remote sensing

MDPI

Review

European Space Agency (ESA) Calibration/Validation Strategy for Optical Land-Imaging Satellites and Pathway towards Interoperability

Fabrizio Niro [1,*], Philippe Goryl [2], Steffen Dransfeld [2], Valentina Boccia [2], Ferran Gascon [2], Jennifer Adams [3], Britta Themann [2], Silvia Scifoni [1] and Georgia Doxani [1]

[1] Serco SpA for European Space Agency (ESA), European Space Research Institute (ESRIN), 00044 Frascati, Italy; Silvia.Scifoni@serco.com (S.S.); Georgia.Doxani@esa.int (G.D.)
[2] European Space Agency (ESA), European Space Research Institute (ESRIN), 00044 Frascati, Italy; Philippe.Goryl@esa.int (P.G.); Steffen.Dransfeld@esa.int (S.D.); valentina.boccia@esa.int (V.B.); ferran.gascon@esa.int (F.G.); Britta.Themann@esa.int (B.T.)
[3] RHEA System SpA for European Space Agency (ESA), European Space Research Institute (ESRIN), 00044 Frascati, Italy; Jennifer.Adams@esa.int
* Correspondence: fabrizio.niro@esa.int

Abstract: Land remote sensing capabilities in the optical domain have dramatically increased in the past decade, owing to the unprecedented growth of space-borne systems providing a wealth of measurements at enhanced spatial, temporal and spectral resolutions. Yet, critical questions remain as how to unlock the potential of such massive amounts of data, which are complementary in principle but inherently diverse in terms of products specifications, algorithm definition and validation approaches. Likewise, there is a recent increase in spatiotemporal coverage of in situ reference data, although inconsistencies in the used measurement practices and in the associated quality information still hinder their integrated use for satellite products validation. In order to address the above-mentioned challenges, the European Space Agency (ESA), in collaboration with other Space Agencies and international partners, is elaborating a strategy for establishing guidelines and common protocols for the calibration and validation (Cal/Val) of optical land imaging sensors. Within this paper, this strategy will be illustrated and put into the context of current validation systems for land remote sensing. A reinforced focus on metrology is the basic principle underlying such a strategy, since metrology provides the terminology, the framework and the best practices, allowing to tie measurements acquired from a variety of sensors to internationally agreed upon standards. From this general concept, a set of requirements are derived on how the measurements should be acquired, analysed and quality reported to users using unified procedures. This includes the need for traceability, a fully characterised uncertainty budget and adherence to community-agreed measurement protocols. These requirements have led to the development of the Fiducial Reference Measurements (FRM) concept, which is promoted by the ESA as the recommended standard within the satellite validation community. The overarching goal is to enhance user confidence in satellite-based data and characterise inter-sensor inconsistencies, starting from at-sensor radiances and paving the way to achieving the interoperability of current and future land-imaging systems.

Keywords: calibration; validation; optical; land; interoperability; traceability; metrology; Fiducial Reference Measurements (FRM)

check for updates

Citation: Niro, F.; Goryl, P.; Dransfeld, S.; Boccia, V.; Gascon, F.; Adams, J.; Themann, B.; Scifoni, S.; Doxani, G. European Space Agency (ESA) Calibration/Validation Strategy for Optical Land-Imaging Satellites and Pathway towards Interoperability. *Remote Sens.* **2021**, *13*, 3003. https://doi.org/10.3390/rs13153003

Academic Editors: Fernando Camacho and Jadu Dash

Received: 25 June 2021
Accepted: 26 July 2021
Published: 30 July 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Recent years have seen an unprecedented growth of space-borne Earth Observation (EO) optical land imaging systems at ever increasing spatial, temporal and spectral resolutions. The advent of the New Space era, with the launch of a multitude of small satellite constellations, further contributes to expanding our ability of monitoring Earth's system

processes at enhanced spatiotemporal scales. Similarly, a fast-growing number of automated cost-effective sensors provides a unique opportunity for gathering valuable in situ data for the validation of satellite-derived bio-geophysical products. All in all, our current capability in generating EO data, both from space and from ground-based or drone-based platforms, far outpace our ability to synergistically exploit their measurements.

To get the most out of such a deluge of EO data in the years to come, we are faced with two major challenges: diversity and quality. Diversity may result, if properly handled, in enhanced observation capabilities by combining complementary observations acquired by a variety of heterogeneous sensors. Quality translates in this context to the need of ensuring that the relevant EO data is provided to the users with exhaustive and fully traceable information about its quality, allowing them to readily evaluate the fitness for purpose for their applications [1]. As long as EO data quality is assessed with metrological rigour and using community-endorsed best practices, the measurements acquired by heterogeneous sensors can be reliably combined, meaning that the corresponding observing systems are interoperable.

Interoperability can be tackled at different levels of complexity, ranging from format definition, products content and metadata, data processing algorithms and quality assurance approaches. Yet, in general terms, two systems can be considered interoperable when their derived EO data come with all the necessary quality information, including a detailed uncertainty estimate, allowing to fully characterise, and eventually correct for, the systematic errors inherent in each system. The interoperability concept is tackled within this paper from a Space Agency's perspective, meaning in terms of end-to-end system performances. The overall quality assurance process is therefore reviewed and discussed, starting in the early phase of the mission, with the pre-launch characterisation activities, and continuing while in-orbit, with the calibration and validation activities (Cal/Val).

Cal/Val and metrology are the two essential elements in order to assess inconsistencies between sensors and consequently address the potential for interoperability. Cal/Val activities provide the reference measurements and the relevant methods for quantitatively estimating the quality (uncertainty) of satellite EO products. The metrology component ensures the generic framework and the internationally agreed standards, ideally traceable to SI (International System of units), to tie the EO data acquired by a variety of sensors to a common reference. The advances and standardisation of Cal/Val methodologies and the implementation of metrological practices are therefore crucial for enabling an informative exploitation of the various EO remote sensing systems, eventually reaching the objective of a global System-of-Systems [2].

This System-of-Systems vision equally applies to in situ reference data, which are used to validate the satellite-based products. This ideal condition is far from being attained for the currently available land validation systems. As a matter of fact, the existing in situ validation data from field campaigns and networks still lack consistency and proper traceability to metrological standards, which limits their integrated use for validation purposes. The harmonisation of best practices across existing ground-based networks would significantly enhance our ability to validate satellite-based EO data in terms of spatiotemporal and geographical coverage.

This document aims to provide a review of the state-of-the-art in the domain of satellites land products validation, with a focus on current and future European Space Agency (ESA) optical imaging sensors and to identify the current challenges, the needs for improvements and the data gaps. Furthermore, the on-going efforts for tackling such gaps are reported and discussed. The status and needs were gathered in a series of international meetings with the science community, in particular in a recent land validation workshop [3]. The overarching objective of this document is to set the stage for building a long-term strategy for ESA optical satellite land products validation, which will help to define and prioritise the upcoming Cal/Val activities in coordination with the other Space Agencies and in close liaison with the Committee on Earth Observation Satellite Land Product Validation (CEOS-LPV) sub-group.

The document starts in Section 2 by introducing the background and the international context in which the ESA operates to coordinate and steer the effort in land Cal/Val activities. Section 3 provides the underlying concepts and terminology, which will be referred to throughout the paper. Section 4 illustrates the current status and challenges in Cal/Val, while Section 5 presents the readiness level in terms of infrastructure and methods and the overall strategy to fill the identified gaps. Concluding remarks are presented in Section 6.

2. Background and International Context

2.1. ESA/Copernicus Current and Future Land-Focused Optical Sensors

ESA, in the frame of the European Copernicus programme [4], has developed two satellite families in the optical domain, namely Sentinel-2 (S2) [5] and Sentinel-3 (S3) [6], to derive the bio-geophysical land variables needed by the land downstream services. Additionally, Copernicus contributing missions, in particular, Proba-V [7], operationally provided real time data to the Copernicus Global Land Service (CGLS) [8] and Copernicus Climate Change Service (C3S) [9] up to June 2020, bridging the gap until the full operations of the Sentinel-3 twin satellites constellation. The payload carried by these missions allows complementary observations of the Solar-reflected and Earth-emitted spectral radiation at different spatiotemporal and spectral resolutions. More specifically, these missions have the following main characteristics:

- Sentinel-2 is a two-satellite constellation operational since March 2017 and ensuring a 5-day repeat cycle at the Equator in a sun-synchronous orbit at 786 km altitude with a local time at the descending node of 10:30 AM [5]. The Multi-Spectral Imager (MSI) on board S2 is a push-broom instrument and has a wide swath of 290 km with 13 spectral bands, covering the Visible (VIS), Near-Infrared (NIR) and Short-Wave Infrared (SWIR) domains at different spatial resolutions at the ground, ranging from 10 m to 60 m. The four bands at 10 m resolution in the VIS and NIR range ensure compatibility with SPOT 4 and 5 and meet the user requirements for land cover classification. The 20 m bands include the spectrally narrow red-edge bands for vegetation status monitoring, and the two SWIR bands for improving snow/ice/cloud discrimination. The bands at 60 m resolution are mainly dedicated to atmospheric correction and cloud screening.
- Sentinel-3 is a twin-satellite constellation operational since April 2018 in a sun-synchronous, near-polar orbit at 814 km with a local time at the descending node of 10:00 AM [6]. The Ocean and Land Colour Instrument (OLCI) on board S3 is a spectrometer imaging in push-broom mode and consisting of five cameras in a fan-shaped arrangement. OLCI provides a native 300 m resolution across a swath of 1270 km with 21 spectral bands in the 400–1020 nm spectral range. S3 also carries the Sea and Land Surface Temperature Radiometer (SLSTR) on board, a dual view conical scanning imaging radiometer designed to retrieve global coverage sea and land surface temperature, active fire monitoring, ice surface temperature, clouds and atmospheric aerosols in support of various Copernicus services. SLSTR provides six solar VIS/NIR/SWIR reflectance bands at 500 m resolution and three thermal infrared bands and two dedicated fire bands at 1 km resolution.
- The Proba-V satellite was launched in March 2013 in a quasi-sun-synchronous orbit at 820 km, with a mean local time in the range of 10:45 AM +/−15 min [7]. The VEGETATION instrument on board Proba-V consists of three overlapping cameras covering a 2250 km swath for near-daily global coverage of landmasses and coastal zones. The Proba-V cameras acquire data into four spectral bands: Blue, Red, NIR and SWIR. The Proba-V operational phase was terminated in June 2020 to limit the impact of orbital drift on mission's archive multi-temporal consistency [10]. By 1 July 2020, the Proba-V mission entered an Experimental Phase scenario with reduced acquisition planning but with the perspective to expand its observation capabilities thanks to the launch of a series of Cubesat Companion satellites [11].

The ensemble of the ESA's operational land products derived from those missions and currently provided to the user community is summarised in Table 1, with a focus on those bio-geophysical variables retrieved from the observations in the solar spectrum domain, which is the main subject of this paper. Future expansion of this portfolio includes the S2 harmonised and fused products in combination with Landsat-8, which are currently in the prototype phase [12]. Supplementary land products derived from thermal observations of the S3 SLSTR instrument, include, notably, the Land Surface Temperature (LST) and Fire Radiative Power (FRP), which are key variables to study evapotranspiration processes and vegetation fire disturbances.

