prec* above 0.08, and *Unc* above 0.09 (for PV-C2) and 0.1 (for PV-C1). The statistical consistency with S3/SYN-VGT is better for PV-C2: although the coefficient of agreement (R^2) is slightly lower, GM regression lines are closer to the 1:1 relation. The histograms also show a more similar shape, especially for the higher NDVI range; the large inconsistencies for the lower NDVI range however remain. Bias histograms are more narrow and less skewed for PV-C2.

PV-C2 LTS (2014-2018) vs. S3A SYN-VGT (2022)



PV-C2 LTS (2014-2018) vs. S3B SYN-VGT (2022)

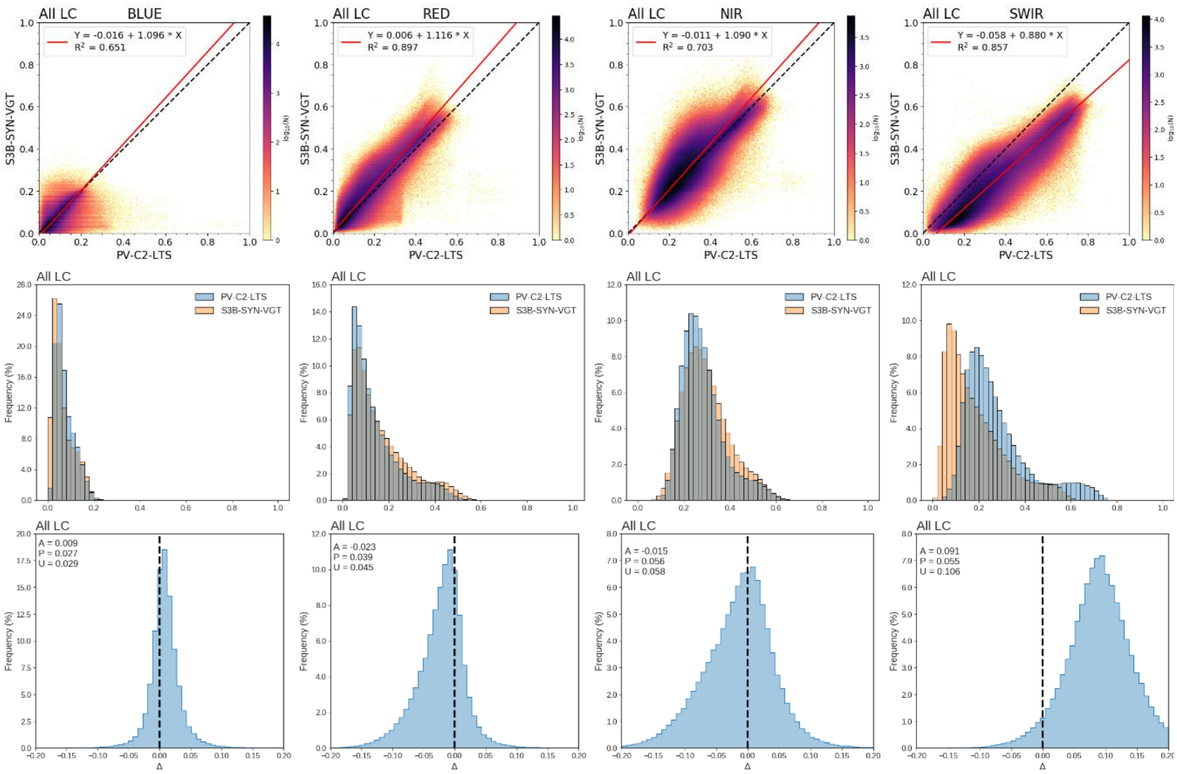


Figure 21: Scatter density plots and GM regression, histograms (grey colour indicates where both histograms overlap) and bias histograms with APU metrics between PV-C2 LTS (2014-2018) (X) and S3A (top) and S3B (bottom) SYN-VGT (2022) (Y) S10-TOC reflectances

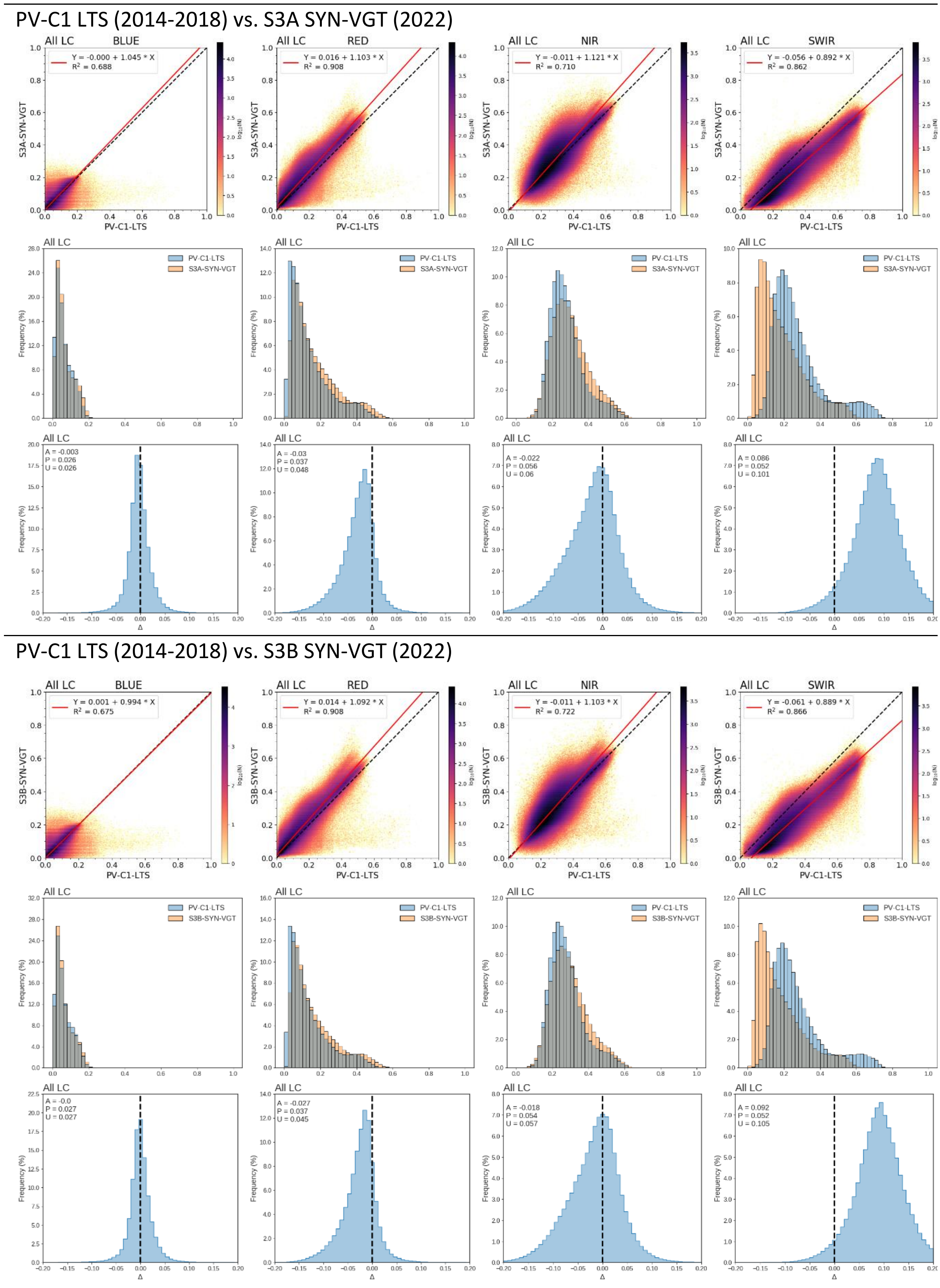


Figure 22: Scatter density plots and GM regression, histograms (grey colour indicates where both histograms overlap) and bias histograms with APU metrics between PV-C1 LTS (2014-2018) (X) and S3A (top) and S3B (bottom) SYN-VGT (2022) (Y) S10-TOC reflectances

PV-C1		PV-C2	
vs. S3A/SYN-VGT	vs. S3B/SYN-VGT	vs. S3A/SYN-VGT	vs. S3B/SYN-VGT

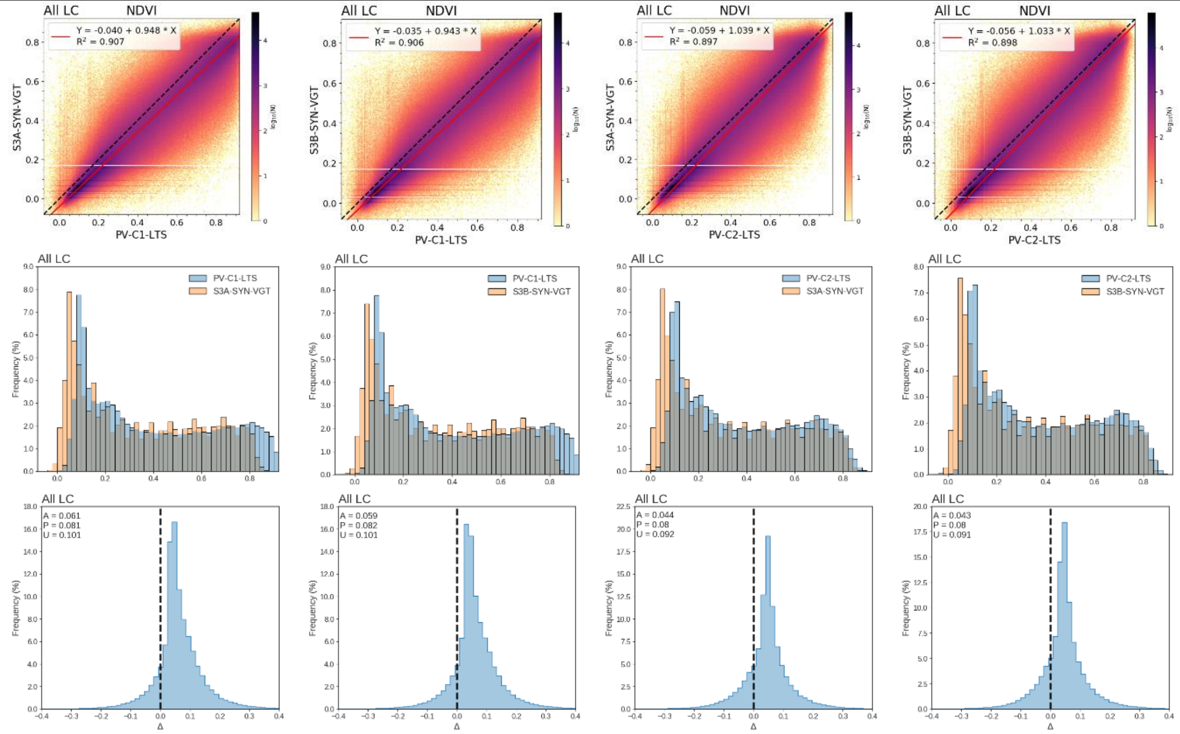


Figure 23: Scatter density plots and GM regression, histograms (grey colour indicates where both histograms overlap) and bias histograms with APU metrics between PV-C1 (left) resp. PV-C2 (left) LTS (2014-2018) (X) and S3A and S3B SYN-VGT (2022) (Y) NDVI

6.3. Spatial pattern of the differences

6.3.1. Between Proba-V and SPOT/VGT

In order to evaluate the spatial pattern of the differences between PV-C2 (and PV-C1) and VGT-C3, the APU metrics are calculated on a global subsample, stratified per biome. The results for the overlap period are shown in Table 9; the results of the comparison of the LTS in Table 10.

For most bands, the metrics show similar values over the different biomes. For PV-C2, the consistency tends to be lower (negative *Acc*, higher *Prec* and *Unc*) for the forest biomes, which is in agreement with the previous results.

6.3.2. Between Proba-V and S3/SYN-VGT

In order to evaluate the spatial pattern of the differences between PV-C2 (and PV-C1) and S3/SYN-VGT, the APU metrics are calculated on a global subsample, stratified per biome. The results for S3A/SYN-VGT are shown in Table 11; the results for S3B/SYN-VGT in Table 12.

For most bands, the metrics show similar values over the different biomes. Nevertheless, it is clear that over all biomes, large inconsistencies are observed for SWIR, NIR and Red, and for the NDVI.

Table 9: APU metrics between PV-C2 (top) resp. PV-C1 (bottom) and VGT-C3 surface reflectances and NDVI, overall and stratified per biome for the overlap period

overlap period										
PV-C2 - VGT-C3		**overall**	**EBF**	**DBF**	**NLF**	**MXF**	**SHR**	**HER**	**CRO**	**BA**
Blue	A	0.017	0.024	0.014	0.020	0.023	0.015	0.016	0.013	0.017
	P	0.025	0.030	0.023	0.035	0.035	0.021	0.023	0.022	0.022
	U	0.030	0.039	0.027	0.040	0.042	0.026	0.028	0.026	0.028
Red	A	0.014	0.010	0.008	0.008	0.011	0.016	0.015	0.011	0.019
	P	0.031	0.027	0.025	0.033	0.031	0.032	0.033	0.026	0.036
	U	0.034	0.029	0.026	0.034	0.033	0.036	0.036	0.029	0.041
NIR	A	0.017	0.006	0.013	0.010	0.016	0.021	0.020	0.015	0.022
	P	0.055	0.085	0.055	0.055	0.053	0.052	0.052	0.051	0.043
	U	0.058	0.085	0.056	0.056	0.055	0.056	0.056	0.054	0.048
SWIR	A	0.004	-0.007	0.000	-0.005	0.000	0.008	0.006	0.004	0.012
	P	0.051	0.058	0.047	0.038	0.036	0.057	0.054	0.050	0.049
	U	0.051	0.058	0.047	0.039	0.036	0.057	0.054	0.050	0.050
NDVI	A	-0.017	-0.043	-0.022	-0.014	-0.022	-0.020	-0.016	-0.018	-0.002
	P	0.060	0.101	0.065	0.088	0.091	0.044	0.046	0.055	0.016
	U	0.062	0.110	0.068	0.089	0.094	0.048	0.049	0.058	0.016
PV-C1 - VGT-C3		**overall**	**EBF**	**DBF**	**NLF**	**MXF**	**SHR**	**HER**	**CRO**	**BA**
Blue	A	0.010	0.011	0.005	0.006	0.006	0.010	0.009	0.007	0.014
	P	0.025	0.035	0.022	0.035	0.031	0.022	0.024	0.021	0.022
	U	0.027	0.036	0.022	0.035	0.031	0.024	0.026	0.022	0.026
Red	A	0.009	-0.001	0.001	0.000	-0.001	0.013	0.012	0.007	0.016
	P	0.031	0.032	0.024	0.035	0.029	0.031	0.032	0.027	0.033
	U	0.033	0.032	0.024	0.035	0.029	0.034	0.034	0.027	0.036
NIR	A	0.014	0.008	0.012	0.002	0.006	0.018	0.017	0.013	0.018
	P	0.054	0.088	0.055	0.055	0.053	0.051	0.051	0.051	0.039
	U	0.056	0.089	0.056	0.055	0.053	0.054	0.053	0.052	0.043
SWIR	A	0.001	-0.008	-0.002	-0.008	-0.005	0.004	0.002	0.000	0.006
	P	0.050	0.058	0.047	0.041	0.039	0.055	0.052	0.048	0.044
	U	0.050	0.059	0.047	0.042	0.039	0.055	0.052	0.048	0.044
NDVI	A	-0.002	0.014	0.006	0.015	0.013	-0.013	-0.009	-0.008	-0.002
	P	0.067	0.123	0.074	0.106	0.105	0.047	0.051	0.059	0.016
	U	0.067	0.123	0.074	0.107	0.106	0.049	0.051	0.060	0.016