Future land-focused missions are currently in preparation; they will be launched in the coming years to complement and expand ESA/Copernicus observation capabilities. We recall here, in particular, the following missions:

- The Fluorescence Explorer (FLEX) mission [13,14], which aims at providing global maps of vegetation fluorescence for studying photosynthetic activity and plant health and stress. The FLEX satellite will orbit in tandem with one of the Sentinel-3 satellites, taking advantage of its optical and thermal sensors to provide complementary measurements, in particular for atmospheric correction and cloud screening. The FLEX satellite will carry on board the FLuORescence Imaging Spectrometer (FLORIS), a high-resolution spectrometer acquiring data in the 500–780 nm spectral range, with a sampling of 0.1 nm in the oxygen bands (759–769 nm and 686–697 nm) and 0.5–2.0 nm in the red edge, chlorophyll absorption and PRI (Photochemical Reflectance Index) bands.
- The Copernicus Hyperspectral Imaging Mission (CHIME) [15] will carry a unique visible to shortwave infrared spectrometer to provide routine hyperspectral observations to support new and enhanced services for sustainable agricultural and biodiversity management, as well as soil property characterisation. The mission will complement Sentinel-2 for applications such as land cover mapping.
- The Copernicus Land Surface Temperature Monitoring (LSTM) [16] mission will carry a high-spatiotemporal-resolution thermal infrared sensor to provide observations of the land surface temperature. Land surface temperature measurements and the derived evapotranspiration are key variables to understand and respond to climate variability, manage water resources for agricultural production, predict droughts and to address land degradation, natural hazards, coastal and inland water management as well as urban heat island issues.

Table 1. List of current ESA Optical Land imaging satellites with their operational land products.

Mission	Sensor	Level	Description	Reference
Sentinel-2	MSI	L1-C	Top-Of-Atmosphere (TOA) reflectances radiometric and geometrically corrected	[17]
		L2-A	Bottom-Of-Atmosphere (BOA) reflectances[1] corrected for atmospheric effects	[17,18]
Sentinel-3	OLCI	L1-B	TOA radiances in the instrument grid radiometrically corrected	[19]
	SLSTR	L1-B	TOA radiances (VIS, NIR, SWIR channels) and TOA brightness temperatures (thermal infrared channels)	[20]
	OLCI	L2 Land	OLCI Global Vegetation Index (OGVI renamed in the future as GI-FAPAR, Green Instantaneous-fAPAR) providing the Fraction of Absorbed Photosynthetically Active Radiation (fAPAR)	[21,22]
	OLCI	L2 Land	OLCI Terrestrial Chlorophyll Index (OTCI) providing indication of the content of Chlorophyll in the canopy	[23,24]

Table 1. *Cont.*

Mission	Sensor	Level	Description	Reference
	OLCI + SLSTR	L2	S3 Synergy Surface Directional Reflectances[1] product (SY_2_SYN), atmospherically corrected surface reflectances over land provided in SLSTR solar reflective bands (except S4 dedicated to cloud detection) and OLCI channels (except for oxygen absorption bands Oa13, Oa14, Oa15 and water vapour bands Oa19 and Oa20)	[25]
Proba-V	Vegetation	L1-C	TOA reflectances in instrument grid unprojected, but radiometrically and geometrically corrected.	[26]
		L3	Top-Of-Canopy (TOC) atmospherically corrected reflectances[1], provided as multi days composite products (S1, S5, S10) at different spatial resolution (1 km, 333 m and 100 m).	[26]

[1] The definition of "surface reflectances" slightly differs among these ESA operational products. For S2, the L2A [18] is an approximate of the HDRF (Hemispherical Directional Reflectance Factor) [27]. The same applies to Proba-V TOC products (provided as multi-days synthesis), while for S3, the Surface Directional Reflectance (SDR) corresponds to the BRF (Bi-directional Reflectance Factor) [25,27]. In the remainder of the paper, the term surface reflectance (SR) is used for simplicity to discuss the generic validation needs for present and future ESA optical land imaging sensors.

2.2. International Context

2.2.1. Inter-Agencies Working Groups

The need for globally distributed reference data and for community-agreed Cal/Val best practices calls for strong international collaborations. The objective of these collaborations is to join forces and exploit synergies among the various Space Agencies and scientific institutions. Various inter-agency bodies are actively working in this respect. Among the various groups, we recall here those most relevant for satellite land products validation and for addressing the need of inter-sensors interoperability in the optical domain:

- The Global Space-based Inter-Calibration System (GSICS) [28] is an international collaborative effort, initiated and promoted by the World Meteorological Organization (WMO), aiming at monitoring, improving and harmonising the quality of observations from operational weather and environmental satellites. The GSICS delivers calibration corrections needed for accurately integrating data from multiple observing systems into products, applications and services. ESA is an active contributor to the GSICS working groups' objectives in the advances and consolidation of best practices for cross-sensor intercalibration.
- The CEOS Infrared and Visible Optical Sensors Subgroup (CEOS-IVOS) [29] is a sub-level working group of the CEOS-WGCV (Working Group on Calibration and Validation) aiming at fostering international collaborations in the frame of Cal/Val for optical sensors and proposing and agreeing on methodologies and standards, identifying test sites and sharing Cal/Val data among the various agencies. The ESA has been an active member of the CEOS-IVOS group for many years and actively supported the definition and implementation of advanced practices and networks for optical sensors Cal/Val.
- The CEOS Land Surface Imaging Virtual Constellation (CEOS-LSI-VC) [30] is a sub-group of the CEOS Virtual Constellations (CEOS-VC), which is responsible for promoting the efficient and comprehensive collection, distribution and application of space-acquired image data of the global land surface. The CEOS-LSI-VC leads the CEOS Analysis Ready Data for Land (CARD4L) initiative [31], including the definition of the individual product family specifications, as well as the assessment process and outreach. ESA is active member of the LSI-VC working group and is currently contributing to the CARD4L initiative.
- CEOS Land Product Validation (CEOS-LPV) [32] is a sub-group of the CEOS-WGCV, which is responsible for coordinating the quantitative validation of satellite-derived land products. The main objective of CEOS-LPV is to establish standardised guidelines and best practice protocols for the validation of a wide range of terrestrial Essential

Climate Variables (ECVs). ESA is actively supporting CEOS-LPV activities in the various focus areas and is promoting the series of Land Products Validation Workshops [3,33], gathering the CEOS-LPV teams for discussing on common issues and collecting recommendations for improvements in Cal/Val best practices for land.

2.2.2. Sentinels Validation Teams

In addition to the inter-agency working groups, ESA is contributing to dedicated Cal/Val activities in the frame of the Sentinels Validation Teams, whose objectives are to gather validation experts to contribute to the quality assessment of the Sentinels mission core products. The following two groups are particularly relevant for the scope of this paper:

- The Sentinel-2 Validation Team (S2VT) [34] is responsible for coordinating and overseeing the quantitative validation of satellite data products derived from the Sentinel-2 mission. S2VT brings together, on an annual basis, leading Cal/Val experts addressing all S2 validation requirements, encompassing a large range of domains from radiometric and geometric calibration to land and water products validation. The recommendations gathered during a series of annual meetings are used to further reshape and improve Cal/Val methodologies, as well as easing the exchange of Cal/Val data among the various scientific teams.
- The Sentinel-3 Validation Team (S3VT) [35] is jointly convened by the ESA and the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) to promote and coordinate the involvement of the international community in the domain of Cal/Val activities for the Sentinel-3 mission. The S3VT is divided in different sub-groups focusing on Land, Ocean Colour, Sea Surface Temperature, Altimetry and Atmosphere. The S3VT meets on a regular basis to review the status of the Cal/Val activities and the quality of the S3 core products, as well as to support the exchanges of ideas and collaboration among the various members and to discuss potential product evolutions.

As part of the Sentinels mission and in order to ensure the highest quality of products, the Mission Performance Centers (MPC, Center to be replaced by Cluster for the next generation MPC from 2022 onwards) are established to control the quality of all generated products. Expert Support Laboratories (ESLs) are involved in calibration and validation activities and provide specific assessments of the products (e.g., analysis of trends, ad hoc analysis of anomalies, etc.). Furthermore, the Quality Working Groups (QWGs) are chaired by the ESA and co-chaired with EUMETSAT for S3 L1 to define and evolve the operational baseline algorithm and contribute to the geophysical validation of the operational products.

3. Underlying Terminology and Concepts

3.1. Calibration and Validation

Following the CEOS-recommended terminology [36], we define calibration as "the process of quantitatively defining a system's responses to known, controlled signal inputs". Validation, on the other hand, is defined as: "the process of assessing, by independent means, the quality (uncertainty) of the data products derived from those system outputs" [37].

The radiometric and geometric calibration of optical land imaging sensors is the fundamental preprocessing step for ensuring that the considered mission provides reliable, accurate and consistent observations over space and time to meet the stringent requirements for scientific applications. The outcomes of the calibration process are, in the optical solar spectrum domain, the radiance or reflectance data at the Top-Of-Atmosphere (TOA) level, which are generally referred to as Level 1 (L1) products.

The validation consists in quantifying the uncertainties of the satellite dataset by analytical comparison with a space-and-time-collocated independent measurement of the same quantity, which is assumed as the reference value (*ground truth*). Validation is a

generic term that applies to any satellite-based product, both Level 1, e.g., validation of at-sensor radiance, and Level 2 (L2) surface reflectance and bio-geophysical products.

The ultimate goal of the validation is to ensure that the users' requirements, initially set as part of the Mission Requirement Document (MRD), in terms of desired products' accuracy, are effectively met in orbit. Therefore, a comprehensive and accurate set of validation reference data is fundamental to verify that the mission successfully fulfilled its scientific objectives and to support corrective actions, e.g., upgrades of the retrieval algorithms and/or calibration methods in the case of shortcomings.

The geophysical validation of satellite-derived land products is an extremely challenging task, owing to the intrinsic heterogeneity and anisotropy of terrestrial surfaces and the complex dependency on observation geometries, illumination conditions, canopy structure and layering. Within this chapter, we review the main concepts and underlying physical challenges associated to the Cal/Val of optical land imaging sensors.

3.2. Traceability and Uncertainty

In the ideal scenario, the space-borne data and the reference validation measurements are both linked to a metrological standard, ideally up to SI units, through an unbroken chain of calibrations or comparisons, each contributing to the stated measurement uncertainty. This is the so-called "traceability" of the measurements, as defined in the International Vocabulary of Metrology (VIM) [38]. The traceability requirement equally applies to all satellite-based EO data and should be verified along the full processing chain from L1 to L2 bio-geophysical products. Examples of visualised traceability chains for various land and atmosphere ECVs [39] were generated as part of the Quality Assurance for Essential Climate Variables (QA4ECV) project [40].

In practice, both the satellite and the reference data are rarely fully traceable, since rigorous calibrations performed in well-controlled laboratory conditions cannot be replicated in the field or in space. In addition, there is a lack of a consistent and metrologically-sound vocabulary within the EO community to describe the data and their quality [41]. The most obvious example are the terms "error" and "uncertainty", which are often misused or used interchangeably. Within this paper, we follow VIM recommended definition, namely [38]:

- Measurement error is the measured quantity value minus a reference quantity value. The measurement error can contain both a random and a systematic component. While the random component averages out over multiple measurements, the systematic component does not. On the other hand, the systematic error, also called bias, if significant in size compared to required measurement accuracy and should be carefully estimated and ideally corrected through the use of an appropriate correction factor.
- Measurement uncertainty is a non-negative parameter describing the dispersion of the quantity values attributed to a measurand, where the measurand is the quantity being measured. The methods for evaluating measurement uncertainty are classified in two main categories: Type A and Type B. Type A uncertainty evaluation is based on the statistical analysis of a series of independent observations of the same quantity. Type B includes all other methods, which do not rely on the analysis of a series of observations but are based on a prior and reliable knowledge about the measurement, e.g., previous data, knowledge about the instrument's properties and behaviour and calibration certificates.

Both error and uncertainty are intrinsic properties of any measurement, which is, by definition, an imperfect estimation of the true value of the measurand. Any satellite-based EO data should therefore be provided as a bias-free measurement, i.e., corrected for known systematic errors (e.g., radiometric or geometric), and with an associated estimate of its uncertainty. For optical land imaging sensors, this requirement results in the need of providing uncertainty per-measurand, namely per-pixel. This information should be provided already at Level-1 and propagated along the full processing chain up to the derived Level-2 bio-geophysical products, following the Guide to the expression of Uncertainty in Measurement (GUM) [42].

The rigorous application of metrological practices to EO fails at this initial stage since none of the currently provided ESA operational Level-1 products of Table 1 have an associated per-pixel uncertainty. This is due to the complexity in rigorously estimating the uncertainties at pixel level and the lack of uncertainty estimates in the used ancillary data, such as pre-flight calibration data or meteorological data. Another key challenge consists in the quantitative assessment of the pixel-to-pixel correlation. Furthermore, the provision of per-pixel uncertainty has an obvious non-negligible impact on the product's size, which cannot be easily accommodated in the frame of an operational Ground Segment infrastructure. At the moment, an attempt of providing per-pixel uncertainty at Level 1, with a qualitative assessment of pixel correlation structures in the spatial, temporal and spectral dimensions, is the S2 Radiometric Uncertainty Tool (RUT) [43,44]. The S2 RUT is currently provided as part of the Sentinels Application Platform (SNAP) toolbox [45] to be run on the user side. The integration of S2 RUT at an operational level will be investigated over the coming years. A similar exercise is currently on-going in the frame of S3 OLCI and SLSTR MPC [46].

3.3. Validation Metrics

The provision of uncertainty for both the reference and satellite data is a prerequisite in order to have a rigorous and meaningful validation. As a matter of fact, the validation problem, in its basic form, consists in verifying whether the two independent measurements of the same measurand are statistically consistent at a given significance level within their combined standard uncertainties. This translates, under the assumption of normally distributed uncertainties, to the following statistical test [47]:

$$|y_{Sat} - y_{Ref}| < k\sqrt{u_{Sat}^2 + u_{Ref}^2} \quad (1)$$

where y_{Sat} and y_{Ref} are the satellite and reference measurements, u_{Sat} and u_{Ref} are their respective uncertainties and k is the so-called coverage factor. Equation (1) provides a mean to quantitatively assess the consistency between the two independent measurements for different values of the coverage factor. Following the frequently used terminology [47], we can state that two measurements are "consistent" if they agree within k = 1, while, if they agree within k = 2, they are in "statistical agreement" to 4.5% significance level. This test can also be used to assess the quality of the uncertainty estimates. For instance, if a large number of data are compared and a fraction, much larger than 4.5%, of this data is significantly different (i.e., the consistency test is false for k = 2), then the provided uncertainties are likely to be underestimated [47]. Equation (1) was further elaborated by Loew et al. [48], introducing an additional uncertainty component, which is induced by the different spatiotemporal representativeness of the satellite and reference measurements. In this formulation, the statistical test takes the form of [48]:

$$|y_{Sat} - y_{Ref}| < k\sqrt{u_{Sat}^2 + u_{Ref}^2 + \Sigma_{mismatch}^2}, \quad (2)$$

where the term Σ, is the collocation mismatch error, which is supposed, in this expression, to be fully uncorrelated from the uncertainties of the satellite and reference measurements. Equation (2) provides a more comprehensive test, which is highly suitable to the validation of satellite-based EO products, owing to the intrinsic spatial and temporal scale mismatch between the measurements acquired from a space-borne and ground-based platform. This issue is of particular relevance for land bio-geophysical products, as further discussed in the next paragraph.

In most of the cases however, uncertainties are not provided as part of the space-borne or ground-based EO data, so the rigorous application of Equation (1) is not viable. In such cases, the general approach consists in using standard statistical metrics to quantify the pairwise agreement of the two datasets. The most commonly used methods are the correlation coefficient, the bias and the root-mean-square deviation (RMSD) [48]. In the

land community, widely used metrics include the accuracy, expressed in terms of mean bias, the precision, representative of the repeatability of the measurements and computed as standard deviation of the estimates around the true values, and the uncertainty, which is a measure of the actual statistical deviation of the estimate from the truth and is expressed as RMSD [49]. This latter terminology is consistent with the generic uncertainty definition provided in the previous paragraph, since it represents a statistical method for evaluating uncertainties, namely a Type A evaluation, in metrological terms [38]. The validity of such statistical approaches increases as long as the population of the ensemble of satellite and ground-truth collocated data increase in size and spatiotemporal representativeness.

3.4. Spatiotemporal Sampling

The representativeness, i.e., the extent to which a set of measurements taken in a given space-time domain reflect the actual conditions in the same or different space-time domain [50], is a fundamental issue in land product validation, owing to the inherent spatial and geographical heterogeneity and temporal variability of terrestrial surfaces. In order to ensure a good representativeness and high accuracy of the validation dataset, the following aspects should be carefully considered.

- Reference measurement—Since the truth is, strictly speaking, unknown, the general validation process consists in assessing the accuracy of the satellite data against a chosen reference measurement. The basic assumption is that the uncertainty associated to the reference is better constrained than the one of the satellite data. For instance, ground-based sensors can be regularly calibrated (pre- and post-deployment) and closely characterised in laboratory conditions, which is not the case in satellite-borne sensors. In the ideal case, the reference measurement should be traceable to SI unit. In reality, reference data are rarely traceable to a standard and can be affected by various sources of uncertainties, which are very often neglected.
- Spatial sampling—The ideal reference data should provide a meaningful estimate of the investigated geophysical quantity to a level, which is comparable to the satellite measurement scale that is subject of the validation process. This generic requirement drives the choice of the validation sites and the location of in situ observations within each site (spatial sampling). The adoption of community-agreed best practices for in situ data collection allows to address this generic issue in satellite validation. In the land community, and specifically for the validation of bio-geophysical products, the most widely used approach consists in repeating the in situ measurements over the representative sampling unit, the so-called Elementary Sampling Unit (ESU), and upscaling the data averaged over this ESU to a larger scale using empirical transfer functions, based on a combination of high and coarse resolution sensors [51].
- Temporal sampling—In the ideal case, Cal/Val data should be acquired continuously, such as by using automated devices, to increase the number of collocation opportunities (match-ups) with satellite overpasses, as well as to enhance the monitoring of dynamic bio-geophysical processes over the considered ecosystem. The increase of match-up opportunities is essential for improving the reliability of the derived validation results by providing more robust statistical metrics. In reality, a large fraction of available Cal/Val data for land are gathered during ad-hoc field campaigns with limited temporal sampling.
- Spatiotemporal mismatch—Despite the effort in defining an optimised sampling in the spatiotemporal domain, the intrinsic difference between the measurement scale for the satellite and the reference in situ data directly translates into a collocation mis-match error (Σ in Equation (2)), which can be the most significant source of errors for highly dynamic geophysical parameters [48]. For the specific case of terrestrial variables, the temporal mismatch is a second order issue with respect to the spatial sampling, since surface properties change slowly with time. On the other hand, the short-term variability in aerosol and cloud contamination can induce a mismatch temporal collocation error in the derived land bio-geophysical parameters.

- Geographical coverage—The uncertainty budget of satellite EO data should ideally be assessed at global scale in order to verify the accuracy of the used retrieval algorithms over a wide range of biomes and eco-climate conditions. This translates into the need of a network of globally distributed sites, providing continuous, representative and consistent measurements of the same geophysical quantity. The fulfilment of this requirement has some non-negligible costs, which are difficult to cover as part of a single mission or Agency's budget. At the present time, except for the Aerosol Robotic Network (AERONET) [52], there are no operational Land Cal/Val networks entirely fulfilling this need. The pragmatic approach to this problem consists in reinforcing international collaborations, leveraging on existing regional and continental infrastructures, with the objective to build a network-of-networks system. This approach is now being considered in the frame of the Copernicus Ground-Based Observations for Validation (GBOV) [53].

3.5. Fiducial Reference Measurements

The need for a traceable uncertainty for satellite measurements is becoming a pressing requirement for the climate community to derive temporally consistent Climate Data Record (CDR) over multi-decade temporal frame [54,55]. Within this context, the legacy methods for satellite products validation are not *fit-for-purpose*, owing to the lack of traceability, the disparity of protocols and procedures and the inadequacy of uncertainty budget estimation. For this purpose, the concept of Fiducial Reference Measurements (FRM), initially proposed for in situ radiometers measurements of sea surface skin temperatures [56], was generalised within the ESA S3 Validation Team as a set of guidelines to be followed in order to provide to the users the required confidence in satellite products validation. The FRM are defined as: "The suite of independent ground measurements that provide the maximum Return on Investment (ROI) for a satellite mission by delivering, to users, the required confidence in data products, in the form of independent validation results and satellite measurement uncertainty estimation, over the entire end-to-end duration of a satellite mission".

The FRM is a generic concept, which applies to the validation of any satellite-based data product. The defining mandatory characteristics for FRM are the following (see Figure 1):

- FRM measurements have documented SI traceability (e.g., via round-robin characterisation and regular (pre- and post-deployment) calibration of instruments using metrology standards (e.g., [56–58]).
- FRM measurements are independent from the satellite geophysical retrieval process.
- An uncertainty budget for all FRM instruments and derived measurements is available and maintained.
- FRM measurement protocols, procedures and community-wide management practices (measurement, processing, archive, documents, etc.) are defined, published openly and adhered to by FRM instrument deployments.
- FRM are accessible to other researchers, allowing for the independent verification of processing systems.

The full compliancy to FRM requirements may be difficult for some terrestrial ECVs, owing to the lack of SI standards for the derived geophysical parameters and the limited knowledge on calibration uncertainties for some of the commonly used Cal/Val sensors. On the other hand, inter-comparison field campaigns and the reinforced focus on pre- and post-deployment calibration and detailed analysis of the error budget shall be addressed to move towards the FRM concept.

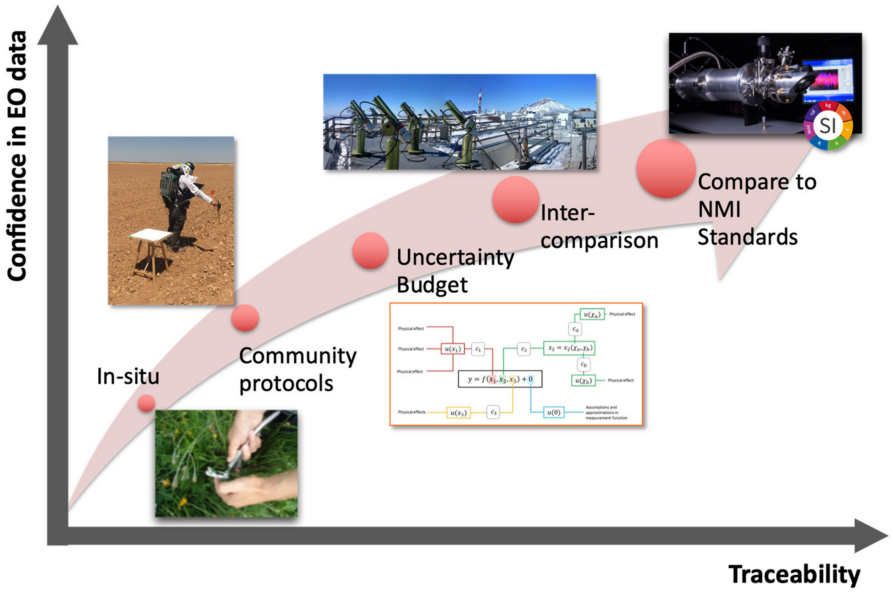


Figure 1. The figure shows a schematic representation of FRM concept. As long as the community-endorsed best practices and metrological guidelines are adopted, e.g., through uncertainty budget analysis and regular inter-comparisons, the confidence in the reference Cal/Val data is enhanced with a resulting increased confidence in the satellite EO products being validated. The traceability to SI is eventually attained when comparisons are made and regularly maintained against internationally agreed upon standards provided by National Metrological Institutes (NMI).