Table 10: APU metrics between PV-C2 (top) resp. PV-C1 (bottom) LTS (2014-2018) and VGT-C3 LTS (2009-2013) surface reflectances and NDVI, overall and stratified per biome for the overlap period

comparison of LTS										
PV-C2 - VGT-C3		**overall**	**EBF**	**DBF**	**NLF**	**MXF**	**SHR**	**HER**	**CRO**	**BA**
Blue	A	0.017	0.023	0.016	0.028	0.031	0.014	0.016	0.013	0.014
	P	0.030	0.025	0.030	0.048	0.047	0.024	0.029	0.024	0.021
	U	0.035	0.034	0.035	0.056	0.056	0.028	0.033	0.028	0.025
Red	A	0.013	0.008	0.009	0.017	0.017	0.012	0.012	0.009	0.017
	P	0.031	0.022	0.028	0.044	0.042	0.029	0.031	0.026	0.029
	U	0.033	0.024	0.030	0.047	0.046	0.031	0.034	0.027	0.034
NIR	A	0.018	0.009	0.016	0.018	0.020	0.018	0.018	0.018	0.023
	P	0.042	0.052	0.041	0.050	0.050	0.038	0.040	0.041	0.033
	U	0.046	0.053	0.045	0.053	0.054	0.042	0.044	0.045	0.040
SWIR	A	-0.003	-0.011	-0.008	-0.008	-0.007	-0.002	-0.003	-0.004	0.008
	P	0.037	0.036	0.031	0.027	0.026	0.039	0.041	0.038	0.036
	U	0.037	0.037	0.032	0.028	0.027	0.039	0.041	0.039	0.037
NDVI	A	-0.011	-0.031	-0.011	-0.024	-0.032	-0.010	-0.009	-0.004	0.002
	P	0.070	0.084	0.073	0.098	0.099	0.062	0.064	0.075	0.022
	U	0.071	0.090	0.074	0.101	0.104	0.063	0.065	0.075	0.022
PV-C1 - VGT-C3		**overall**	**EBF**	**DBF**	**NLF**	**MXF**	**SHR**	**HER**	**CRO**	**BA**
Blue	A	0.007	0.009	0.003	0.009	0.006	0.006	0.006	0.004	0.012
	P	0.030	0.028	0.024	0.045	0.039	0.022	0.033	0.030	0.021
	U	0.031	0.029	0.024	0.046	0.040	0.022	0.033	0.030	0.024
Red	A	0.006	-0.003	-0.002	0.002	-0.002	0.008	0.008	0.004	0.015
	P	0.031	0.025	0.023	0.042	0.036	0.027	0.035	0.031	0.025
	U	0.032	0.025	0.024	0.042	0.036	0.028	0.035	0.031	0.030
NIR	A	0.012	0.007	0.011	0.004	0.003	0.013	0.013	0.015	0.017
	P	0.040	0.052	0.038	0.046	0.045	0.034	0.041	0.042	0.028
	U	0.042	0.052	0.039	0.046	0.045	0.037	0.043	0.044	0.032
SWIR	A	-0.004	-0.012	-0.008	-0.009	-0.010	-0.003	-0.004	-0.005	0.005
	P	0.034	0.034	0.028	0.028	0.027	0.035	0.038	0.034	0.031
	U	0.034	0.036	0.030	0.029	0.029	0.035	0.039	0.035	0.031
NDVI	A	0.008	0.022	0.022	0.020	0.023	-0.001	0.000	0.008	-0.003
	P	0.073	0.096	0.072	0.110	0.108	0.062	0.064	0.073	0.022
	U	0.073	0.099	0.075	0.112	0.110	0.062	0.064	0.074	0.022

Table 11: APU metrics between PV-C2 (top) resp. PV-C1 (bottom) LTS (2014-2018) and S3A/SYN-VGT surface reflectances and NDVI, overall and stratified per biome

PV-C2 LTS - S3A/SYN-VGT		overall	EBF	DBF	NLF	MXF	SHR	HER	CRO	BA
Blue	A	0.006	0.015	0.009	0.009	0.007	0.005	0.006	0.009	-0.002
	P	0.027	0.030	0.026	0.034	0.036	0.022	0.028	0.027	0.024
	U	0.028	0.033	0.028	0.035	0.037	0.022	0.028	0.028	0.024
Red	A	-0.026	-0.006	-0.013	-0.008	-0.009	-0.031	-0.030	-0.028	-0.036
	P	0.040	0.028	0.028	0.033	0.031	0.042	0.043	0.040	0.039
	U	0.047	0.029	0.031	0.034	0.032	0.052	0.053	0.049	0.053
NIR	A	-0.020	-0.023	-0.017	-0.011	-0.009	-0.026	-0.023	-0.029	-0.010
	P	0.058	0.071	0.056	0.053	0.056	0.059	0.060	0.065	0.042
	U	0.061	0.074	0.059	0.054	0.057	0.064	0.064	0.072	0.043
SWIR	A	0.086	0.088	0.083	0.062	0.062	0.082	0.085	0.081	0.104
	P	0.055	0.036	0.037	0.037	0.032	0.059	0.060	0.055	0.062
	U	0.102	0.095	0.091	0.072	0.070	0.101	0.104	0.097	0.121
NDVI	A	0.044	0.013	0.045	0.041	0.044	0.049	0.051	0.048	0.045
	P	0.080	0.100	0.086	0.105	0.104	0.075	0.078	0.098	0.023
	U	0.092	0.101	0.097	0.112	0.113	0.089	0.093	0.109	0.051
PV-C1 LTS - S3A/SYN-VGT										
Blue	A	-0.003	0.001	-0.002	-0.005	-0.009	-0.003	-0.004	-0.001	-0.004
	P	0.026	0.029	0.023	0.031	0.030	0.022	0.027	0.027	0.024
	U	0.026	0.029	0.023	0.031	0.032	0.022	0.027	0.027	0.024
Red	A	-0.030	-0.017	-0.021	-0.018	-0.021	-0.033	-0.033	-0.032	-0.037
	P	0.037	0.029	0.025	0.029	0.026	0.040	0.041	0.039	0.037
	U	0.048	0.034	0.033	0.034	0.034	0.051	0.053	0.050	0.052
NIR	A	-0.022	-0.022	-0.018	-0.015	-0.013	-0.029	-0.026	-0.031	-0.014
	P	0.056	0.070	0.054	0.050	0.054	0.057	0.057	0.064	0.039
	U	0.060	0.073	0.057	0.052	0.056	0.064	0.063	0.071	0.042
SWIR	A	0.086	0.087	0.085	0.063	0.065	0.082	0.085	0.082	0.103
	P	0.052	0.034	0.034	0.035	0.031	0.056	0.057	0.050	0.060
	U	0.101	0.093	0.091	0.073	0.072	0.100	0.103	0.096	0.120
NDVI	A	0.061	0.069	0.077	0.088	0.098	0.054	0.059	0.062	0.041
	P	0.081	0.105	0.084	0.105	0.100	0.074	0.078	0.098	0.025
	U	0.101	0.126	0.114	0.137	0.140	0.092	0.098	0.116	0.048

Table 12: APU metrics between PV-C2 (top) resp. PV-C1 (bottom) LTS (2014-2018) and S3B/SYN-VGT surface reflectances and NDVI, overall and stratified per biome

PV-C2 LTS - S3B/SYN-VGT		overall	EBF	DBF	NLF	MXF	SHR	HER	CRO	BA
Blue	A	0.009	0.016	0.011	0.010	0.007	0.007	0.008	0.011	0.004
	P	0.027	0.030	0.027	0.035	0.036	0.022	0.028	0.026	0.024
	U	0.029	0.034	0.029	0.036	0.037	0.023	0.029	0.028	0.024
Red	A	-0.023	-0.006	-0.011	-0.007	-0.008	-0.027	-0.026	-0.024	-0.033
	P	0.039	0.029	0.029	0.034	0.031	0.041	0.042	0.038	0.040
	U	0.045	0.030	0.031	0.034	0.032	0.049	0.050	0.045	0.052
NIR	A	-0.015	-0.017	-0.010	-0.006	-0.003	-0.021	-0.018	-0.023	-0.008
	P	0.056	0.069	0.055	0.052	0.054	0.057	0.058	0.063	0.043
	U	0.058	0.071	0.055	0.052	0.054	0.061	0.061	0.067	0.044
SWIR	A	0.091	0.091	0.088	0.064	0.065	0.089	0.092	0.088	0.108
	P	0.055	0.036	0.036	0.036	0.031	0.058	0.059	0.053	0.062
	U	0.106	0.098	0.095	0.074	0.072	0.106	0.109	0.103	0.125
NDVI	A	0.043	0.017	0.045	0.042	0.046	0.046	0.049	0.045	0.043
	P	0.080	0.102	0.085	0.104	0.100	0.075	0.077	0.096	0.024
	U	0.091	0.103	0.096	0.112	0.110	0.088	0.091	0.107	0.049
PV-C1 LTS - S3B/SYN-VGT										
Blue	A	0.000	0.002	-0.001	-0.005	-0.008	0.000	-0.001	0.001	0.002
	P	0.027	0.030	0.023	0.032	0.032	0.022	0.028	0.028	0.024
	U	0.027	0.030	0.023	0.032	0.033	0.022	0.028	0.028	0.024
Red	A	-0.027	-0.017	-0.019	-0.017	-0.019	-0.029	-0.029	-0.029	-0.034
	P	0.037	0.030	0.025	0.031	0.027	0.039	0.041	0.038	0.037
	U	0.045	0.034	0.031	0.035	0.033	0.048	0.050	0.047	0.051
NIR	A	-0.018	-0.016	-0.011	-0.011	-0.007	-0.024	-0.021	-0.025	-0.012
	P	0.054	0.069	0.052	0.049	0.052	0.055	0.056	0.062	0.040
	U	0.057	0.071	0.053	0.050	0.053	0.060	0.059	0.067	0.042
SWIR	A	0.092	0.090	0.090	0.066	0.068	0.090	0.092	0.089	0.108
	P	0.052	0.034	0.034	0.035	0.031	0.056	0.056	0.049	0.060
	U	0.105	0.096	0.096	0.075	0.074	0.106	0.108	0.102	0.124
NDVI	A	0.059	0.072	0.076	0.088	0.096	0.051	0.056	0.059	0.038
	P	0.082	0.107	0.083	0.107	0.101	0.075	0.077	0.097	0.025
	U	0.101	0.129	0.112	0.138	0.139	0.091	0.096	0.114	0.046

CHAPTER 7 COMPARISON TO EXTERNAL DATASETS

7.1. INTRODUCTION

This chapter focuses on the consistency of Proba-V C2 with external datasets, focusing on the NDVI. In first instance, the consistency between Proba-V C2 NDVI and external datasets (LSA-SAF ENVI10, see §4.5, and MOD13A3 NDVI, see §4.6) is evaluated for the entire operational Proba-V period. We assess the magnitude of the differences (i.e. statistical consistency) and the spatial pattern of the differences (i.e. spatial consistency).

Over a longer time period (2009-2022), the combined NDVI series of SPOT/VGT, Proba-V and S3/SYN-VGT is compared to the external datasets. Although there are intrinsic differences between datasets derived from different sensors (linked to differences in spectral response functions, calibration, processing, etc.), the temporal behaviour of the differences provide useful information on product stability.

The evaluation is done for the S10-TOC 1 km NDVI at global scale on a systematic subsample (see §3.3.1), except for the temporal plots generated over LANDVAL sites (see §3.3.2). The comparison with MODIS is performed on monthly MVC NDVI.

The key questions that are answered in this chapter are:

Q3.1 *What is the statistical consistency between Proba-V C2 NDVI and external data?*
Q3.2 *What is the spatial pattern of the differences?*
Q3.3 *What is the spatio-temporal evolution of the differences between a combined NDVI series from SPOT/VGT C3, Proba-V C2 and S3/SYN-VGT and external data?*
Q3.4 *What is the temporal consistency between a combined NDVI series from SPOT/VGT C3, Proba-V C2 and S3/SYN-VGT and external data over LANDVAL sites?*

7.2. MAGNITUDE OF THE DIFFERENCES

The scatterplots, geometric mean regression, histograms, bias histograms and APU metrics for the comparison between ENDVI10 and PV-C2 NDVI (Figure 24) indicate a systematic bias (*Acc*) of 0.02 between ENDVI10 and PV-C2 NDVI (ENDVI10 > PV-C2 NDVI). Nevertheless, the coefficient of determination is above 0.9 and the regression line slope is almost 1. The histograms show a similar shape, but PV-C2 NDVI is shifted towards lower values. The bias histogram shows some skewedness towards negative values, and the scatterplot shows more scatter above the 1:1 line, probably related to cloud omission in the ENDVI10 product. Some uncertainty is related to the geometric accuracy of MetOp-B and the pixel shift that was applied in the data extractions (see §4.5).

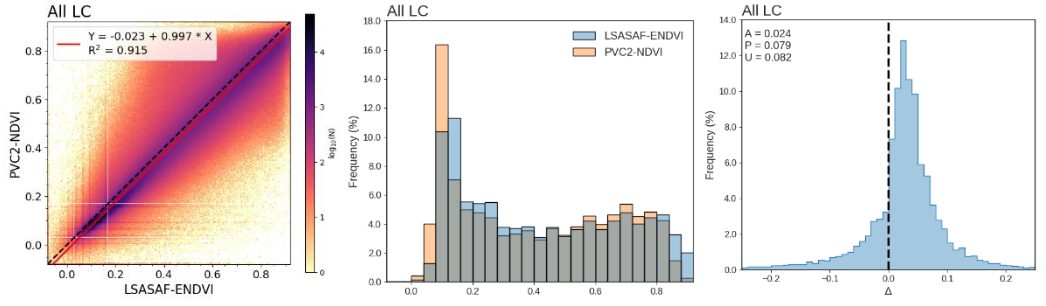


Figure 24: Scatter density plot and GM regression, histogram (grey colour indicates where both histograms overlap) and bias histogram with APU metrics between LSASAF ENDVI10 (X) and PV-C2 (Y) S10-TOC NDVI

The intercomparison with MOD13A3 NDVI shows larger statistical consistency, with lower systematic bias between PV-C2 NDVI and MOD13A3 NDVI (*Acc* = 0.005). The bias histogram peaks around zero. The histograms show a more similar shape, and the same peak at low NDVI values. The intercept is also closer to zero compared to the analysis with ENDVI10. For the higher NDVI range, MOD13A3 NDVI is higher. The scatterplot seems to show a slightly non-linear relation, possibly related to differences in spectral response between both sensors (see Annex A) and a difference in compositing method (see §4.6). Also some asymmetric scattering is observed below the 1:1 line, possibly due to cloud omission in the PV-C2 data.

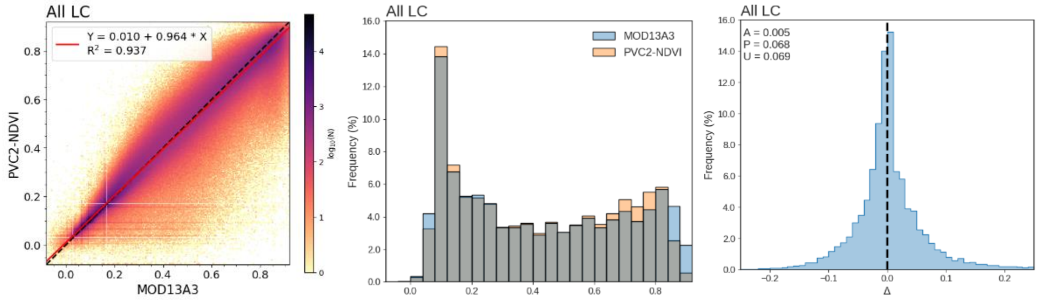


Figure 25: Scatter density plot and GM regression, histogram (grey colour indicates where both histograms overlap) and bias histogram with APU metrics between MOD13A3 NDVI (X) and PV-C2 (Y) monthly MVC NDVI

7.3. SPATIAL PATTERN OF THE DIFFERENCES

Figure 26 and Figure 27 show the statistical consistency analysis between respectively ENVI10 and MOD13A3 NDVI versus PV-C2 NDVI, stratified per biome.