3.6. Validation Stage and Maturity Matrix

As the set of validation reference data increases in spatial and temporal representativeness, and the documented traceability and uncertainty estimate are provided following FRM best practices, the confidence in the relevant EO satellite products increases, but how can we best define this confidence level for EO data? This question was tackled under several projects in recent years [59–63], and various approaches were proposed on how to express the maturity and usability of a given EO dataset. Among the various attempts, the following are worth mentioning and will be considered for elaborating the ESA Cal/Val strategy for land:

- CEOS LPV validation stage [59]—One of the most appropriate and suitable definitions of confidence level for terrestrial ECVs is the validation stage concept, defined and put forward by the CEOS-LPV sub-group. The increase in the stage, from 0 to 4, is driven by the use of more representative sets of reference data and by the adoption of rigorous approaches for the uncertainty estimate. The validation stage should be confirmed and endorsed by the scientific community through peer-reviewed publication, and the validation result should be periodically reviewed and maintained as long as new versions of satellite EO data become available.
- CCI Maturity Matrix [60]—The Maturity Matrix (MM) concept, tailored in the frame of ESA Climate Change Initiative (CCI) [61], is a systematic approach for assessing the usability of any CDR. The CCI MM offers a standard and efficient way to quantify for a given climate dataset six attributes: software readiness, data format, documentation, uncertainty characterisation, accessibility and usability. The maturity level is rated via a score (from 1 to 6), which is assigned on each aspect of the data record. The CCI MM provides an overview on the status of climate datasets, highlighting areas for further improvement and allowing to compare the fitness for purpose of different products.
- GAIA CLIM Maturity Matrix [62]—The GAIA-CLIM MM was specifically designed for assessing the level of maturity of ground-based atmospheric composition networks. The assessment is based on the following aspects: metadata, documentation, uncertainty characterisation, public access, feedback and update, usage, sustainability

and software. The MM can be used both to aid user decision making as to which measurement series are suitable for particular use cases and support networks' owners in improving the maturity of their measurements, facilitating interoperability across existing ground-based measurement systems.

- EDAP Product Quality Evaluation Matrix [63]—The Earthnet Data Assessment Pilot (EDAP) matrix, which is an evolution of the QA4ECV matrix [40], aims at providing to the users a readily accessible overview of the relevant EO satellite data products quality and fitness for purpose. The quality assessment addresses specifically the following main EO products characteristics: format and accessibility, products and algorithm documentation and quality flags, uncertainty characterisation and validation results. For each feature of the identified product, a ranking is provided with colour coding, providing easy and quick information on the overall quality of the considered EO data. While the previous definitions are mainly focusing on bio-geophysical variables, the EDAP quality matrix is a cross-cutting concept, which can be equally applied to any satellite-based product level. This matrix contributes to the effort of standardisation in the quality assessment procedures for EO satellite data. The EDAP matrix is therefore pivotal in moving towards enhanced interoperability in the EO domain.

3.7. Analysis Ready Data

A pragmatic approach to the interoperability of satellite-based EO data consists in identifying a minimum set of requirements, to which all data providers should adhere, in order to ease the synergistic use of products from a variety of sensors. The Space Agencies have started addressing this issue with increasing priority through the definition of a new paradigm in satellite products: Analysis Ready Data (ARD). The definition of ARD is still not fully consolidated by the various data providers, since several interpretations were formulated. Specifically, when confronted with the emerging commercial data providers, the ARD concept becomes debatable, and it was the subject of discussion during the recent ESA Very-High Resolution (VHR) Workshop [64]. On the other hand, there is a shared understanding of the underlying basic principle, namely, ARD is meant to reduce the user's effort in the preparation and preprocessing of satellite data by providing a globally and temporally consistent stream of data, calibrated and processed to a standardised level, enabling the users to directly generate added-value information from the spatial data, e.g., maps, time-series.

This basic principle translates into a number of requirements on the data processing and products definition, such as:

- Radiometric and geometric calibration issues should be solved following best practices approaches.
- A minimum level of processing should be applied in order to lower the bar of expertise on the user side. For optical sensors, this typically includes cloud screening and atmospheric correction.
- Spatially explicit per-pixel quality flags should be provided to enable users to perform informative processing on suitable good quality data.

There are on-going international efforts to further tailor these general principles, providing detailed requirements for ARD. In particular, the CARD4L initiative [31,65] is the most advanced and is progressively becoming the standard within the Institutional Space domain. The CARD4L requirements were defined for a number of product families, among which the land surface reflectances are extremely relevant for discussing the Cal/Val needs for satellite-based Level-2 products. The last reprocessed Landsat-8 dataset, generated by the United States Geological Survey (USGS) adhere to the CARD4L product specifications for surface reflectances [66], and the same will be achieved with the upcoming ESA Sentinel-2 reprocessing activity.

3.8. Multi-Sensor Data Harmonisation

The compliancy to ARD requirements is the first fundamental step for addressing the concept of interoperability at the satellite level. Yet, two interoperable datasets cannot be merged seamlessly until the sensor-specific features are appropriately characterised and adjusted by applying the relevant correction factors. For optical sensors, these features typically include: radiometric calibration, geometric co-registration, spectral bands definition, spatial resolution and sun-viewing conditions. Harmonisation is the process of solving such sensor-specific features in order to blend different mission datasets in a consistent manner. The need for multi-mission harmonisation is a generic concept, which is relevant to any satellite product level (L1 and L2), and it is notably being addressed in the context of climate studies and regional applications, as further discussed below.

- Harmonisation in a climate context—Harmonisation has a strong relevance for the generation of CDR, when multi-sensors datasets are merged to derive a consistent series of observations, spanning a time frame relevant for climate studies, typically more than 20 years. In this context, special attention is given to rigorous estimation of uncertainties to reliably disentangled sensor-related artefacts from climate-relevant trends [55]. A generic framework for harmonising the series of AVHRR (Advanced Very-High-Resolution Radiometer) radiance observations was developed in the frame of the Fiduceo project [67,68]. This project has further developed its own terminology, distinguishing between "harmonisation" and "homogenisation" of satellite data series. Following Fiduceo's terminology, a harmonised satellite series is one where all the calibrations of the sensors have been made consistent with a reference dataset, which can be traced back to known reference sources, in an ideal case back to SI. Unlike harmonisation, homogenisation consists in adding the constraint that all satellites would give the exact same signal when looking at the same location at the same time. According to Fiduceo's recommendations, homogenisation should be avoided for CDR generation, since the application of adjustment factors can induce scene-dependent biases that are difficult to characterise a-posteriori.
- Harmonisation in an application context—In addition to climate studies, harmonisation is increasingly relevant for regional applications, where observations from a virtual constellation of similar sensors are blended with the primary goal to enhance revisit time. The most obvious examples of such a strategy are the NASA Harmonised Landsat and Sentinel-2 (HLS) product [69] and the ESA Sen2Like project [10], aiming at similar objectives of harmonising and fusing the Sentinel-2 and Landsat-8 data streams. Harmonisation within these two projects means providing a seamless stream of sensor-agnostic data from the two missions to the users, gridded to a common map projection, atmospherically corrected and cloud screened using a single algorithm, normalised to a fixed sun-viewing geometry and spectrally adjusted.

4. Current Status and Identified Needs

The state-of-the-art in terms of Cal/Val infrastructure and methods for optical land imaging sensors and the current needs and gaps are reported and discussed within this chapter. The needs are identified and grouped per products level, i.e., starting from Level 1 and moving to Level 2 surface reflectances and bio-geophysical variables. This is in line with the ESA's bottom-up approach to Cal/Val, which starts with the verification of cross-sensors consistency at TOA level and continues along the processing chain up to the bio-geophysical variables. This approach ensures full traceability and detailed characterization of the various uncertainty contributions and their propagation through the different products levels.

4.1. Level 1 TOA Radiometry and Geometry

The need for blending optical data from multiple sources directly translates into the need for ensuring that the two datasets are consistent in terms of geometric and radiometric calibration. Any geometric misregistration will induce a geometric systematic error when

mixed-source temporal series have to be used for regional applications. Likewise, inter-sensor radiometric inconsistencies will translate into biases in the derived harmonised multi-mission data record. While the root cause of the radiometric or geometric errors can be considered systematic, their combined impact into the projected and calibrated radiances or reflectances (Level 1 product) have a more complex nature, notably owing to the interpolation and regridding process.

From the ESA's perspective, ensuring radiometric and geometric consistency at Level 1 across its fleet of optical land sensors is of the utmost importance, and it is considered the first fundamental step for tackling the issue of interoperability. This approach is shared among the other Space Agencies, and it is the main principle followed in the frame of the CEOS-LSI-VC and GSICS working groups.

The issue of inter-sensor radiometric consistency has been addressed for several decades, and there is now a good level of maturity in terms of the required pre- and post-launch Cal/Val activities. On the other hand, we have not yet reached the stringent accuracy requirements for some applications, notably for climate applications, and challenges remain in properly characterising geometric mismatch errors for cross-sensor consistency analysis and when propagating the uncertainty at Level 1 into higher-level spatially-aggregated climate datasets [55].

Within this section, we present the current status and highlight the remaining needs in terms of infrastructure and methodologies for the validation of Level 1 TOA products. The needs identified in the text and the derived recommendations are highlighted in Table 2 at the end of the section.

4.1.1. Pre-Flight Calibration

Radiometric calibration is the result of a long and complex process, starting with the rigorous pre-launch characterisation activities. The aim of these activities is to characterise the sensor response in a range of operational conditions that mimic those that will be experienced in space. Any flaws in the acquisition and analysis of pre-launch calibration data or lack of specific characterisation measurements could induce a potential systematic error in the resulting higher-level data products. This error will be difficult to correct in-orbit, since the relevant characterisation measurements could not be feasible in space or could be performed with limited accuracy as compared to the one that can be attained in laboratory conditions.

It is therefore essential to pay extreme attention and ensure appropriate support to pre-launch characterisation activities. In addition, it is important from a user perspective, to guarantee full visibility of the results of the pre-launch characterisation measurements, both the raw data and the derived calibration factors, to allow a full traceability along the processing chain. As a matter of fact, several recommendations were identified during the Fiduceo project [67] on the importance of pre-flight calibration activities and on the need to store and preserve this information for future mission reprocessing campaigns. The Space Agencies should embrace such recommendations and ensure that metrological best practices are embedded in the early phase of the mission design.

4.1.2. In-Orbit Calibration

The SI-traceability breaks down, in rigorous terms, when the sensor is launched into orbit and exposed to the harsh space environment. As a matter of fact, while full traceability to SI units can be ensured and regularly maintained on-ground, through comparison to primary standards in well-controlled laboratory conditions, the same SI-traceability cannot be achieved in orbit. In practice, dedicated on-board devices are designed to allow the calibration factors to be regularly updated and improved post-launch. Despite not being fully SI-traceable, these on-board devices allow to account for the natural ageing of the satellite sensor during the mission lifetime and to closely assess its radiometric and spectral response.

In-orbit radiometric calibration of optical sensors in the visible and short-wave infrared domains is carried out routinely using diffuser panels, which have nearly-Lambertian reflectance properties and are exposed on a regular basis to the Sun to provide the reference measurement for calibrating the sensor radiometric and spectral response of the Earth-view. In the real case scenario, the on-board diffusers can exhibit a non-Lambertian behaviour, e.g., due to imperfections of the used coating material, and can be subject of ageing due to long exposure to the space environment. The availability of secondary diffusers, which are exposed less frequently to ensure pristine reference measurements, allow for improving the accuracy of the in-flight radiometric calibration and for accurately characterising the ageing of the operational diffuser [70,71].

However, some space-borne systems lack for on-board calibration devices, see for instance Proba-V [72] and the large majority of small-satellite constellations. Furthermore, the on-board calibration system provides a reference source, which is, by definition, applicable only to the considered optical sensor and cannot be transferred to another sensor. In such cases, the use of independent calibration targets (*vicarious targets*), accessible by all sensors, is crucial for absolute radiometric assessment in case of missing or damaged on-board diffuser(s) and for establishing a common radiometric scale to estimate inter-sensor radiometric biases.

The vicarious targets are temporally stable, spatially and spectrally uniform Earth and planetary targets whose radiometric signal can be accurately modelled with a high degree of confidence. In the case of Earth's targets, the vicarious calibration methods rely on a Radiative Transfer Model (RTM), which is used to simulate the at-sensor signal over the considered site. This simulated value is used to validate the radiometry of the considered sensors, to monitor its variation over time and to inter-compare various sensors observing the same targets [73]. Vicarious approaches are systematically used at ESA for radiometric assessment and intercalibration of S2 and S3 optical sensors, using the DIMITRI (Database for Imaging Multi-spectral Instruments and Tools for Radiometric Intercomparison) tool [74].

An exhaustive review of the natural targets, which are currently used for vicarious calibration, is out of scope for this paper. We recall in the following paragraphs the most commonly used methods and the current and planned initiatives to address the remaining challenges for assessing radiometric accuracy at TOA level.

4.1.3. Vicarious Calibration with Desert Targets

The most widely used vicarious targets are the Pseudo-Invariant Calibration Sites (PICS); these are desert sites, which were carefully selected based on their low cloud coverage, the high reflectance, and the good stability and uniformity in terms of temporal, spatial and spectral characteristics [75]. The six CEOS-endorsed reference PICS sites are Libya-4, Mauritania-1, Mauritania-2, Algeria-3, Libya-1 and Algeria-5 [76], which were chosen based on the large number of data sets from multiple sensors that already exist in the archives and the long history of characterisation performed over these sites. The PICS have high reflectance and are usually made up of sand dunes with climatologically low aerosol loading and practically no vegetation. Consequently, these PICS are highly suitable to evaluate the long-term stability of an optical sensor in the visible and short-wave infrared spectral regions and to facilitate inter-comparison with other sensors [73].

However, the PICS are not fully invariant, since dunes topography or sand grain size and composition can induce residual directional and shadowing effects, with intra-annual variations related to the seasonal changes of illumination conditions and potential inter-annual changes caused by dune displacement [77]. It is thus generally recognised that the list of endorsed PICS sites should be critically reviewed in order to improve the accuracy of the derived radiometric assessment. This critical review was recently undertaken as part of an ESA-funded study [78] and it is also being pursued in the frame of a CEOS supported initiative [79].

Another challenge in vicarious approaches using PICS resides in the inaccuracies of the RTMs used to simulate the at-sensor signal. These are induced by approximations in modelling the atmosphere–surface radiative interactions (e.g., plane parallel atmosphere) and the use of inaccurate ancillary data (e.g., atmospheric concentration profiles, molecular cross-sections, aerosol properties). As a matter of fact, the current discrepancies between commonly used RTMs for simulating the radiance over the Libya-4 desert site has been demonstrated to be in the order of 2 to 3% [80]. These challenges were the main motivations to start the development of a new open-source Monte-Carlo ray-tracing 3-D RTM, called Eradiate, which is currently being supported by ESA and the European Commission (EC) [81].

4.1.4. Vicarious Calibration with Deep Convective Clouds

Deep Convective Clouds (DCCs) are vertically extended and extremely opaque clouds (optical thickness of the order of 100). They exhibit bright and almost white spectra from the visible to the near infrared with nearly isotropic reflectivity. They are therefore considered the ideal invariant Earth target for vicarious calibration, owing to the very high signal-to-noise ratio [82].

DCCs have a number of advantages over PICS, since they are very close to be Lambertian, except for large viewing angles, when Bidirectional Reflectance Distribution Function (BRDF) correction should be applied [83]. Moreover, DCCs are naturally located at very high altitudes, which strongly simplifies the atmospheric correction procedure, thus reducing the uncertainties associated with PICS. Finally, DCCs are more evenly distributed globally in the tropics, which is particularly important for geostationary satellites, as many of these satellites cannot view PICS under favourable conditions.

DCCs have been used for inter-band calibration for more than two decades, mainly for coarse resolution satellites (see pioneering work from Vermote et al. [84]), and they were also successfully applied to S2 and S3 missions (see recent work by Lamquin et al. [85,86]). The relevant DCC calibration protocols and the used modelling approaches are well consolidated and mature enough, allowing the assessment of inter-band radiometric accuracy at the level of 1–2% and also to study long-term ageing effects. Some remaining challenges include the refinement of BRDF correction approaches for large viewing angles to remove underlying seasonal variability [83]. The GSICS group is currently working to address this remaining challenge and consolidate the DCC protocol for calibration.

4.1.5. Vicarious Calibration with the Moon

Another category of invariant vicarious targets are the celestial bodies, such as the Moon [87] or the stars [88], many of which are truly invariant for most of the practical purposes. They offer radiometric references that are independent of the Earth's surface reflectance and atmospheric variations, therefore reducing the intrinsic uncertainties associated with the Earth's vicarious targets. Among the various celestial bodies, the Moon is an ideal target for stability monitoring since its variations in brightness due to phase angle, non-uniform albedo, distance and the lunar librations can be accurately characterised with models, such as the widely used Robotic Lunar Observatory (ROLO) model, developed by USGS [87].

The Moon has been and still is routinely used for monitoring the on-orbit calibration stability of various optical sensors, such as the Moderate Resolution Imaging Spectroradiometer (MODIS) [89], the Landsat-8 Operational Land Imager (OLI) [90] and Proba-V [72]. For these sensors, the use of the Moon has proven extremely successful, providing trending accuracy better than PICS. For instance, the MODIS Terra Collection 6 recalibration was largely based on the Moon for long-term trending analysis in the solar bands [89]. Similarly, the moon observations will be exploited in the future S2 and S3 satellites. Regular acquisitions of the Moon are now becoming standard also in the operational planning of commercial Smallsat constellations, such as for the Planet Dove satellites [91].

Despite the interest in using the Moon as a vicarious target, the currently available ROLO model's accuracy (estimated uncertainty is 5–10%) does not meet the requirement for an absolute radiometric assessment [92–94], hence, this method is still predominantly used for stability monitoring and intercalibration. In order to overcome this limitation, EUMETSAT, in collaboration with USGS, developed an improved version of the ROLO model. The new lunar model, named GIRO (GSICS Implementation of the ROLO model) [95,96], has been endorsed as the established community-available reference for lunar calibration. Future updates of the GIRO model are currently being considered with the target to attain the required radiometric accuracy (better than 1%) for using the Moon as an absolute calibration target. In parallel, ESA has run a project dedicated to improving current lunar observations and modelling [97]. The resulting Lunar Irradiance Model ESA (LIME), has a radiometric uncertainty at 2-sigma level of about 2%, as estimated with a Monte-Carlo method [98]. The LIME model was recently used for calibrating S3 OLCI data [99], and its accuracy is being evaluated in the frame of a current GSICS inter-comparison exercise with ROLO, GIRO and other Space Agencies' lunar models [100].

4.1.6. Cross-Sensors Intercalibration and Tandem Phase

The basic assumption of cross-sensors intercalibration is that two instruments should provide the same TOA signal when viewing the same target on-ground at the same time and with the identical spatial and spectral response [73]. Eventual temporal and spectral mismatches should be accurately accounted for and corrected to properly assess radiometric biases [101]. The vicarious targets introduced in the above paragraphs are largely exploited for intercalibration using the RTM as a transfer reference to compare the radiometric signal of the two systems under analysis.

Among the various intercalibration methodologies, the simultaneous nadir overpass (SNO) approach has proven very effective, in particular in the thermal infrared domain, as well as for solar reflective bands [73]. The advantages of SNO as compared to other methods is that the same target on Earth is viewed at nearly the same time and in nadir viewing conditions, therefore minimising the impact of residual atmospheric and directional effects in the RTM simulations. SNO approaches can yields to accurate intercalibration coefficients with an uncertainty of about 1% in the VNIR range [102].

The drawbacks of the SNO techniques are the limited spatial coverage of space and time collocated overlaps, notably for polar-orbiting satellites. Collocation criteria can be relaxed both in space and time domains, and the corresponding natural variability can be modelled through RTMs, albeit at the expense of the degraded accuracy of the inter-comparison analysis. When instruments are placed in the exact same orbit flying at a short time interval, a very accurate intercalibration analysis can be performed with full characterisation of the uncertainty budget. This is the concept of the Tandem Phase, which was successfully implemented during the Commissioning Phase of Sentinel-3 twin satellite constellation with the objective to ensure consistency of the data provided by the two satellites. The Tandem Phase has proven to be the best way to ensure the traceability of the measurements and the in-depth characterisation and validation of their estimated uncertainties. These are critical issues for building a reliable long-term data record as required for climate applications [103]. For S3 optical sensors (OLCI and SLSTR), this phase was essential for advancing in the harmonisation and homogenisation of the data provided by the two satellites (A/B), both at radiometric level and at geophysical products level [104–106].

4.1.7. Instrumented Sites for TOA Radiometric Calibration

The RadCalNet network [107] ensures free and open access to SI-traceable TOA spectrally resolved reflectances with associated uncertainties over five globally spread sites. These sites provide nadir-view TOA reflectances every 30 min from 9:00 AM to 3:00 PM local time at 10 nm intervals in the spectral range from 400 nm to 2500 nm [108]. The TOA reflectances are computed from ground-based spectral measurements, using a common

methodology for propagating the surface reflectance to TOA level [107]. Each member site takes responsibility for the quality assurance of the surface/atmosphere measurements provided and is subject to peer review and rigorous comparison to ensure site-to-site consistency and SI traceability.

The RadCalNet initiative is a successful example of international collaboration between Space Agencies coordinated in the frame of the CEOS-WGCV, fulfilling a clear need from the scientific community. RadCalNet measurements are now being extensively used for systematic assessment of TOA radiometry for a large number of institutional and commercial space missions, with a main focus on high-spatial resolution sensors [109–112]. Evolutions of the RadCalNet service were suggested by the EO community [109,110], including, in particular, the provision of TOA measurements at enhanced spectral sampling intervals and the inclusion of a full BRDF characterisation to improve the radiometric assessment of off-nadir observations. These proposed improvements are now being considered within RadCalNet evolution activities, together with the inclusion of new sites to increase global coverage.

4.1.8. SI-Traceability in Space

Our ability to accurately assess TOA radiometry will make an unprecedented leap forward in the future with the launch of the Traceable Radiometry Underpinning Terrestrial- and Helio- Studies (TRUTHS) mission [113]. This UK-lead ESA mission will allow ultimately to attain the required accuracy (<1% at surface reflectance level), which is needed for climate and solar studies [114]. The TRUTHS platform will carry on board a hyperspectral imager measuring spectrally resolved incoming and reflected solar radiation (in the 320 nm to 2350 nm range) at high spatial resolution. TRUTHS and CLARREO Pathfinder [115,116] will be the first missions to be calibrated directly traceable to SI units, in TRUTHS's case, via a primary standard in orbit, therefore providing, for the first time, an absolute high-accuracy traceable calibration reference to anchor any optical sensor to a common radiometric scale. TRUTHS will be the final missing element to complete ESA's Cal/Val strategy at Level 1, complementing the existing elements (vicarious targets, ground-based network, on-board devices, pre-flight calibrations) and allowing to reach the final goal of SI-traceability.