A similar positive *Acc* between ENDVI10 and PV-C2 NDVI is observed over all biomes (except EBF), which means that the spatial pattern of the bias is homogeneous. These biomes also show a clear linear relationship. All histograms have the same shift between both datasets. Except for BA, the bias histograms all show some skewedness towards negative values, related to insufficient cloud screening in the ENDVI10 product.

The intercomparison with MOD13A3 shows similar shapes in the histograms over all biomes, although for some forest biomes PV-C2 NDVI seem to have a longer tail towards lower NDVI values. Except for EBF, the histograms peak around the same value. Average bias (*Acc*) fluctuates around 0.01 for most biomes, except the EBF (*Acc* = 0.05), NLF (*Acc* = 0.04) and MXF (*Acc* = 0.03). The spatial pattern of the bias is thus not homogeneous. Although bias histograms peak close to zero, some skewedness towards negative values is observed over all biomes, except BA, possibly due to cloud omission in the PV-C2 data.

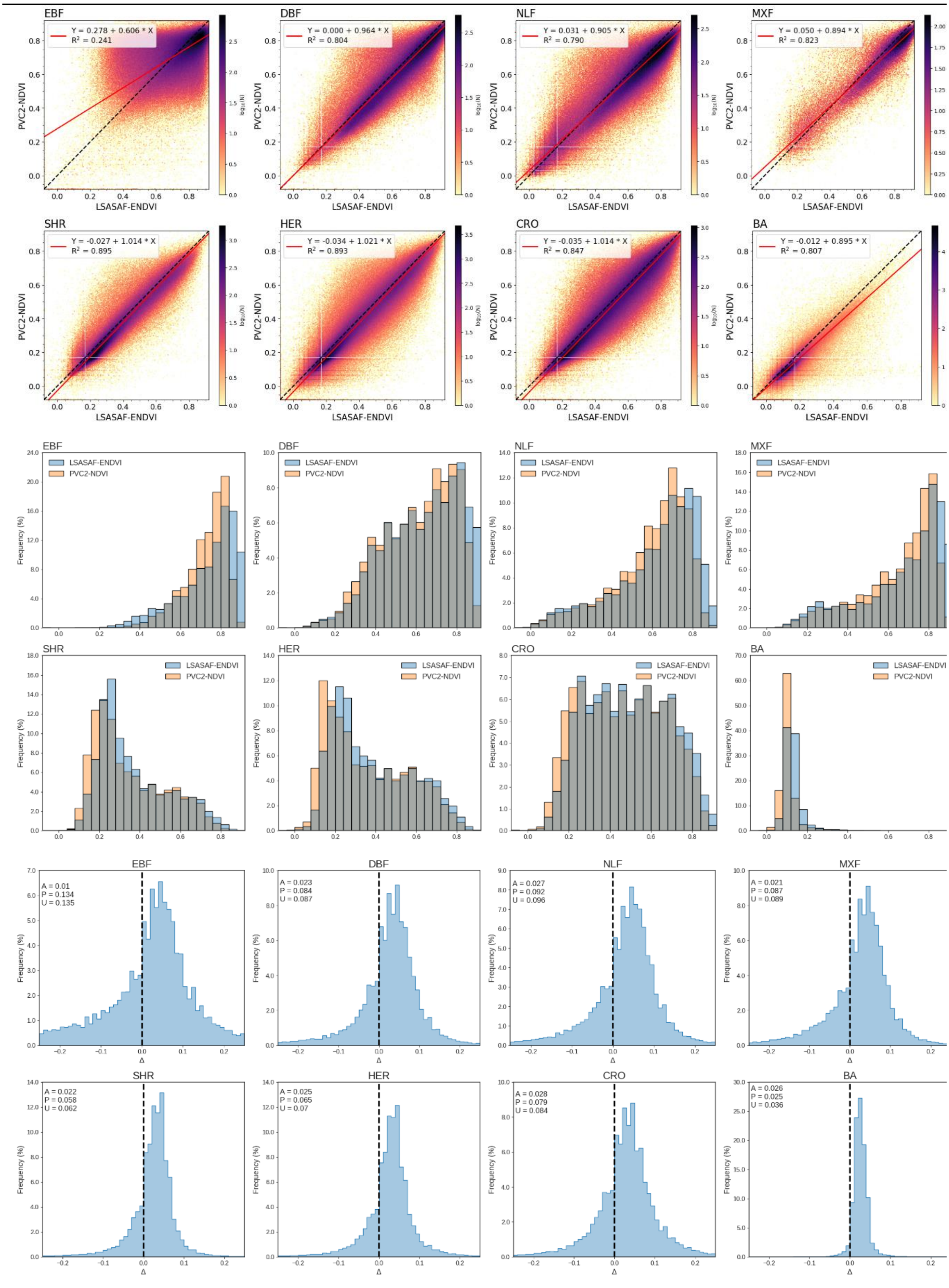


Figure 26: Scatter density plots and GM regression (top), histograms (middle, grey colour indicates where both histograms overlap) and bias histograms with APU metrics (bottom) between LSASAF ENDVI10 (X) and PV-C2 (Y) S10-TOC NDVI, stratified per biome

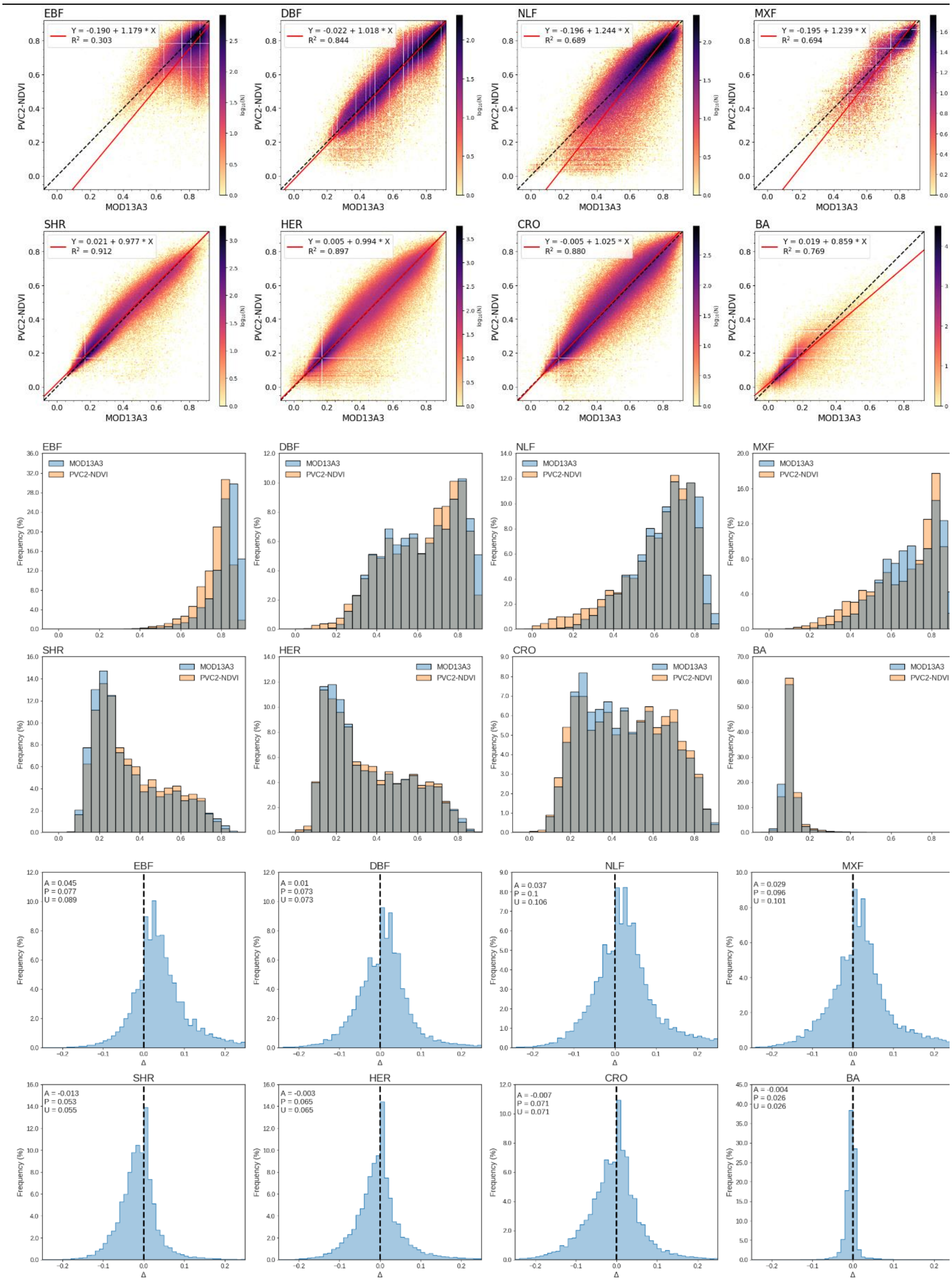


Figure 27: Scatter density plots and GM regression (top), histograms (middle, grey colour indicates where both histograms overlap) and bias histograms with APU metrics (bottom) between MOD13A3 NDVI (X) and PV-C2 (Y) monthly MVC NDVI, stratified per biome

7.4. SPATIO-TEMPORAL EVOLUTION OF THE DIFFERENCES

In order to evaluate the spatio-temporal evolution of the differences, Hovmöller plots (see §3.7) are generated for the intercomparison of the combined NDVI series of SPOT/VGT-C3 (2009-2013), Proba-V C2 (2014-June/2020) and S3A/SYN-VGT (July/2020-2022), hereafter called VGT-PV-S3, with LSA-SAF ENDVI10 and MOD13A3 NDVI.

Figure 28 shows the spatio-temporal evolution of the mean NDVI per time step and per latitude band. All plots show a clear seasonal variation, and higher vegetation densities around the equator and in summer periods at higher latitudes, and low NDVI between 20°N and 30°N (dominated by the Sahara region). The plot for ENDVI10 shows lower product completeness in both the northern and southern hemisphere winter periods. Product coverage of the combined VGT-PV-C3 series is more comparable to MOD13A3.

Whereas the plots for ENDVI10 and for MOD13A3 exhibit strong temporal stability, this is not the case for the combined VGT-PV-S3 series: the switch from VGT-C3 to PV-C2 is hardly visible, but a strong discontinuity is observed at the switch to S3/SYN-VGT (July/2020). This is because until May/2021, the S3/SYN-VGT products were erroneously based on TOA reflectances. Also for the more recent period, there are a number of quality issues that affect the consistency with the VGT-C3 and PV-C2 archives, leading to lower S3/SYN-VGT NDVI values (see also CHAPTER 6 and §4.4). The combined VGT-PV-S3 data archive thus not provide the required spatio-temporal consistency over the entire time series under investigation.

This is also obvious from the spatio-temporal evolution of APU metrics of the intercomparison of the VGT-PV-C3 series with the external datasets. The S3/SYN-VGT period shows high *Acc* and *Unc*, especially for the period July/2020 – May/2021. But also after this, there is high bias with both ENDVI10 and MOD13A3 NDVI. This means there is no continuity after Proba-V with the S3/SYN-VGT products.

The Hovmöller plots of *Acc* also show discontinuities at the switch from SPOT/VGT to Proba-V (January/2014), related to differences in atmospheric correction inputs in Proba-V C2, resulting in lower consistency with SPOT/VGT-C3 (see also CHAPTER 6). There is an abrupt difference in overpass time between VGT-C3 and Proba-V (see Annex B). A gradual effect towards lower values of *Acc* is visible in the last years of SPOT/VGT, probably related to the sensor drift (see Annex B) and the associated change in illumination conditions. A similar pattern is slightly noticeable in the last years of Proba-V. TOC NDVI is known to increase with an increase in solar zenith angles (Deering et al., 1999), because a higher fraction of solar radiation is intercepted by the vegetation canopy (Nagol et al., 2014). However, although the impact of orbital drift on Proba-V surface reflectances was shown to be heterogeneous along the across-track swath, the asymmetry in the Red and NIR bands is largely smoothed out in the NDVI ratio (Niro, 2021). While the *Prec* remains roughly stable over the SPOT/VGT and Proba-V period for the comparison with MOD13A3, an increase of *Prec* of the comparison with ENDVI10 is visible for the Proba-V period.

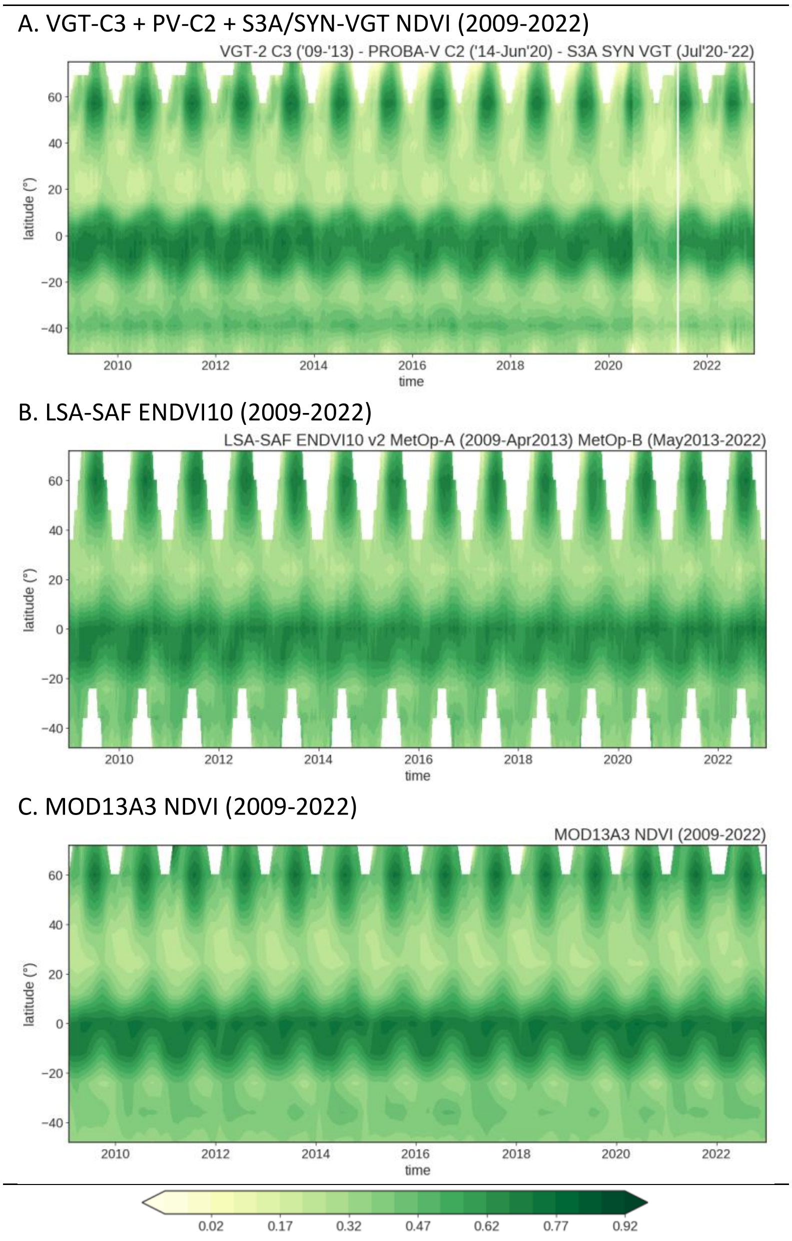


Figure 28: Hovmöller plots of the mean NDVI of A. the combined series of VGT-C3 (2009-2013), PV-C2 (2014-Jun/2020) and S3A/SYN-VGT (Jul/2020-2022); B. LSASAF-ENDVI10 (2009-2022); and C. MOD13A3 NDVI (2009-2022)