Table 2. Identified needs and recommendations for assessing the quality of Level 1 TOA products.

ID	Status and Needs	Recommendations
IDN-01	Pre-launch calibration is critical for ensuring accurate satellite observations, meeting the requirements of the science community. Though additional effort should be spent in fostering the adoption of metrological practices during pre-launch characterisation activities.	Space Agencies should embrace metrological practices in the early phase of the mission design by including relevant requirements in Mission and System Requirement Documents (MRD, SRD) and allocating appropriate budget to pre-launch calibration activities.
IDN-02	Pre-launch calibration database is most of the time undisclosed to users or protected by property rights with resulting difficulty in applying a full traceability tree in the Level 1 processing.	Space Agencies should ease traceability of calibration processing by: requiring instrument providers to identify measurement equation and develop traceability chain, archiving, maintaining and making accessible to users the relevant pre-launch characterisation database.
IDN-03	While radiometric uncertainty is well characterised, theoretical advances needs to be made for fully characterising geometric errors and spatial correlation induced by interpolation and regridding.	Space Agencies should foster advances of theoretical approaches for fully characterising geometric and spatially correlated errors in Level 1 products, such as those induced by orthorectification, regridding, projection.

Table 2. *Cont.*

ID	Status and Needs	Recommendations
IDN-04	There is currently a need for providing per-pixel uncertainty estimates at Level 1. This is a critical input for consistently assimilating satellite observations into geophysical and weather forecast models.	Space Agencies should work towards the provision of uncertainty information at pixel-level in Level 1 products for optical imaging sensors.
IDN-05	PICS are invaluable targets for assessing radiometric accuracy and temporal stability and for cross-calibration. Though, used RTMs should be improved in order to meet the accuracy requirement set by the climate community (1%).	Space Agencies, in coordination with CEOS-WGCV, should work towards improving accuracy of current RTM simulations for Cal/Val applications with initial focus on TOA radiometric assessment over PICS.
IDN-06	The Moon is the best vicarious target for optical sensor stability monitoring and inter-calibration. However, the accuracy of the current (ROLO) lunar irradiance model still prevents its use for absolute calibration. New observations are urgently required for using the Moon as an absolute radiometric target.	Space Agencies, in coordination with GSICS and CEOS-WGCV, should support improvements in Lunar irradiance models and inter-comparison exercises with the target of using the Moon as absolute target for radiometric calibration, reaching the target of 1% in radiometric accuracy.

4.2. Level 2 Land Surface Reflectances

The first satellite-based Level 2 products for land applications are the Bottom-Of-Atmosphere (BOA) products, often referred as Surface Reflectance (SR). The BOA products are the output of the Atmospheric Correction (AC) process, consisting in the estimation and removal of the atmospheric contribution to the TOA at-sensor radiometric signal. The uncertainty budget associated to the atmospheric correction includes all approximations used to simulate the inherent radiative transfer process in the coupled surface–atmosphere system, the uncertainties in the estimation of the atmospheric trace gases and particle concentrations and their scattering and absorption properties. Typically, in the solar spectrum domain and for the commonly used VIS, NIR and SWIR spectral bands, the uncertainty budget is dominated by the uncertainty in the estimation of the aerosol and water vapour optical properties. The rigorous estimates of uncertainties in SR products are a key input to be propagated into the subsequent bio-geophysical retrieval algorithms. Nevertheless, as of today, no operational ESA SR products are provided with an uncertainty estimate at pixel level, and this is considered a major drawback for a correct assimilation of this dataset into geophysical models.

Furthermore, BRDF correction of SR products and normalisation to a standard illumination and viewing geometry is a prerequisite for blending data from multiple sensors and for minimising high frequency "noise" in temporal series [117]. This additional correction is not included yet as part of the currently available ESA SR products (see Table 1), which are provided as atmospherically corrected reflectances at the sun-viewing conditions of the satellite measurement. The provision of BRDF-corrected SR products, normalised to a standard sun-viewing geometry, is being addressed in the frame of the Sen2Like project for harmonising Sentinel-2 and Landsat-8 time-series [12], and it is under development for both the S2 and S3 SR products as part of a recent ESA/Copernicus project, started in early 2021 and named COPA (Copernicus: 4 Core Products Algorithm Studies).

Within this section, we present the current status and highlight the needs of infrastructure and methodologies for the validation of SR products. The needs identified in the text are listed in the Table 3 at the end of the section with derived recommendations.

4.2.1. SR Assessment within ACIX Initiative

The lack of a global network of ground-based surface reflectance measurements over land is currently considered the most urgent data gap for assessing the radiometric accuracy of current and future ESA optical land imaging sensors at the BOA level, as currently pointed out in a recent workshop [3]. In the absence of such ground-based infrastructure, the common approach is an indirect validation, as developed in the frame

of the joint ESA-NASA Atmospheric Correction Inter-Comparison Exercise (ACIX) with focus on the Sentinel-2 and Landsat-8 (L8) missions [118].

The validation of SR within ACIX is performed over a set of AERONET stations [52], globally spread to be representative of different surface and climatological conditions. The AERONET network provides a reliable, globally representative and consistent dataset of atmospheric variables that allows for the validation of the performance of an atmospheric correction processor using common statistical metrics. In the frame of ACIX, a "ground-truth" dataset of synthetic SR was computed with the 6SV RTM [119] using as input data the Aerosol Optical Thickness (AOT), the aerosol micro-physical properties and Total Column Water Vapor (TCWV) provided by AERONET retrievals. A protocol was consolidated within ACIX allowing to quantitatively estimate the uncertainty budget in SR products in terms of Accuracy, Precision and Uncertainty (APU) metrics [49] and to compare this estimate against mission requirements.

The first ACIX initiative, completed in 2018, was successful in identifying drawbacks and limitations of current AC codes for S2 and L8 missions, helping the algorithm providers to improve their methods. Typical limitations include the use of Dark Dense Vegetation (DDV) approaches for AOT estimation that fail over non-vegetated scenes or the choice of predefined aerosol models. Owing to this success, a new ACIX-II initiative was designed, building on the lessons learnt gathered during the first exercise [120]. Within ACIX-II, a clear distinction was made, and different inter-comparisons were run for land and aquatic reflectances and cloud masks. In the land domain, the ensemble of the considered AERONET stations was significantly increased, with the objective to verify the accuracy of the AC algorithms at a global scale. Furthermore, the ground-based measurements were included for the direct validation of both the land and water reflectances. Over land, these observations were taken from the RadCalNet network [107]. ACIX-II is now completed, and the relevant papers are now published for the aquatic reflectances [121] and in preparation for the land domain and for the cloud mask. As a result of ACIX-II, additional recommendations and needs were identified on how to pursue the effort in assessing the uncertainty budget of satellite-derived SR products. We recall here the recommendations, which are relevant in the land domain.

The use of AERONET-derived SR is a pragmatic solution, and it provides a robust and fair protocol for inter-comparing the different AC codes. On the other hand, there is a clear need for having a direct validation approach, using independent ground-based surface reflectance measurements. The RadCalNet network [107], which was primarily designed for assessing TOA radiometry, cannot entirely fulfil this need, mainly because the RadCalNet stations are generally located over non-vegetated, flat and spatially homogeneous areas, except for La Crau station, with very limited cloud coverage and aerosol loading. This scenario represents the almost ideal conditions and is thus not suitable to properly challenge the performances of the AC codes for the majority of conditions. The requirement for SR validation is to have vegetated sites representative of a wide range of surface and atmosphere realistic conditions and potentially also spatial heterogeneous sites to study the adjacency effects, which are difficult to properly model within the AC algorithms.

The adoption of the 6SV RTM [119] for generating the "ground-truth" over AERONET stations could potentially induce biases when assessing the accuracy of those codes, which use different RTMs for estimating and subtracting the atmospheric contribution from the TOA signal. Typical RTMs being used by the ACIX community include LibRadTran [122] and MODTRAN [123]. Overall, there is a need of inter-comparing the accuracy of different RTMs against a community-agreed benchmark algorithm. While similar exercises were done in the past [124], the benchmarking should now be repeated to assess how much RTM discrepancies propagate in the uncertainty budget of the output SR products, with focus on the current and future optical land imaging sensors. To this purpose, ESA in coordination with the EC, is currently supporting and promoting the RAdiation transfer Model Intercomparison for Atmosphere (RAMI4ATM) initiative [125]. The primary goal of RAMI4ATM will be to document the variability between coupled surface–atmosphere

RTMs under well-controlled, but realistic, conditions. Within RAMI4ATM, the surface properties will be defined by the simple homogeneous scenes, as defined within the RAMI-V exercise [126].

4.2.2. Cloud Mask Assessment within CMIX

All AC codes in the optical domain typically work under the assumption of clear sky-conditions since the radiative impact of the clouds is not accounted for within the RTM. As a result, failure in accurately masking cloud-contaminated pixels will lead to uncertainties in the atmospherically corrected reflectances. Therefore, as part of ACIX-II, a dedicated initiative, named Cloud Mask Inter-Comparison Exercise (CMIX) [127], was defined to assess the accuracy of the currently used cloud detection schemes for the S2 and L8 missions. Despite not being an objective metric, the intercomparison of cloud masks is a widely used method to assess the accuracy of the cloud detection schemes, and it was the subject of various papers in recent years [128,129]. The first CMIX exercise is now completed, and the outcomes are being summarised in a scientific paper. Some of the recommendations gathered as part of the CMIX, which are relevant for land applications, are here reported and discussed.

While cloud contamination is generally easily detected for optically thick clouds, using the spectral information available within S2 and L8 bands, issues arise in the case of optically thin clouds (often referred to as semi-transparent) or for pixels nearby a cloud. In these cases, the performances of the different cloud mask algorithms strongly differ. Two conditions generally occur: on one side, clear-sky conservative algorithms overestimate the cloud cover; on the other side, cloud-conservative algorithms tend to underestimate the cloud presence. The choice between these two extremes is mostly driven by the needs of the subsequent application and the sensitivity of higher-level algorithms to residual cloud-contaminated pixels. As a rule of thumb, clear-sky conservative approaches are better suited for land and water applications, while for studying cloud optical properties, cloud-conservative approaches may be the preferred choice. With respect to pixels at the cloud edges, a practical method consists in dilating the cloud mask for a certain number of pixels to limit the impact of adjacency effects, i.e., increase of radiance for pixels nearby cloud boundaries, this is for instance the approach used by some of the CMIX algorithms [130,131].

The problem of the cloud mask ultimately reduces to the basic question of defining what is a cloud, or rather, what is a cloud for the target application. This is a non-trivial question, and it has not been fully answered yet for optical land imaging sensors. In the meteorological community, a cloud is generally defined as: "a visible aggregate of minute water droplets or ice particles in the atmosphere above the Earth's surface" [132]. The issue remains however, that being visible depends on the observation conditions and on the remote sensing device being used. For instance, active Light Detection and Ranging (LiDAR) sensors provide enhanced capabilities for observing cloud optical and micro-physical properties not feasible with passive sensing techniques.