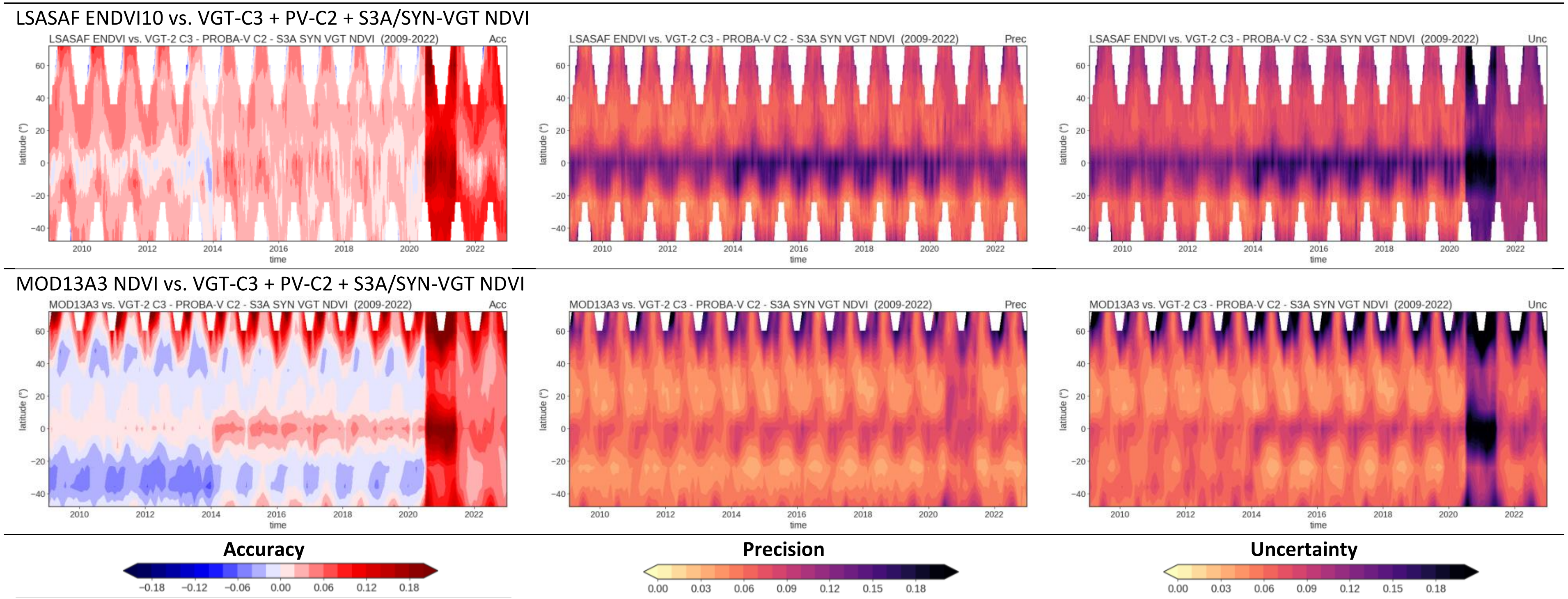


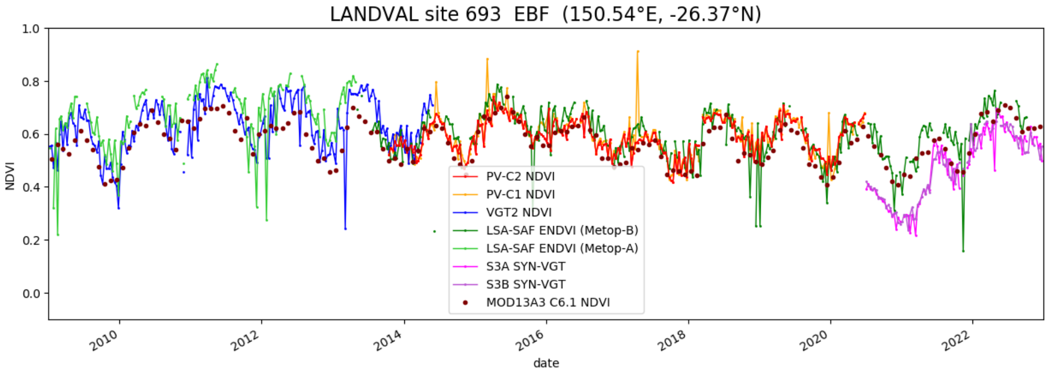
Figure 29: Hovmöller plots of the APU metrics between resp. ENDVI10 (top) and MOD13A3 NDVI (bottom) and the combined NDVI series of VGT-C3 (2009-2013), PV-C2 (2014-Jun/2020) and S3A/SYN-VGT (Jul/2020-2022)

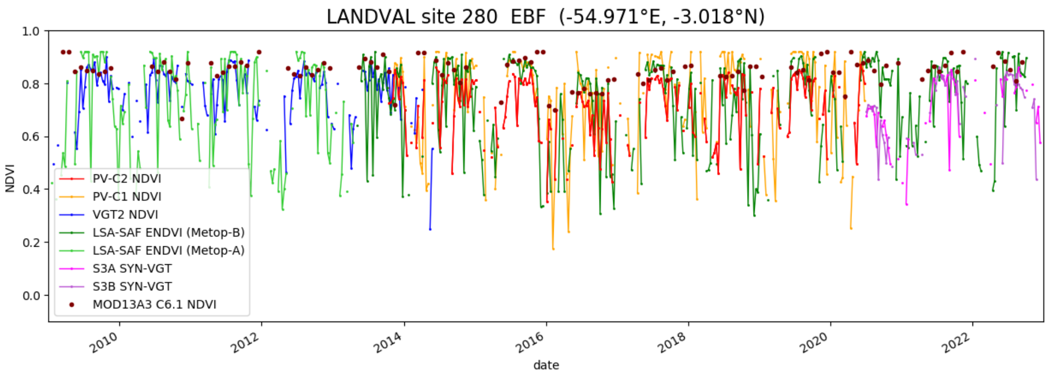
7.5. TEMPORAL CONSISTENCY OVER LANDVAL SITES

A selection (2 per biome, see §3.3.3) of temporal NDVI plots of PV-C2, PV-C1, VGT-C3, S3/SYN-VGT, ENDVI10 and MOD13A3 NDVI over the LANDVAL sites (see §3.3.2) is illustrated in Figure 30. All plots, over all LANDVAL sites, are available as a digital annex to this report (see Annex F).

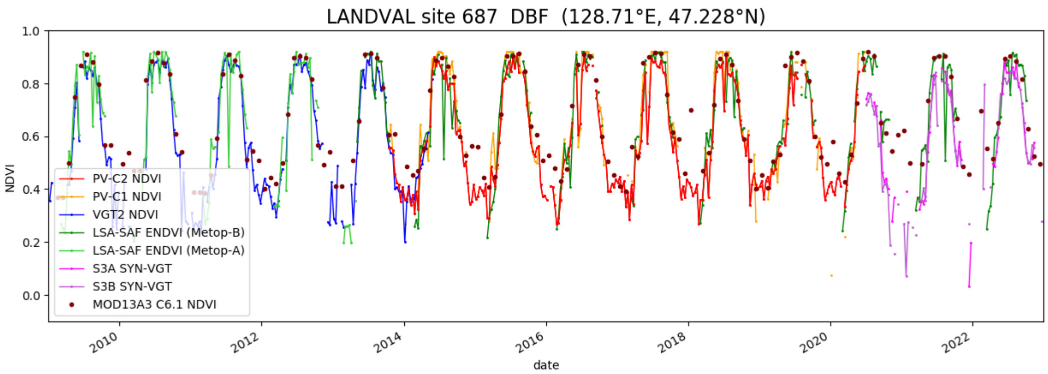
The temporal profiles again show the large inconsistency with S3/SYN-VGT NDVI, both with the SPOT/VGT and Proba-V archives, and the external datasets. All other datasets show large overall temporal consistency, with similar timing and amplitude of fluctuations in vegetation density. A clear offset (or systematic bias) is in some cases visible between ENDVI10 and the other datasets. The (monthly) MOD13A3 profiles show less residual cloud contamination, especially e.g. over EBF. The 10-daily datasets show high frequency variations, probably related to residual cloud cover and anisotropy effects. Some spurious high NDVI values in PV-C1 NDVI disappear in PV-C2. Sometimes spurious values are observed in the MOD13A3 series, e.g. during winter periods for MXF sites, or for BA sites.

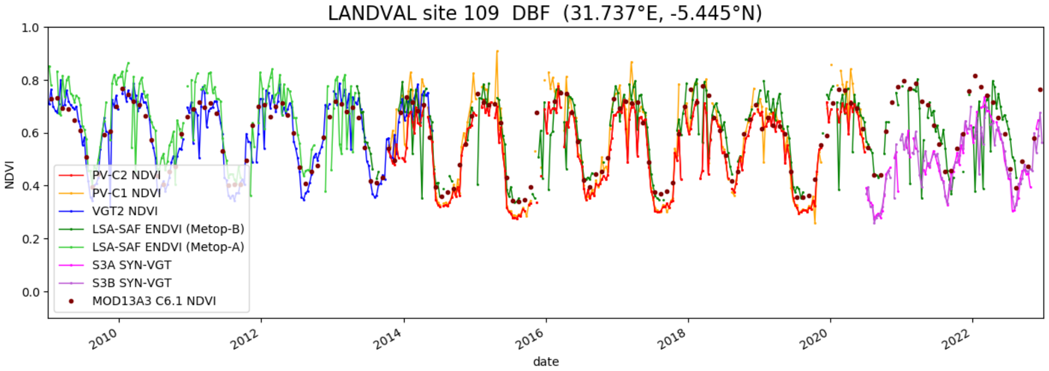
EBF



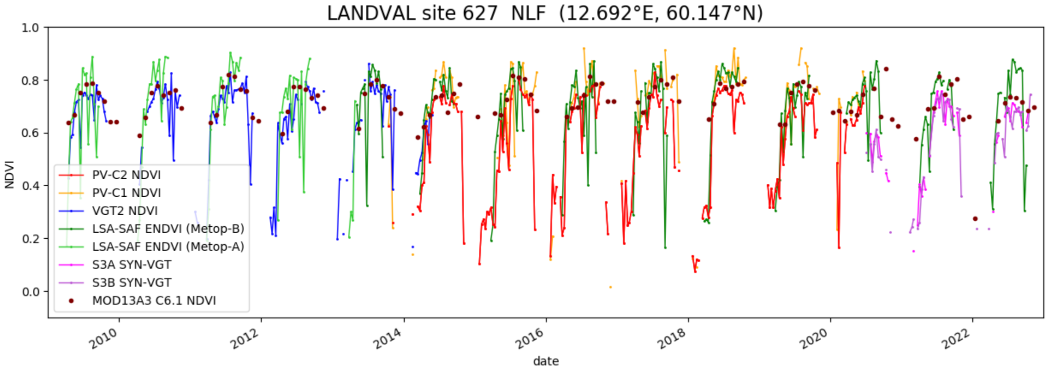


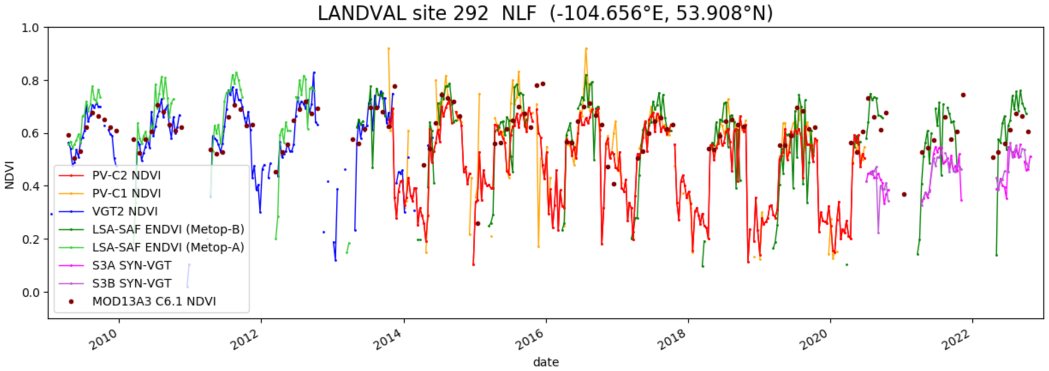
DBF



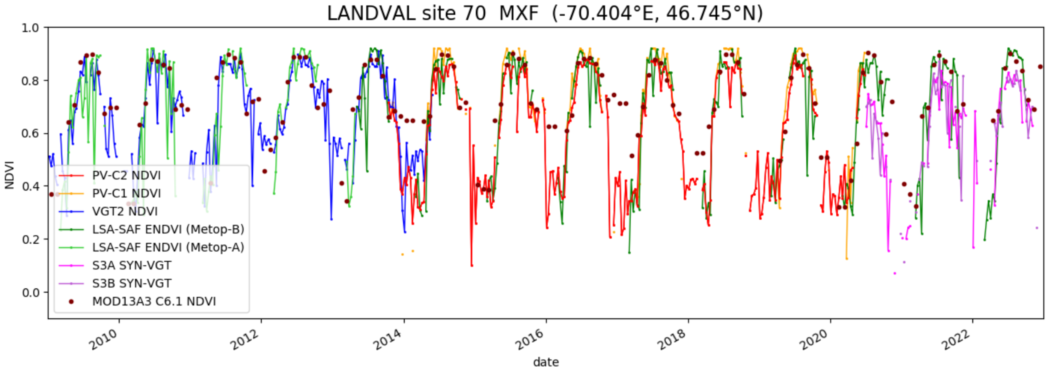

LANDVAL site 109 DBF (31.737°E, -5.445°N)
NDVI
date
PV-C2 NDVI
PV-C1 NDVI
VGT2 NDVI
LSA-SAF ENDVI (Metop-B)
LSA-SAF ENDVI (Metop-A)
S3A SYN-VGT
S3B SYN-VGT
MOD13A3 C6.1 NDVI

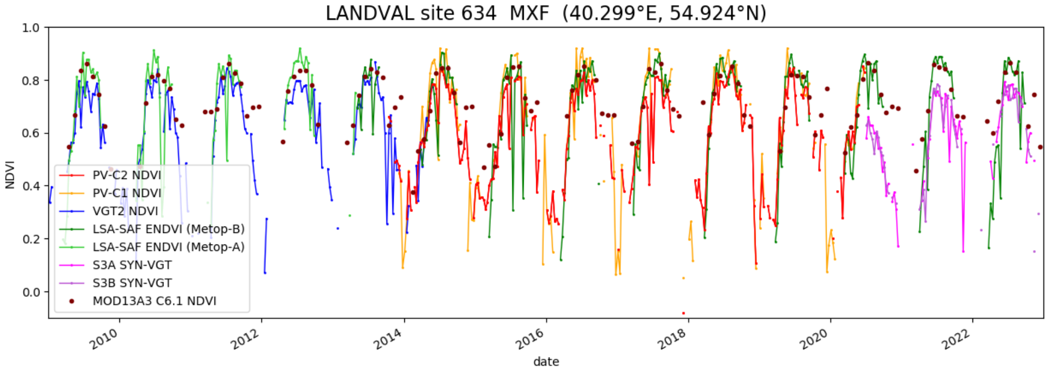
NLF


LANDVAL site 627 NLF (12.692°E, 60.147°N)
NDVI
date
PV-C2 NDVI
PV-C1 NDVI
VGT2 NDVI
LSA-SAF ENDVI (Metop-B)
LSA-SAF ENDVI (Metop-A)
S3A SYN-VGT
S3B SYN-VGT
MOD13A3 C6.1 NDVI