In the absence of an agreed definition of a "cloud", CMIX adopted a pragmatic approach for validating and inter-comparing the different cloud masks using best practices and the reference data available in the literature. The approach was mostly based on the use of manually labelled S2 and L8 scenes, such as the Hollestein dataset for S2 [133] and the Biome dataset for L8 [134]. The labelling of the scenes into different classes (cloud, semi-transparent, clear pixel) was performed by experts. This labelling was considered to be the ground-truth against which all the cloud masks were compared. This approach clearly lacks an objective and physically based definition of cloud properties, since the manual labelling process is subjective and error-prone. Furthermore, inconsistencies were verified among the currently available labelled datasets for S2 and L8, owing to the intrinsic subjectivity of the photo-interpretation process.

The take-home message from CMIX was that our ability of inter-comparing cloud-masks is currently hampered by the lack of a community-agreed cloud definition and an

associated reference dataset. In the ideal case, this reference dataset should be collected from independent ground-based measurements and should provide a physically based measure of the cloud properties, such as the optical thickness. One cost-effective solution, based on the use of ground-based fish-eye cameras, was recently proposed by NASA and initially used within CMIX [135]. In the absence of independent datasets, there is an urgent need for an accurately labelled ground-truth dataset, specifically tailored to the validation needs. In order to address the remaining needs and gaps, a follow-up CMIX-II initiative is being defined by ESA in coordination with the CEOS-WGCV.

4.2.3. HYPERNETS Project

The HYPERNETS project [136,137] aims at developing a global automated network of ground-based hyperspectral radiometers, measuring water and land bidirectional surface reflectance for supporting the validation of satellite-based SR products. The radiometers will be equipped with an embedded calibration device and a pointing system, allowing it to cover a wide range of viewing and azimuth angles for the full characterisation of the surface BRDF.

The HYPERNETS network will ensure accurate SR validation data over land and water sites, supporting the needs of any space-borne optical mission and filling a well-recognised data gap in the current Cal/Val ground-based infrastructure. In the water domain, the AERONET-OC network [138] has been providing long-term continuous data for validation of water leaving reflectance. However, AERONET-OC multi-spectral measurements were specifically tailored to the needs of ocean colour narrow-band sensors, such as MERIS, OLCI or SeaWIFS, and they are not well suited for validating broad-band high-resolution sensors, such as S2 or L8. For these sensors, the common approach consists of accounting for the differences in spectral band definition using band adjustment procedures [139,140], although this introduces uncertainties in the validation analysis. Hyper-spectral measurements from HYPERNETS will allow users to overcome these limitations, by ensuring suitable validation data for any spectral bands in the solar spectrum domain and by filling the Cal/Val needs of future hyperspectral missions, such as ESA's CHIME mission.

At the time of writing this paper, preparatory activities are on-going for setting up the future HYPERNETS network; this includes the deployment of first autonomous hyperspectral radiometers in coastal and inland waters and initial exploitation for validation of aquatic reflectances derived from high-resolution satellite data [141]. The project will continue in the coming years with evolution of the instrumental set-up, using more advanced hyperspectral systems with additional SWIR channels for land observations and with the progressive deployment in a number of land and water sites. The location of the land sites was defined to cover a wide range of surface types, aerosol loading and ecoclimatic areas for challenging the performances of the AC algorithms in realistic conditions.

Table 3. Identified needs and derived recommendations for assessing quality of land SR products.

ID	Status and Needs	Recommendations
IDN-07	The lack of ground-based surface reflectance measurements over land is currently limiting our capability to validate SR products and assess the quality of AC algorithms. Such measurements should ideally be: continuous (increase match-up), hyper-spectral (multi-mission) and multi-angular (for BRDF characterisation).	Space Agencies, in coordination with CEOS-LPV, should support on-going efforts (e.g., HYPERNETS) working towards setting up a globally representative network of ground-based measurements for supporting the validation of satellite-derived land SR products.

Table 3. *Cont.*

ID	Status and Needs	Recommendations
IDN-08	Validation of cloud masks is currently based on the use of manually labelled reference ground-truth data with an obvious lack of an objective and a physically based approach.	Space Agencies, in coordination with CEOS-WGCV, should develop advanced methods for cloud mask validation using ground-based independent measurements, working towards the set-up of an autonomous globally representative network for both medium and coarse resolution sensors.
IDN-09	There are inconsistencies among the publicly available manually labelled databases for S2 and L8 pixel classification, owing to the bias introduced by the photo-interpretation process. These inconsistencies are currently limiting their synergistic use for cloud mask validation.	Space Agencies, in coordination with CEOS-WGCV, should develop a reference of the manually labelled database specifically tailored to cloud mask validation. The database should include optically thick, semi-transparent and cloud shadow classes.
IDN-10	There is a lack of unanimously agreed definition of cloud in the optical land community, notably for optically thin or semi-transparent clouds. Furthermore, this definition is generally application-dependent or related to the sensitivity of the downstream algorithm to residual cloudy-contaminated pixels.	Space Agencies, in coordination with CEOS-WGCV, should agree on a common and physically based definition of clouds, for easing inter-comparison of different cloud mask algorithms for optical sensors.
IDN-11	The presence of undetected clouds or adjacency effects at the cloud edges can strongly impact accuracy of downstream products. On the other hand, there have been few theoretical studies to understand this problem and to quantitatively estimate the uncertainty associated to cloud screening in higher level land products.	Space Agencies, in coordination with the science community, should foster theoretical advances in order to rigorously estimate the uncertainty associated to the cloud mask, both for partially cloudy pixels and for clear pixels nearby a cloud.
IDN-12	Terrestrial vegetated surfaces are generally highly anisotropic and heterogeneous at the medium resolution pixel scale. The two effects combine when assessing the quality of satellite-based SR products and it is difficult to disentangle the spatial and directional contribution.	Space Agencies, in coordination with the science community, should support theoretical advances, to fully understand the uncertainty associated to the validation of directional surface reflectance over typical land heterogenous vegetated surfaces.

4.3. Level 2 Bio-Geophysical Land Products

A large number of bio-geophysical parameters are derived from the SR products currently provided by ESA optical land imaging sensors (see, in particular, the products systematically generated in the frame of CGLS to monitor the vegetation, the water cycle, the energy budget and the terrestrial cryosphere [8]). A full review of Cal/Val needs for all relevant land bio-geophysical products is out of scope for the present paper, and we instead focus our attention on a subset of these parameters, including key terrestrial ECVs, which are retrieved from ESA optical imaging sensors in the solar spectrum domain, namely:

- Land Surface Albedo is defined as the ratio of the radiant flux reflected from a unit surface area into the whole hemisphere to the incident radiant flux of hemispherical angular extent. Albedo measures include black-sky albedo, defined in the absence of a diffuse component (no atmospheric scattering), white-sky albedo, under isotropic illumination, and blue-sky albedo, under ambient conditions. Long-term historical observations from the Proba-V mission and the current S3 measurements from OLCI and SLSTR are being used within the C3S to derive coarse resolution global land surface albedo products [142].
- Fraction of absorbed photosynthetically active radiation (fAPAR) is defined as the fraction of solar photosynthetically active radiation (PAR) reaching the surface in the 400–700 nm spectral region that is absorbed by vegetation. fAPAR can be considered at a given time (e.g., instantaneous fAPAR) or daily integrated. All vegetation elements contribute to PAR absorption, both green elements, as well as other non-green elements such as trunks, senescent material and flowers. Satellite-based estimates of

fAPAR currently differ in terms of the adopted definition (green, total, instantaneous, integrated), and this lack of a common definition is inducing discrepancies when inter-comparing products from different missions, although the largest source of differences is caused by the different underlying RTMs [143]. OLCI OGVI products (in the future to be renamed as GI-FAPAR), based on the legacy MERIS algorithm [21,22], provide a measure of instantaneous green fAPAR in the plant canopy.

- Leaf Area Index (LAI) is defined as one half of the total green (i.e., photosynthetically active) leaf area per unit horizontal ground surface area. Different definitions of LAI can be considered: the GAI (Green Area Index) has the same definition as LAI but includes only green elements, and the PAI (Plant Area Index) includes all elements, green and non-green. Similar to fAPAR the lack of consistent definition may cause discrepancies when inter-comparing satellite-based LAI products generated using different algorithms. The CGLS systematically provides global estimation of LAI at 1/3 km resolution derived from Proba-V [144] and now from Sentinel-3 observations.
- Terrestrial Chlorophyll Content. Chlorophyll is one of the most important foliar biochemicals, where the content within a vegetation canopy is related positively to both the productivity of the vegetation and the depth and width of the chlorophyll absorption feature in the reflectance spectra [23]. The S3 OTCI product, which is based on the legacy MERIS Terrestrial Chlorophyll Index (MTCI) algorithm, provide information on the chlorophyll content of vegetation (amount of chlorophyll per unit area of ground). This is a combination of information on area of leaves per unit area on ground (LAI) and the chlorophyll concentration of those leaves [23,24].
- Sun-Induced Fluorescence (SIF) is considered the most direct remote sensing signal to track photosynthetic activity and its dynamics [145,146]. The ESA FLEX mission aims to observe SIF in order to monitor the photosynthetic activity of different biomes and their interannual changes related to climate variability [13,14]. There is currently no active network to support systematic and global validation of SIF. ESA has been supporting the development of dedicated ground-based instrument for this purpose, called Fluorescence Box (Flox) [147,148], and is planning to deploy these instruments in various locations in order to properly support FLEX Cal/Val activities.

There is, in principle, a considerable number of reference data available for validating satellite-based estimation of these bio-geophysical variables. This in situ data is typically gathered during field campaigns or operationally collected in the frame of existing permanent sites and distributed as part of regional and continental networks. However, the main challenge is that these geophysical variables are not directly measured in the field, but rather inferred from indirect observations (see, in particular, fAPAR). Furthermore, the procedures for making the field measurements and for sampling the region of interest generally differ, which can cause large discrepancies over heterogeneous vegetated surfaces when comparing the point data to the collocated satellite-based observations. This lack of common Cal/Val best practices is limiting our ability of synergistically exploiting the currently available ground-based data for the global validation of satellite-derived terrestrial ECVs [41]. Furthermore, a wide range of potentially complementary Cal/Val initiatives are on-going or are planned under different umbrellas (ESA, EC, CEOS), but there is a lack of an integrated solution, since synergies across current and future missions and cross-fertilization among existing Cal/Val projects are largely under-exploited, with resulting duplication of efforts and increase of costs.

Within this section, we review the current status and highlight the remaining needs in terms of infrastructure and methodologies for the validation of land bio-geophysical products. The needs identified in the text are listed with derived recommendations in the Table 4 at the end of the section.

4.3.1. CEOS-LPV Best Practices

In order to address the issue of disparity in the currently used Cal/Val practices, the CEOS-LPV sub-group is defining standardised approaches for in situ measurements,

sampling and upscaling procedures. The definition of CEOS-LPV best practices started in 2014 with the LAI protocol [149], and it now includes endorsed protocols for the following variables: Albedo, Land Cover, Above Ground Biomass, Burned Areas, Soil Moisture and Land Surface Temperature [150,151]. The definition of a Cal/Val protocol is a laborious and collective work, involving contributions from a large team of internationally recognised scientific experts. When a protocol is issued and endorsed at CEOS-LPV level, it has to be promoted within the scientific community and the consensus should be expanded outside the CEOS-LPV contributors. This is a long process, and the Space Agencies should contribute to broadening consensus within the science community. A side effect of this lengthy process is that a protocol may become obsolete, notably in case of latest technological or theoretical advances in the relevant domain. This is the case, for instance, of the LAI protocol [149]. Hence, the protocols should be considered living documents that need to be periodically reviewed and updated in case of major advances in the field. To be noted however, that not all required protocols are ready, for instance, there is not yet an agreed protocol for SR, fAPAR and phenology variables [3]. This gap of best practices is currently being tackled with high priority, in particular for SR and fAPAR within CEOS-LPV sub-group.