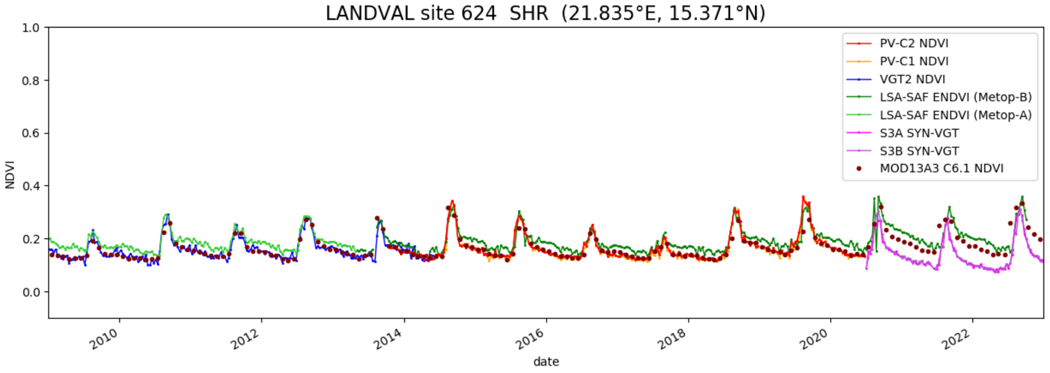

LANDVAL site 292 NLF (-104.656°E, 53.908°N)
NDVI
date
PV-C2 NDVI
PV-C1 NDVI
VGT2 NDVI
LSA-SAF ENDVI (Metop-B)
LSA-SAF ENDVI (Metop-A)
S3A SYN-VGT
S3B SYN-VGT
MOD13A3 C6.1 NDVI

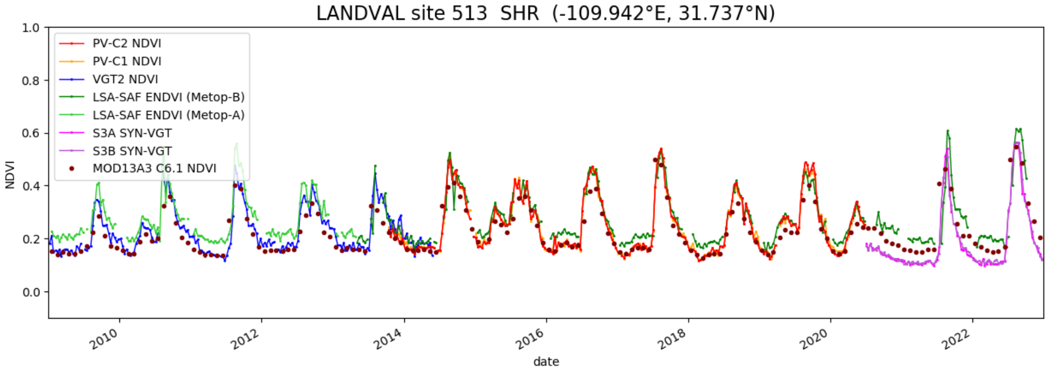
MXF


LANDVAL site 70 MXF (-70.404°E, 46.745°N)
NDVI
date
PV-C2 NDVI
PV-C1 NDVI
VGT2 NDVI
LSA-SAF ENDVI (Metop-B)
LSA-SAF ENDVI (Metop-A)
S3A SYN-VGT
S3B SYN-VGT
MOD13A3 C6.1 NDVI

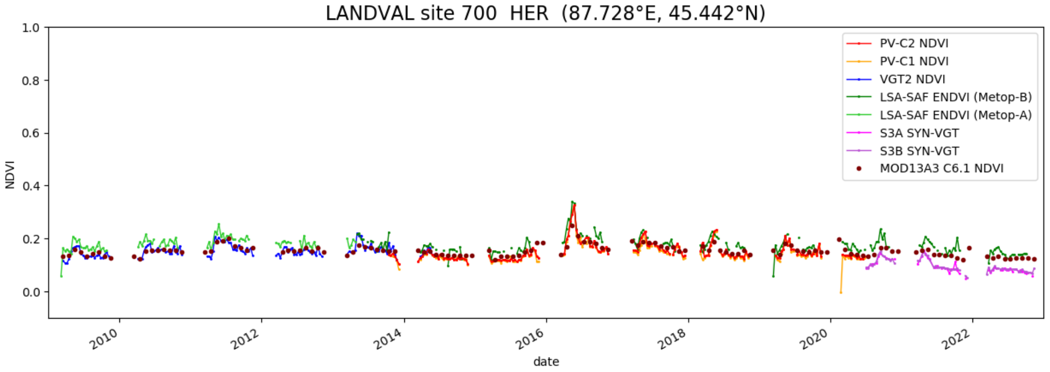

LANDVAL site 634 MXF (40.299°E, 54.924°N)
NDVI
date
PV-C2 NDVI
PV-C1 NDVI
VGT2 NDVI
LSA-SAF ENDVI (Metop-B)
LSA-SAF ENDVI (Metop-A)
S3A SYN-VGT
S3B SYN-VGT
MOD13A3 C6.1 NDVI

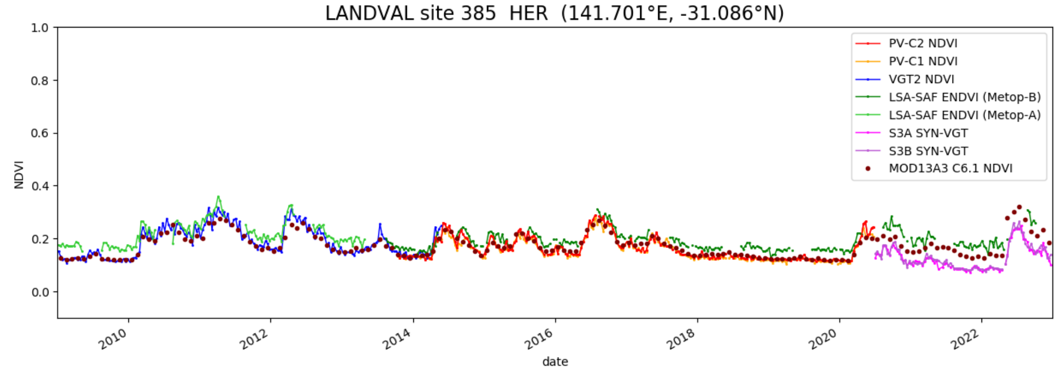
SHR


LANDVAL site 624 SHR (21.835°E, 15.371°N)
NDVI
date
PV-C2 NDVI
PV-C1 NDVI
VGT2 NDVI
LSA-SAF ENDVI (Metop-B)
LSA-SAF ENDVI (Metop-A)
S3A SYN-VGT
S3B SYN-VGT
MOD13A3 C6.1 NDVI

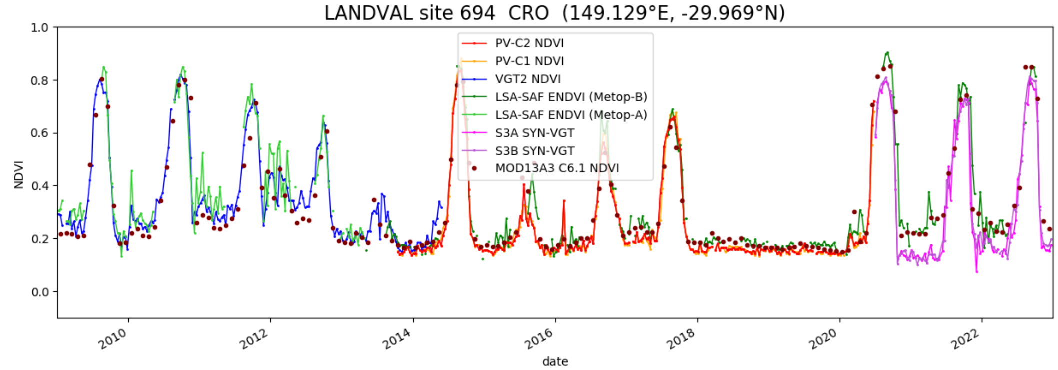

LANDVAL site 513 SHR (-109.942°E, 31.737°N)
NDVI
date
PV-C2 NDVI
PV-C1 NDVI
VGT2 NDVI
LSA-SAF ENDVI (Metop-B)
LSA-SAF ENDVI (Metop-A)
S3A SYN-VGT
S3B SYN-VGT
MOD13A3 C6.1 NDVI

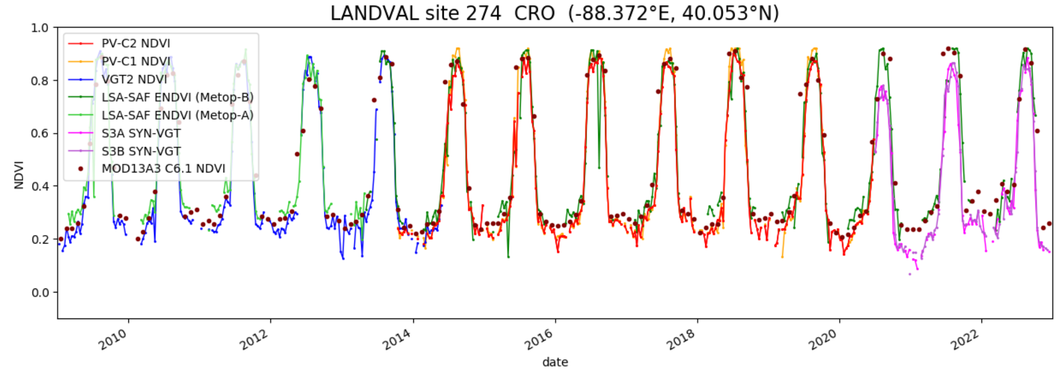
HER


LANDVAL site 700 HER (87.728°E, 45.442°N)
NDVI
date
PV-C2 NDVI
PV-C1 NDVI
VGT2 NDVI
LSA-SAF ENDVI (Metop-B)
LSA-SAF ENDVI (Metop-A)
S3A SYN-VGT
S3B SYN-VGT
MOD13A3 C6.1 NDVI

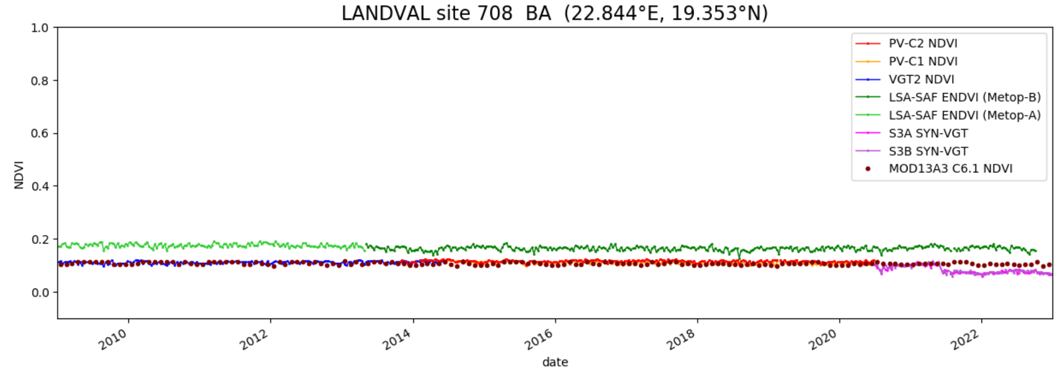

LANDVAL site 385 HER (141.701°E, -31.086°N)
PV-C2 NDVI
PV-C1 NDVI
VGT2 NDVI
LSA-SAF ENDVI (Metop-B)
LSA-SAF ENDVI (Metop-A)
S3A SYN-VGT
S3B SYN-VGT
MOD13A3 C6.1 NDVI
NDVI
date

CRO


LANDVAL site 694 CRO (149.129°E, -29.969°N)
PV-C2 NDVI
PV-C1 NDVI
VGT2 NDVI
LSA-SAF ENDVI (Metop-B)
LSA-SAF ENDVI (Metop-A)
S3A SYN-VGT
S3B SYN-VGT
MOD13A3 C6.1 NDVI
NDVI
date


LANDVAL site 274 CRO (-88.372°E, 40.053°N)
PV-C2 NDVI
PV-C1 NDVI
VGT2 NDVI
LSA-SAF ENDVI (Metop-B)
LSA-SAF ENDVI (Metop-A)
S3A SYN-VGT
S3B SYN-VGT
MOD13A3 C6.1 NDVI
NDVI
date

BA


LANDVAL site 708 BA (22.844°E, 19.353°N)
PV-C2 NDVI
PV-C1 NDVI
VGT2 NDVI
LSA-SAF ENDVI (Metop-B)
LSA-SAF ENDVI (Metop-A)
S3A SYN-VGT
S3B SYN-VGT
MOD13A3 C6.1 NDVI
NDVI
date

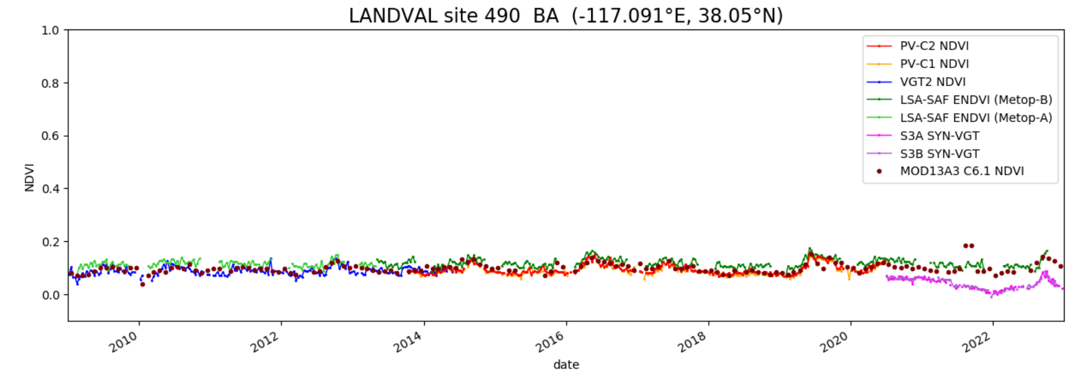


Figure 30: Selection of NDVI profiles over LANDVAL sites comparing PV-C2, PV-C1, VGT-C3, S3/SYN-VGT, ENDVI10 and MOD13A3 NDVI over 2009-2022

CHAPTER 8 CONCLUSIONS

This report presents the outcomes of the evaluation done after the reprocessing campaign of the Proba-V C1 archive, resulting in Proba-V C2. The most important modifications in the Proba-V processing chain from C1 to C2, with largest impact on the product content, are related to updated radiometric calibration, an improved cloud detection algorithm and improved atmospheric correction.

The evaluation is based on analysis over the entire operational phase of Proba-V, i.e. more than 7.5 years of S1-TOA (at 1 km resolution), S5-TOC (at 100 m resolution) and S10-TOC (at 300 m and 1 km resolution) data at global scale. Data was subsampled by taking a systematic spatial subsample and by extracting information over the LANDVAL sites. Analyses included consistency checks with SPOT/VGT, S3/SYN-VGT and external datasets from MetOp/AVHRR and Terra/MODIS.

In this concluding section, a summarized answer to a number of key questions is formulated:

1. <u>**On the comparison of the reprocessed archive (C2) with the previous archive (C1) over the entire operational phase of Proba-V**</u>

→ What is the completeness of the data, in terms of spatial pattern and temporal evolution? What is the difference in flag occurrences between Proba-V C1 and C2?

Proba-V C2 shows more clear observations, because less pixels are flagged as cloud. Slightly more pixels are detected as snow/ice. Overall, this leads to a general increase of 'good' (clear pixels with good radiometric quality) pixels, at all resolutions. Spatial patterns of quality flag occurrences are similar for C1 and C2, with slightly higher amounts of good observations in C2, especially at high latitudes and in tropical regions. Also temporal patterns are similar for C1 and C2, with largest differences but similar seasonal variations for cloud occurrences.

→ What is the magnitude of the difference between Proba-V C2 and C1? What is the spatial and temporal pattern of the difference?

Differences at TOA level are very small, and in line with updates to the radiometric calibration. The overall temporal patterns match with the temporal patterns of changes in the absolute calibration coefficients. At the level of surface reflectance, the overall differences are larger. Largest bias is observed for Blue and Red bands, related to the adaptations to the atmospheric correction inputs. The average bias is opposite for Red and NIR, leading to lower NDVI in C2. This is most pronounced over areas and periods with more dense vegetation cover.

2. <u>**On the comparison of Proba-V C2 with related datasets derived from SPOT/Vegetation and Sentinel-3/SYN-VGT**</u>

→ What is the statistical consistency of Proba-V C2 with SPOT/VGT C3 and Sentinel-3 SYN-VGT? What is the magnitude of the difference (i) between Proba-V and SPOT/VGT C3 for the

overlapping period, (ii) between the LTS of Proba-V and SPOT/VGT C3, and between the LTS of Proba-V and recent data of S3/SYN-VGT?

Overall, there is high correspondence between Proba-V and SPOT/VGT. In contrast, there is high systematic bias between Proba-V and S3/SYN-VGT, especially for the SWIR, but also for Red and NIR bands. This leads to low consistency for the NDVI. The current S3/SYN-VGT products still suffer from a number of quality issues, that have an important impact on product consistency (see §4.4). A consistent archive can only be achieved when these quality issues are solved and a complete reprocessing of the S3/SYN-VGT archive is performed.

→ **What is the spatial consistency of Proba-V C2 with SPOT/VGT C3 and S3/SYN-VGT?**

Proba-V and SPOT/VGT show slightly less consistency over forest biomes. The systematic low consistency between Proba-V and S3/SYN-VGT for SWIR, NIR and Red is observed over all biomes.

→ **How do the results of Proba-V C2 compare with those of Proba-V C1?**

Proba-V C1 shows slightly better consistency with SPOT/VGT, especially for the Blue band, related to a more similar atmospheric correction scheme. Also on the consistency with S3/SYN-VGT similar results are observed for Proba-V C2 and C1, with C2 showing slightly better results for NDVI.

3. <u>On the comparison to reference time series from external datasets</u>

→ **What is the statistical consistency between Proba-V C2 NDVI and external data?**

Proba-V C2 shows a strong relation with LSA-SAF ENDVI10, although Proba-V NDVI is systematically lower. Large consistency with MOD13A3 NDVI is observed, with a slight non-linear behaviour.

→ **What is the spatial pattern of the differences?**

The spatial pattern of the bias between Proba-V C2 and LSA-SAF ENDVO10 is homogeneous, with similar bias observed over (nearly) all biomes. Histograms over forest biomes show some deviations between Proba-V C2 and MOD13A3 NDVI.

→ **What is the spatio-temporal evolution of the differences between a combined NDVI series from SPOT/VGT C3, Proba-V C2 and S3/SYN-VGT and external data?**

The combined NDVI series of SPOT/VGT-C3 (2009-2013), Proba-V C2 (2014-June/2020) and S3A/SYN-VGT (July/2020-2022) shows a strong discontinuity at the switch to S3/SYN-VGT. High bias with LSA-SAF ENDVI10 and MOD13A3 NDVI is observed for the S3/SYN-VGT period. Discontinuities are also observed at the switch from SPOT/VGT to Proba-V, related to an abrupt change in overpass time and differences in atmospheric correction. Small gradual effects are also visible in the last years of SPOT/VGT and Proba-V, related to sensor drift. The combined VGT-PV-S3 data archive thus not provide the required spatio-temporal consistency over the entire time series under investigation, especially – but not only – related to quality issues in the S3/SYN-VGT data.

→ **What is the temporal consistency between a combined NDVI series from SPOT/VGT C3, Proba-V C2 and S3/SYN-VGT and external data over LANDVAL sites?**

The temporal profiles show large inconsistencies between S3/SYN-VGT NDVI and both the SPOT/VGT and Proba-V archives and the external datasets. Overall, all other datasets show large temporal consistency. In some cases a systematic bias is observed between ENDVI10 and the other datasets.

In summary, the Proba-V reprocessing campaign was successful, yielding the expected impacts in terms of product completeness, and differences with the previous products. Proba-V C2 products show large consistency with the SPOT/VGT-C3 data archive and external datasets (except S3/SYN-VGT). Since the current S3/SYN-VGT products still suffer from important quality issues, users are advised not to use these products in combination with Proba-V or SPOT/Vegetation products.

Users are strongly recommended to update their Proba-V archive with Collection 2 (.V2 in the file naming).

References

Buchhorn, M., Smets, B., Bertels, L., Lesiv, M., Tsendbazar, N.-E., Herold, M., Fritz, S., 2019. Copernicus Global Land Service: Land Cover 100m: epoch 2015: Globe. https://doi.org/10.5281/zenodo.3939038

Deering, D.W., Eck, T.F., Banerjee, B., 1999. Characterization of the reflectance anisotropy of three Boreal forest canopies in spring-summer. Remote Sens. Environ. 67, 205–229. https://doi.org/10.1016/S0034-4257(98)00087-X

Deronde, B., Debruyn, W., Gontier, E., Goor, E., Jacobs, T., Verbeiren, S., Vereecken, J., 2014. 15 years of processing and dissemination of SPOT-VEGETATION products. Int. J. Remote Sens. 35, 2402–2420. https://doi.org/10.1080/01431161.2014.883102

Dierckx, W., Sterckx, S., Benhadj, I., Livens, S., Duhoux, G., Van Achteren, T., Francois, M., Mellab, K., Saint, G., 2014. PROBA-V mission for global vegetation monitoring: standard products and image quality. Int. J. Remote Sens. 35, 2589–2614. https://doi.org/10.1080/01431161.2014.883097

Dierckx, W., Swinnen, E., Kempeneers, P., 2015. Validation of spectral continuity between PROBA-V and SPOT-VEGETATION global daily datasets. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch. 40, 1155–1162. https://doi.org/10.5194/isprsarchives-XL-7-W3-1155-2015

Duveiller, G., Fasbender, D., Meroni, M., 2016. Revisiting the concept of a symmetric index of agreement for continuous datasets. Sci. Rep. 6, 1–14. https://doi.org/10.1038/srep19401

Fell, F., Bennartz, R., Loew, A., 2015. Validation of the EUMETSAT Geostationary Surface Albedo Climate Data Record -2- (ALBEDOVAL-2).

Francois, M., Santandrea, S., Mellab, K., Vrancken, D., Versluys, J., 2014. The PROBA-V mission: the space segment. Int. J. Remote Sens. 35, 2548–2564. https://doi.org/10.1080/01431161.2014.883098

Fuster, B., Sánchez-Zapero, J., Camacho, F., García-Santos, V., Verger, A., Lacaze, R., Weiss, M., Baret, F., Smets, B., 2020. Quality Assessment of PROBA-V LAI, fAPAR and fCOVER Collection 300 m Products of Copernicus Global Land Service. Remote Sens. 12, 1017. https://doi.org/10.3390/rs12061017

Galvao, L.S., Ponzoni, F.J., Epiphanio, J.C.N., Rudorff, B.F.T., Formaggio, A.R., 2004. Sun and view angle effects on NDVI determination of land cover types in the Brazilian Amazon region with hyperspectral data. Int. J. Remote Sens. 25, 1861–1879. https://doi.org/10.1080/01431160310001598908

Gelaro, R., McCarty, W., Suárez, M.J., Todling, R., Molod, A., Takacs, L., Randles, C.A., Darmenov, A., Bosilovich, M.G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A.M., Gu, W., Kim, G.K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J.E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S.D., Sienkiewicz, M., Zhao, B., 2017. The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). J. Clim. 30, 5419–5454. https://doi.org/10.1175/JCLI-D-16-0758.1

Korras-Carraca, M.B., Gkikas, A., Matsoukas, C., Hatzianastassiou, N., 2021. Global clear-sky aerosol speciated direct radiative effects over 40 years (1980–2019). Atmosphere (Basel). 12, 1–30. https://doi.org/10.3390/atmos12101254

Loew, A., Bennartz, R., Fell, F., Lattanzio, A., Doutriaux-Boucher, M., Schulz, J., 2016. A database of global reference sites to support validation of satellite surface albedo datasets (SAVS 1.0). Earth Syst. Sci. Data 8, 425–438. https://doi.org/10.5194/essd-8-425-2016

Meroni, M., Atzberger, C., Vancutsem, C., Gobron, N., Baret, F., Lacaze, R., Eerens, H., Leo, O., 2013. Evaluation of Agreement Between Space Remote Sensing SPOT-VEGETATION fAPAR Time

Series. IEEE Trans. Geosci. Remote Sens. 51, 1951–1962. https://doi.org/10.1109/TGRS.2012.2212447

Meyer, D.J., 1996. Estimating the effective spatial resolution of an AVHRR time series. Int. J. Remote Sens. 17, 2971–2980. https://doi.org/10.1080/01431169608949122

Nagol, J., Vermote, E., Prince, S., 2014. Quantification of Impact of Orbital Drift on Inter-Annual Trends in AVHRR NDVI Data. Remote Sens. 6, 6680–6687. https://doi.org/10.3390/rs6076680

Niro, F., 2021. Evaluation of Orbital Drift Effect on Proba-V Surface Reflectances Time Series. Remote Sens. 13, 2250. https://doi.org/10.3390/rs13122250

Rahman, H., Dedieu, G., 1994. SMAC: a simplified method for the atmospheric correction of satellite measurements in the solar spectrum. Int. J. Remote Sens. 15, 123–143. https://doi.org/10.1080/01431169408954055

Sellers, P.J.J., 1985. Canopy reflectance, photosynthesis and transpiration. Int. J. Remote Sens. 6, 1335–1372. https://doi.org/10.1080/01431168508948283

Smith, D., 2020. Assessment of visible and short wavelength radiometric calibration using vicarious calibration methods. Preparation and Operations of the Mission Performance Centre (MPC) for the Copernicus Sentinel-3 Mission.

Swinnen, E., Verbeiren, S., Henry, P., Deronde, B., Henry, P., 2014. Assessment of the impact of the orbital drift of SPOT-VGT1 by comparison with SPOT-VGT2 data. Int. J. Remote Sens. 35, 49–52. https://doi.org/10.1080/01431161.2014.883100

Toté, C., Swinnen, E., Sterckx, S., Adriaensen, S., Benhadj, I., Iordache, M.-D., Bertels, L., Kirches, G., Stelzer, K., Dierckx, W., Van den Heuvel, L., Clarijs, D., Niro, F., 2018. Evaluation of PROBA-V Collection 1: Refined Radiometry, Geometry, and Cloud Screening. Remote Sens. 10, 1375. https://doi.org/10.3390/rs10091375

Toté, C., Swinnen, E., Sterckx, S., Clarijs, D., Quang, C., Maes, R., 2017. Evaluation of the SPOT/VEGETATION Collection 3 reprocessed dataset: Surface reflectances and NDVI. Remote Sens. Environ. 201, 219–233. https://doi.org/10.1016/j.rse.2017.09.010

Toté, C., Swinnen, E., Wolters, E., Smets, B., 2019. Validation Report - Normalized Difference Vegetation Index (ENDVI10) - LSA-420. EUMETSAT LSA SAF, https://landsaf.ipma.pt/en/products/vegetation/endviv2/.

Walthall, C.L., Norman, J.M., Welles, J.M., Campbell, G., Blad, B.L., 1985. Simple equation to approximate the bidirectional reflectance from vegetative canopies and bare soil surfaces. Appl. Opt. 24, 383–387. https://doi.org/10.1364/AO.24.000383

Wolters, E., Swinnen, E., Toté, C., Sterckx, S., 2016. SPOT-VGT Collection 3 Products User Manual. VITO, https://docs.terrascope.be/#/DataProducts/SPOT-VGT/ProductsOverview.

Wolters, E., Toté, C., Sterckx, S., Adriaensen, S., Henocq, C., Bruniquel, J., Scifoni, S., Dransfeld, S., 2021. Icor atmospheric correction on sentinel-3/OLCI over land: Intercomparison with aeronet, radcalnet, and syn level-2. Remote Sens. 13, 1–26. https://doi.org/10.3390/rs13040654

Wolters, E., Toté, C., Swinnen, E., 2020. Algorithm Theoretical Basis Document for Normalized Difference Vegetation Index (ENDVI10) version 2. EUMETSAT LSA SAF, https://landsaf.ipma.pt/en/products/vegetation/endviv2/.

Wu, X., Naegeli, K., Wunderle, S., 2020. Geometric accuracy assessment of coarse-resolution satellite datasets: a study based on AVHRR GAC data at the sub-pixel level. Earth Syst. Sci. Data 12, 539–553. https://doi.org/10.5194/essd-12-539-2020

ANNEX A: COMPARISON OF RELATIVE SPECTRAL RESPONSE FUNCTIONS

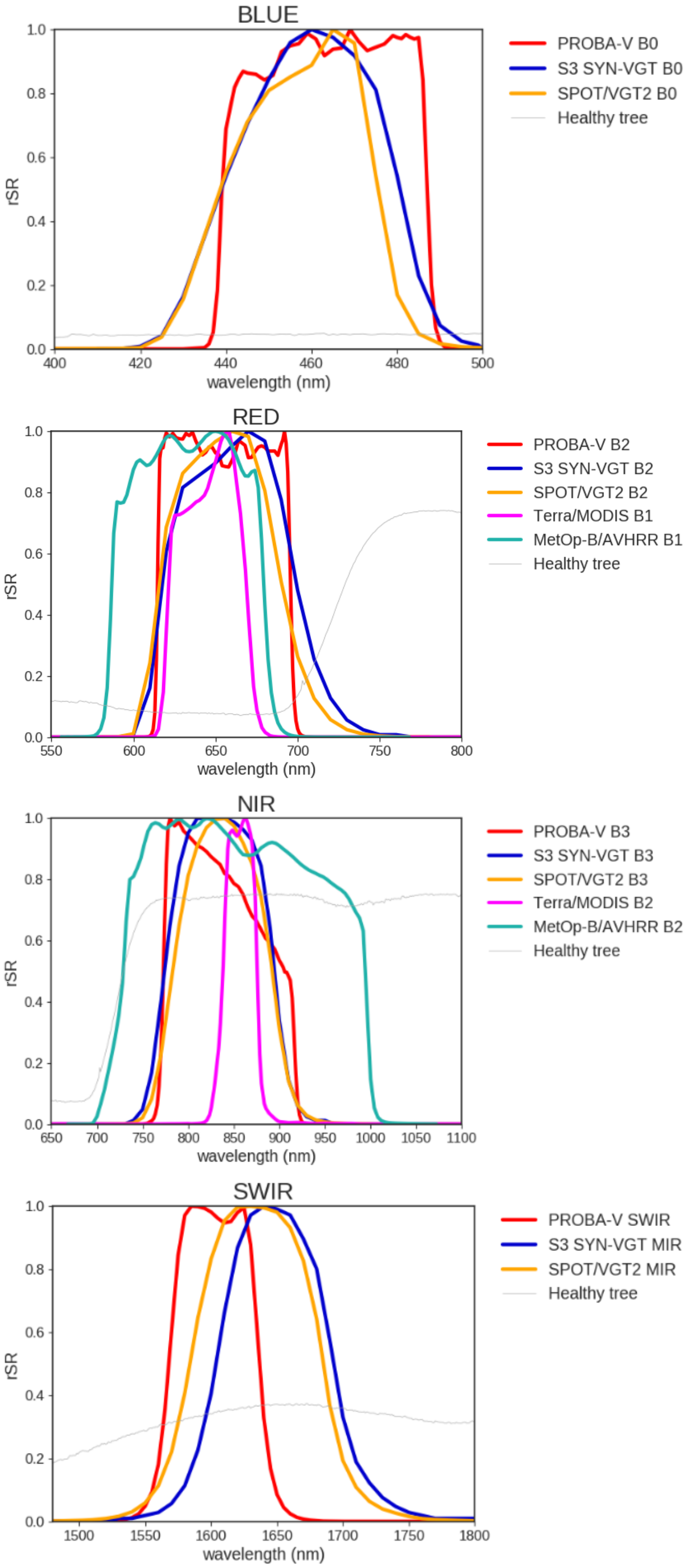


Figure 31: Comparison of relative spectral response (rSR) for Blue, Red, NIR and SWIR bands of datasets used in this report

ANNEX B: COMPARISON OF EQUATOR LOCAL OVERPASS TIMES

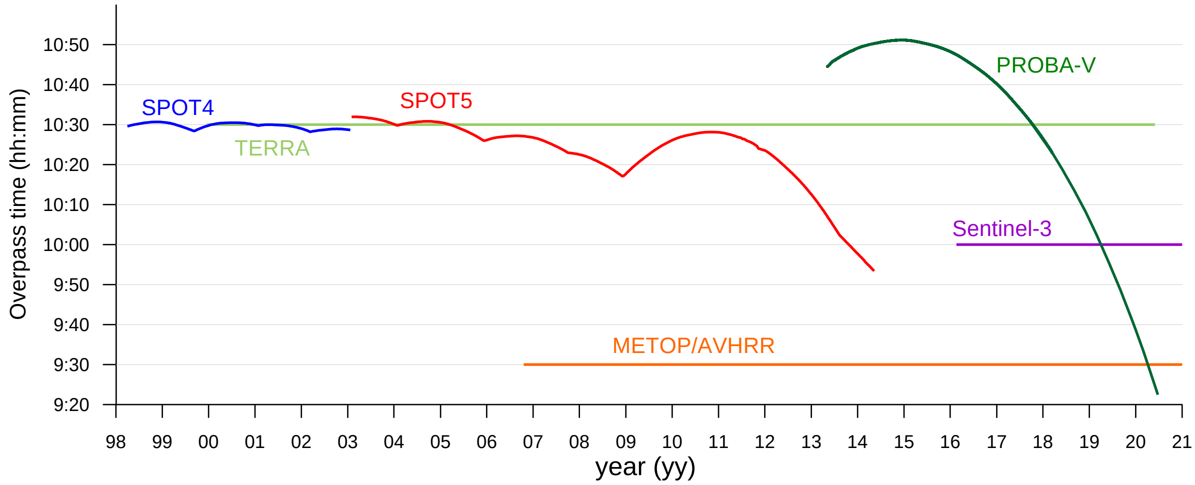


Figure 32 Equator local overpass time of Proba-V, SPOT4, SPOT5, Terra, METOP-A and Sentinel-3

ANNEX C: SPATIAL DISTRIBUTION OF STATUS MAP LABEL OCCURRENCES

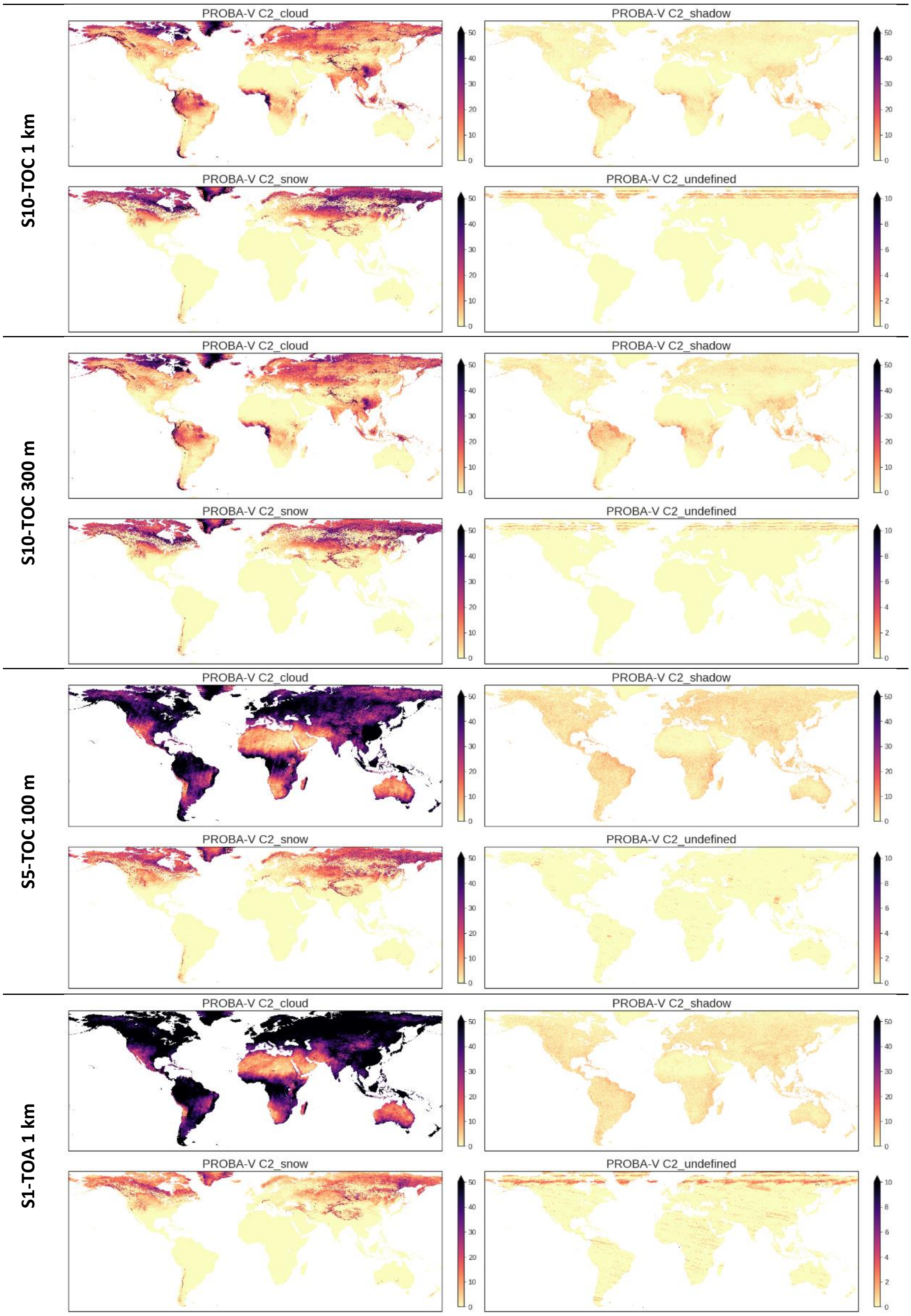


Figure 33: Spatial distribution of the occurrences (%) of cloud, shadow, snow/ice and undefined flags in PV-C2 for different products (October/2013 – June/2020)

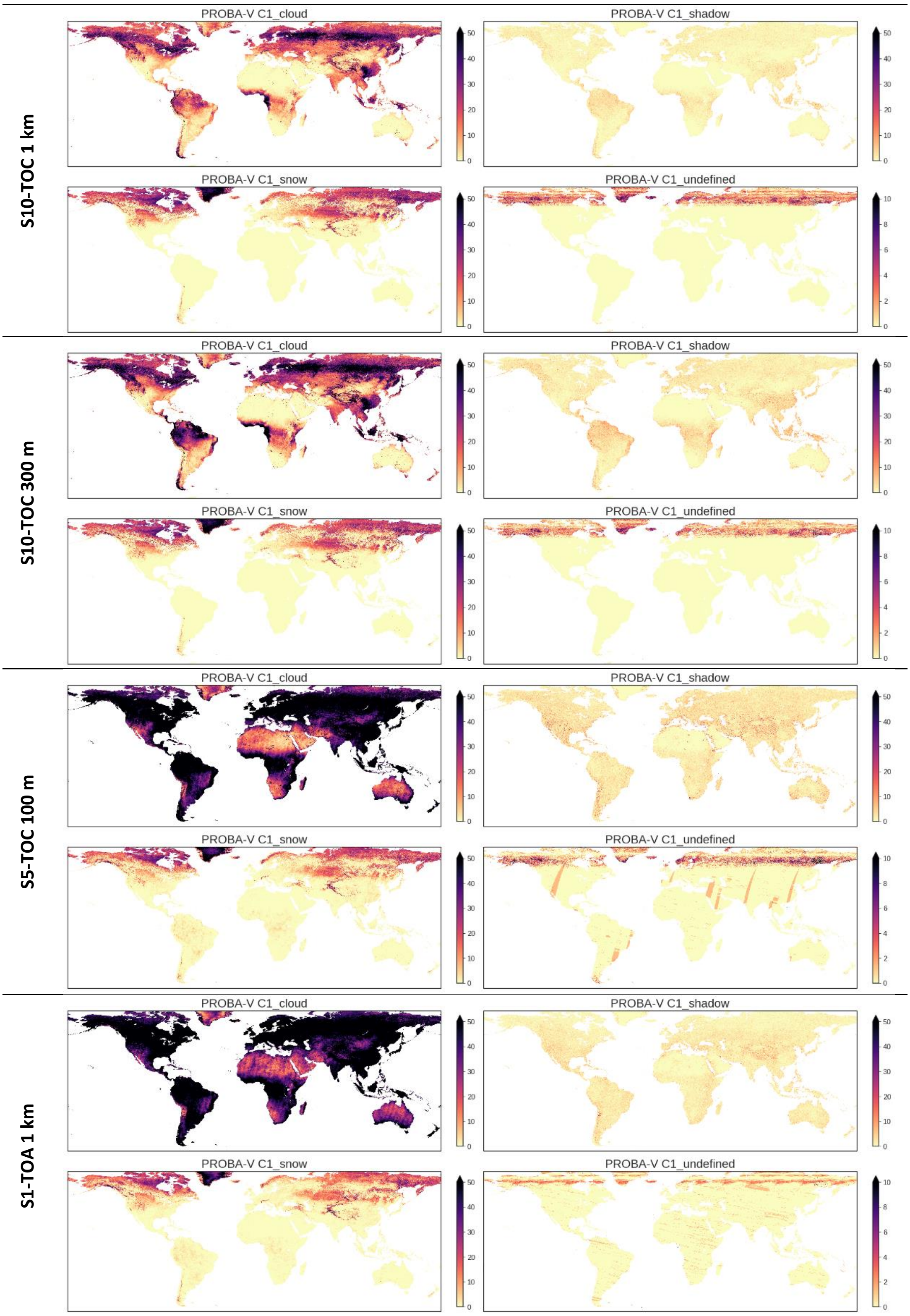


Figure 34: Spatial distribution of the occurrences (%) of cloud, shadow, snow/ice and undefined flags in PV-C1 for different products (October/2013 – June/2020)

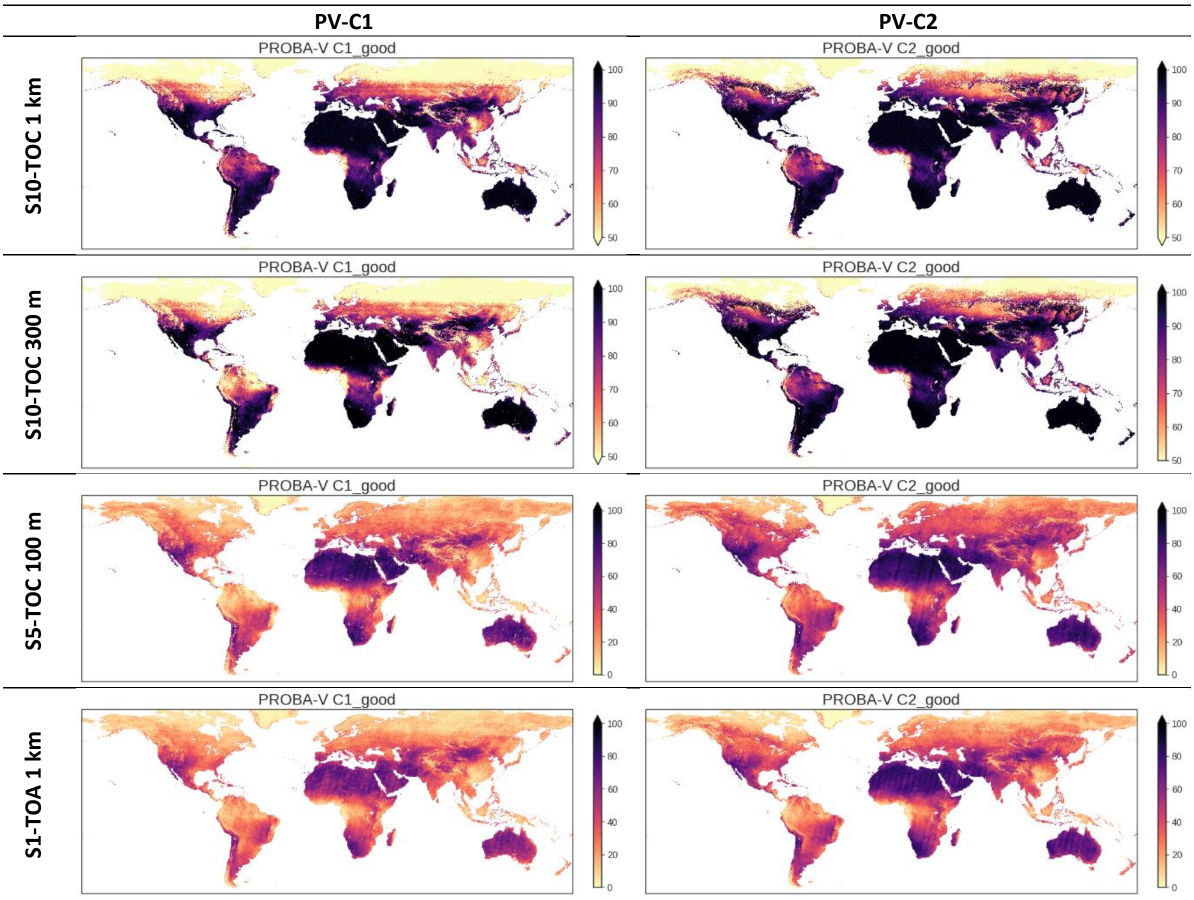
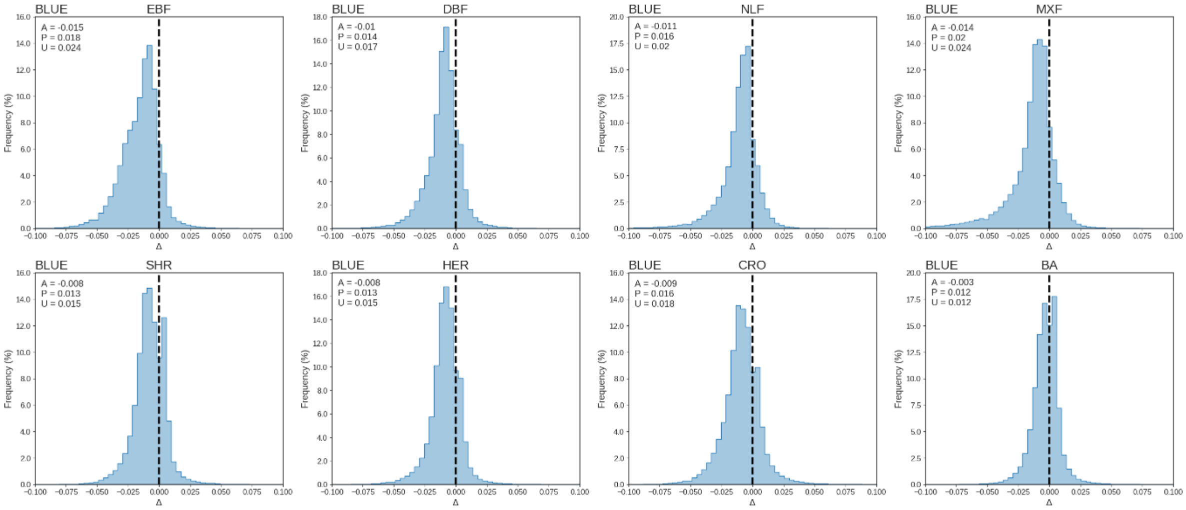


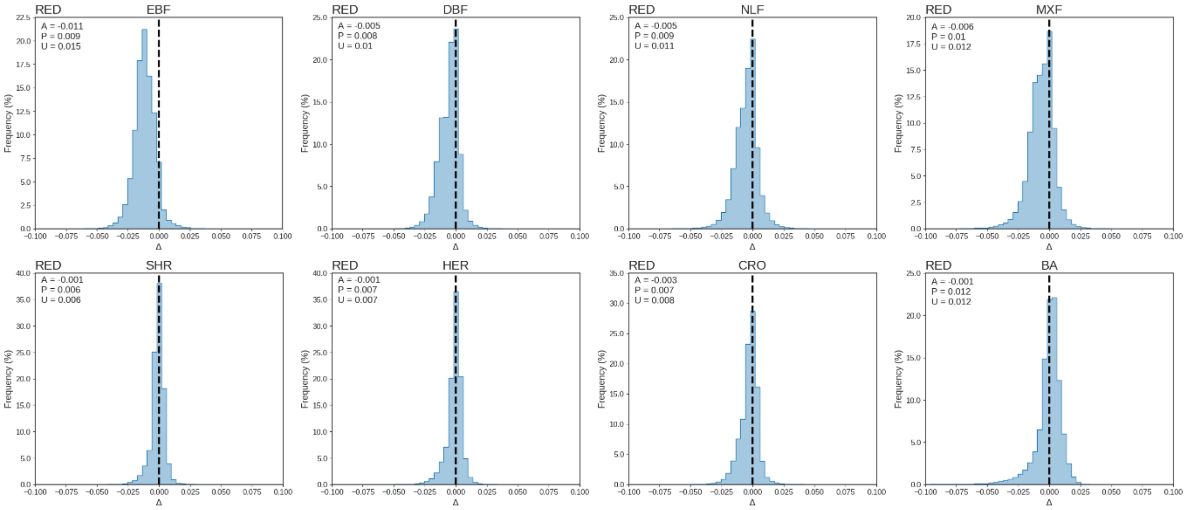
Figure 35: Spatial distribution of the occurrences (%) of good pixel status (clear and good radiometric quality) in PV-C1 (left) and PV-C2 (right) for different products (October/2013 – June/2020)

ANNEX D: BIAS HISTOGRAMS OF TOC REFLECTANCES PER BIOME

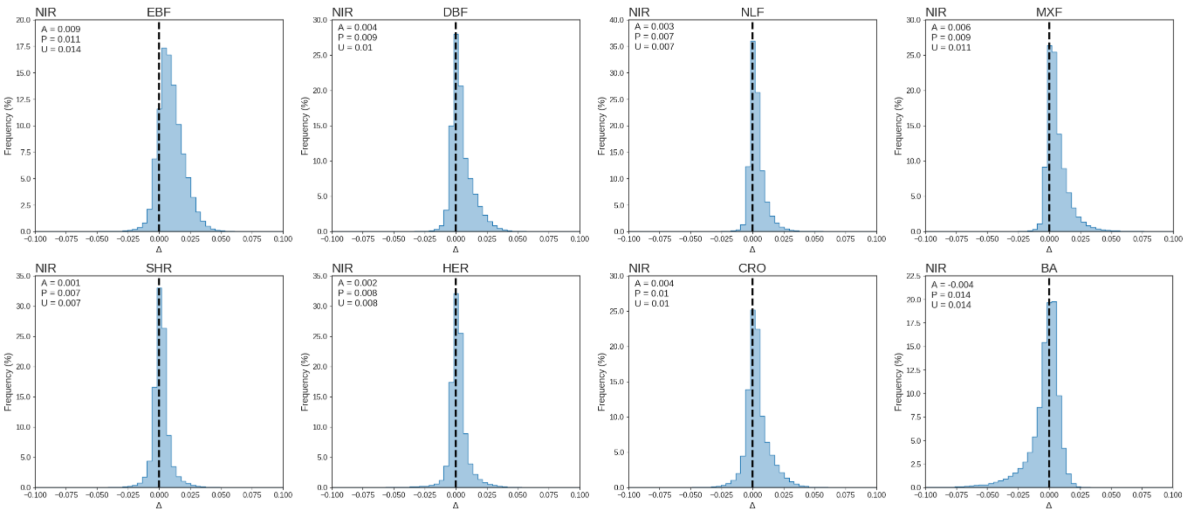
BLUE



RED



NIR



SWIR

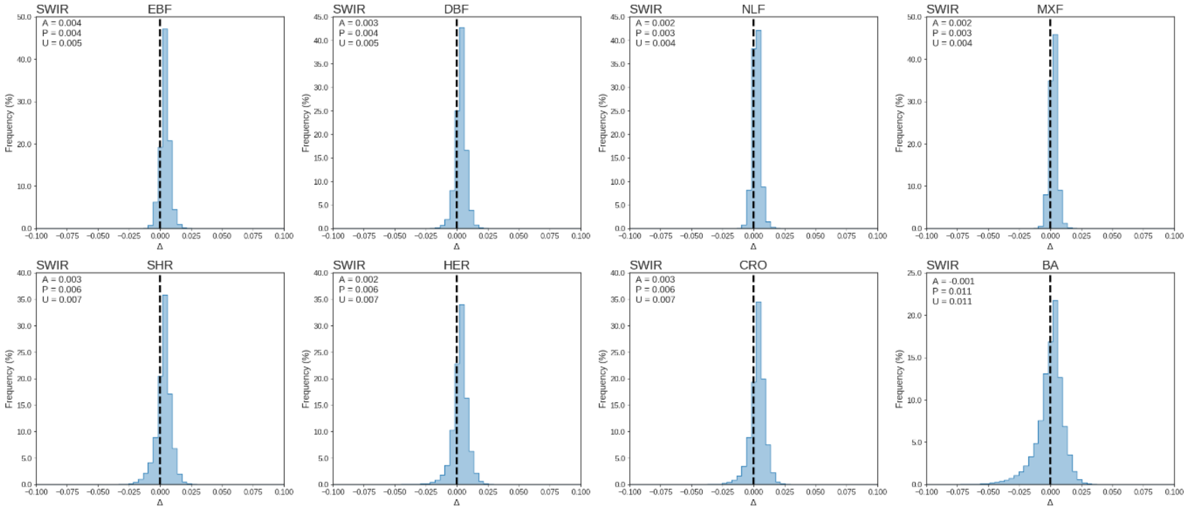


Figure 36: Bias histograms between PV-C1 and PV-C2 S10-TOC reflectances at 1 km resolution, stratified per biome (October/2013 – June/2020)

ANNEX E: TEMPORAL EVOLUTION OF APU METRICS

Temporal evolution of APU metrics of the pairwise intercomparison between Proba-V C1 and Proba-V C2 surface reflectances overall (Figure 37) and per camera (Figure 38).

S10-TOC 1 km

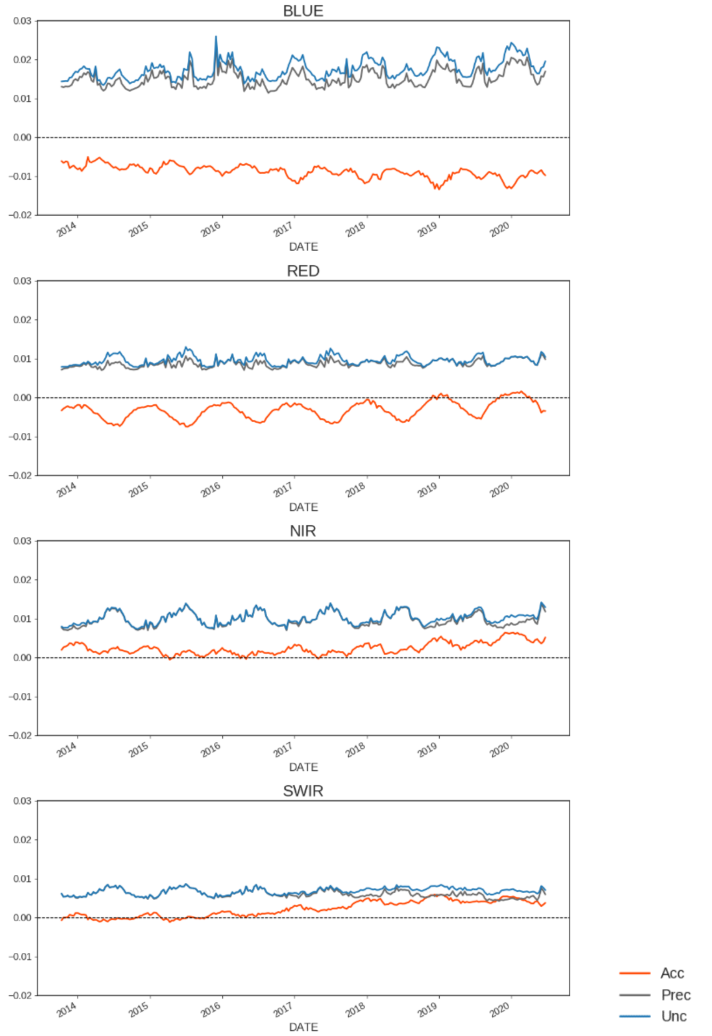


Figure 37: Temporal evolution of APU metrics between C2 and C1 S1-TOA reflectance (left) and S10-TOC reflectance (right) at 1 km resolution calculated over October/2013 – June/2020 per band for all cameras, using all clear pixels that have the same observation day.

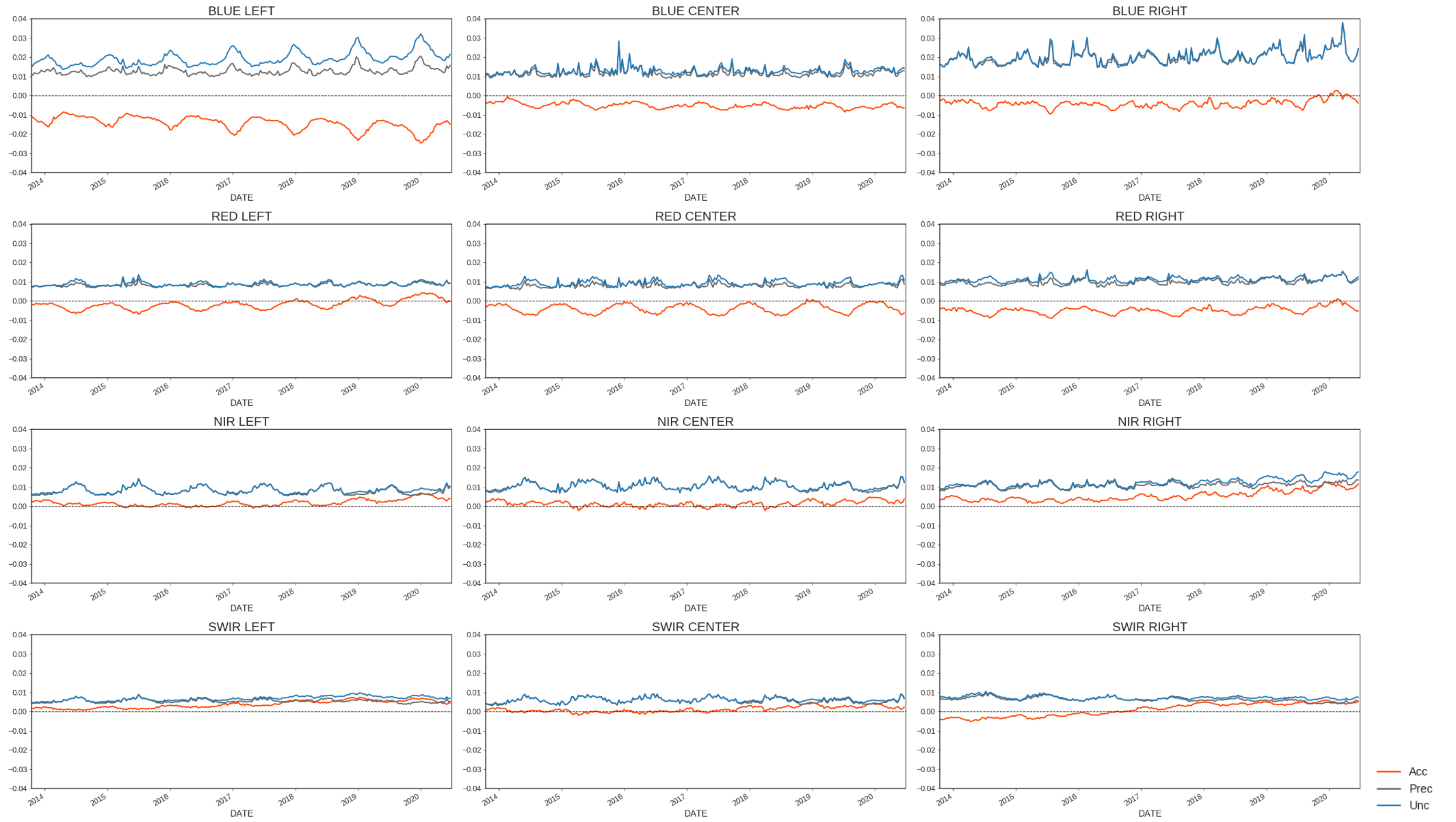


Figure 38: Temporal evolution of APU metrics between C2 and C1 S10-TOC reflectance at 1 km resolution calculated over October/2013 – June/2020 per band and per camera, using all clear pixels that have the same observation day.

ANNEX F: DESCRIPTION OF THE DIGITAL ANNEX

This validation report is accompanied by a digital annex, containing the temporal profiles of NDVI, comparing Proba-V C2, Proba-V C1, SPOT/VGT-C3, S3A&B SYN-VGT V10, ENDVI10 v2 and MOD13A3 C6.1 NDVI.

A number of profiles were selected and visualized in this report (see §7.5).

The annex consists of 680 PNG files, with following naming convention:

NDVI_profile_LANDVAL_site_<siteID>_<biome>_2009-2022.png

The <biome> refers to the biome, derived from the aggregated Copernicus Global Land Service (CGLS) – Land Cover at 100 m (LC100), epoch 2015 (Buchhorn et al., 2019), as described in §3.3.3. The number of sites per biome is shown in Table 13.

Table 13: Number of LANDVAL sites per biome

Biome	Number of LANDVAL sites
EBF	70
DBF	59
NLF	77
MXF	5
SHR	54
HER	155
CRO	130
BA	101
OTH	69