4.3.2. FRM4VEG Best Practices

The Fiducial Reference Measurements for Vegetation (FRM4VEG) project was initiated by ESA in 2018 [152] with the objective of establishing the protocols required for traceable in-situ measurements of vegetation-related parameters. In this respect, FRM4VEG actively contributes to CEOS-LPV endeavour in defining best practices for Cal/Val in the land domain with a reinforced focus on the need of traceability and uncertainty estimate, following FRM generic principles. The main operational goal of the FRM4VEG project was to support the validation of ESA/Copernicus land products generated from the Sentinel-2, Sentinel-3 and Proba-V missions, namely the SR products and the S3 OTCI and OGVI products.

The first phase of FRM4VEG project was concluded at the end of 2019 with the elaboration of initial best practices for in situ measurements and methodologies for validation of satellite-derived SR, fAPAR and Chlorophyll content products. This first phase also included the preparation and execution of two field campaigns over two vegetated sites in Europe in the Wytham Woods forest site (UK) and the Barrax agricultural site (Spain). These sites are compliant to FRM generic principles (FRM sites) since they are well characterised and the field measurements within these sites are traceable and follow well-established best practices. A detailed uncertainty budget could therefore be reliably estimated for the validation of satellite-based geophysical products over these FRM sites. An example of such approach was demonstrated in a recent paper, focusing on the validation of S2 and Proba-V SR products over the Barrax cropland site [153]. This paper demonstrated that the adoption of rigorous metrological practices for uncertainty characterisation allows for evaluating the statistical consistency of satellite and in situ data and to assess the quality of the stated uncertainties.

The second phase of FRM4VEG started in February 2021 for two years with a planned new field campaign, and with the overall objective of evolving Cal/Val best practices, notably for SR. In the frame of this new phase, a CEOS-endorsed initiative called the Surface Reflectance Intercomparison Exercise (SRIX4VEG) will be carried out during Summer 2022 [154]. The SRIX4VEG will be a field inter-comparison campaign calling for the contribution of different Cal/Val teams around the world with the objective of working towards the definition of community-agreed guidelines for drone-based surface reflectance product validation.

4.3.3. CEOS-LPV Supersites

CEOS-LPV sub-group has identified a new set of sites, so-called supersites, useful for the validation of satellite land products [155]. The CEOS-LPV supersites are super charac-

terised sites (canopy structure and bio-geophysical variables) following well established protocols, useful for the validation of multiple satellite land geophysical variables (at least 3) and for 3-D radiative transfer modelling approaches. The need for measuring multiple variables was specifically driven by the requirement to assess physical inter-consistency within ECVs.

The supersites were selected primarily from well-known and established networks (e.g., ICOS, NEON). All proposed sites were evaluated for their suitability by ranking them first based on the availability of data (active site) and their spatial representativeness. The sites were additionally classified, based on how many key variables could be validated within a given location, whether structural 3D information was available for the site and if atmospheric and other properties were measured. This process resulted in the selection of 55 globally spread supersites, which were endorsed at CEOS-LPV level [155].

4.3.4. Existing Land Ground-Based Networks

There is a large number of existing ground-based networks, which could be potentially used for satellite land products validation. We list here those that are currently active and widely used in the land domain:

- The AERONET network [52] is providing globally-spread, continuous and readily accessible Cal/Val data of aerosol optical, microphysical and radiative properties. Despite the fact that AERONET is focused on atmospheric measurements, it has been largely used in the land community, both as ancillary data as well as for indirect validation of SR products, as proposed within ACIX [118]. The major strength of AERONET is its spatiotemporal and geographical coverage and the consistency across the whole network, allowing to perform unbiased, comprehensive and statistically robust validation of satellite-based EO data at global scale.
- The National Science Foundation's National Ecological Observatory Network (NEON) [156,157] provides open data from 81 field sites located in different ecosystems across the United States (US). Data collection methods are standardised across the sites and include automated measurements, observational field sampling and airborne remote sensing surveys. The NEON data catalogue includes over 175 data products, providing key information about plants, soil, freshwater and the atmosphere. The validation of vegetation variables, i.e., fAPAR, LAI, fraction of vegetation cover (fCover), within GBOV [53] (c.f. Section 4.3.5) mostly relies on NEON data, with clear limitation in terms of global coverage, since NEON observations are only representative of US biomes.
- The Surface Radiation Budget Network (SURFRAD) [158,159] was established with the primary objective to support climate research with accurate, continuous, long-term measurements of the surface radiation budget over US. Independent measures of upwelling and downwelling, solar and infrared radiation are the primary SURFRAD products. Observations from SURFRAD have been used for evaluating satellite-based estimates of surface radiation, as in GBOV [53], and for validating hydrologic, weather prediction and climate models.
- The Integrated Carbon Observation System (ICOS) [160] is a pan-European network aiming at quantifying and understanding the greenhouse gas balance of Europe and neighbouring regions. ICOS infrastructure includes more than 100 measurement stations in twelve European countries. These stations measure greenhouse gas concentrations in the atmosphere and fluxes over the terrestrial and marine ecosystems [161]. ICOS measurements are used within GBOV [53] to validate surface reflectance and albedo products. Inclusion of ICOS data for fAPAR, LAI and fCover, as well as LST, is currently under investigation within GBOV, although some adaptation of the measurement protocols will be required [3].
- FLUXNET [162] is a global network of micrometeorological tower sites, providing information for validating remote sensing products for net primary productivity (NPP), evaporation, and energy absorption. The network includes more than 800 active

and historic flux measurement sites, dispersed across most of the world's climate space and representative biomes [163]. At each FLUXNET tower site, the eddy covariance method is applied to quantify water, carbon and energy fluxes between the biosphere and atmosphere. FLUXNET data are used to complement the GBOV [53] database for validation of surface reflectance and albedo variables.

- The Terrestrial Ecosystem Research Network (TERN) [164] is Australia's infrastructure for long-term ecosystem observations. The TERN database includes a large ensemble of validation data gathered during field campaigns (>1000 sites) [165]. In addition, as part of the TERN network, 10 sites were identified, which were intensively characterised (super sites) and are suitable for multi-instrumental land product validation and algorithm development. TERN data are used in the frame of GBOV [53] to expand geographical coverage in Australia continent for energy budget and soil moisture validation.

Despite the availability of this large ensemble of ground-based networks, the full exploitation of this wealth of data is still hindered by the disparity in the site characteristics, the used measurement practices, the adopted quality control procedures. One of the reasons is that the majority of these networks, except AERONET, were not primarily designed for validation purposes, but rather for ecosystem monitoring at regional and continental scale. There is a clear need for enhancing the interoperability of this ensemble of networks and to standardise the used protocols and quality assessment approaches. The use of the MM concept, as prototyped in the frame of GAIA-CLIM [62], can contribute to the endeavour of harmonising quality information and metadata across these networks and for supporting the network owners' in enhancing their readiness for supporting satellite Cal/Val needs.

4.3.5. Ground-Based Observations for Validation (GBOV)

The GBOV [53] project is currently leveraging the data collected from this ensemble of existing ground based networks to populate a multi-years database of high quality in-situ measurements to validate seven core land products, which are generated in the frame of CGLS [8], namely: Top-of-canopy reflectances, Land surface albedo, fAPAR, LAI, fCover, LST and Soil Moisture. As part of GBOV, an initial effort was undertaken to improve interoperability across these networks by compiling and maintaining a centralised repository of Cal/Val data using a common format definition and quality metadata approach. Furthermore, in order to facilitate the practical use of GBOV database, common upscaling procedures were implemented to provide reference values representative of areas large enough to cover several pixels of typical coarse resolution satellite imagers.

In the frame of GBOV, a detailed review of the state-of-the-art was performed to identify gaps and needs for the validation of the considered geophysical variables. The most obvious drawback at the moment is the geographical data gap over Africa, South America and Asia, currently preventing a consistent validation of CGLS vegetation products at global scale [3]. Within GBOV, a dedicated task is on-going to start filling these data gaps, notably with the set-up of six new Cal/Val sites. On the other hand, it is clear that an international effort will be required in the coming years involving all Space Agencies in coordination with CEOS-LPV sub-group to enhance geographical coverage and representativeness of ground-based measurements for land products validation.

4.3.6. Field Campaigns and Advanced Technological Solutions

In addition to the operational ground-based networks, the validation of satellite-based land products is largely based on ad-hoc field campaigns. While operational networks typically rely on well-consolidated technology and standardised measurement procedures, to ensure robustness and consistency across the sites, new measurement solutions and practices can be tested during ad-hoc field campaigns, for in-depth validation of satellite EO data over specific ecosystems.

One very popular example of such advanced solutions are drones, also called Unmanned Aerial Vehicles (UAV), which are a flexible and cost-effective solution for field

surveys, notably if compared to airborne campaigns. UAV-based measurement systems are being tested in various land (and water) remote sensing studies since they allow for enhanced spatial and multi-angular sampling [166] over the region of interest. UAV-based techniques also permit the surveillance of remote locations difficult to access, such as wetland, coastlines and dense tropical forest. The UAV flight plan, both the altitude and the acquisition paths, can be defined pre-launch and progressively optimised in successive flights to attain the best spatial sampling over the considered ecosystem. UAV-based hyperspectral sensors are nowadays considered extremely valuable to complement in-situ and airborne spectrometry for the validation of satellite SR products and for ecological monitoring, as demonstrated in recent activities over the Canadian Mer Bleue peatland site [167,168]. The relevance of these measurement systems is also demonstrated by the recent SRIX4VEG initiative [154], which opens a call for a field inter-comparison campaign of UAV-based hyperspectral sensors, to contribute towards developing community-agreed best practices for satellite SR product validation.

Similar to UAV-based systems, new low-cost automated devices are becoming a very interesting alternative to enhance temporal sampling over the region of interest. Such sensors are usually based on inexpensive commercial off-the shelf (COTS) technologies, such as miniaturised sensors controlled with small single-board computers, which can be easily deployed and replicated at different locations. The most widely used sensors are the automated digital hemispherical photography (DHP), which allows tracking the phenological cycle of the canopy at very high temporal interval [169,170]. The use of these devices in the land domain has significantly grown in the past decade for complementing the in-situ traditional field measurements.

Another technique receiving an increasing interest in the land community, is active sensing, using LiDAR systems, both terrestrial-based (TLS) or UAV-based systems (UAV-LS). Attaching a LiDAR scanner to a moving platform, such as a terrestrial pointing system or UAV-based platform, allows 3D mapping of large and multi-layered canopy structures, notably over forest sites [171]. The derived point cloud measurements can be used to retrieve land bio-geophysical parameters, such as LAI [172], Leaf Area Density [173] or detailed tree volume for Above-Ground Biomass (AGB) estimation [174]. In this latter study, the combination of UAV-based platform and LiDAR systems, has shown to be extremely effective for overcoming the limitations of TLS time-demanding fieldwork, while providing comparable accuracy.

To conclude, there is an ever-increasing number of new advanced solutions for land validation purposes. However, the best practices are not consolidated yet and the estimated accuracy of such methods, both for Cal/Val and for biophysical products estimate still needs to be properly characterised and benchmarked against traditional methods.

4.3.7. Validation Tools and Database

The aim of an operational validation system is to ensure that satellite-based EO products can be validated and inter-compared using standardised methodologies, following community-endorsed best practices against a set of globally representative FRM. This framework was first prototyped within the CEOS-LPV sub-group, resulting in the On Line Validation Exercise (OLIVE) project [175]. The ambition of the OLIVE tool was to provide a centralised platform, allowing any user to perform a full validation exercise on a new product, deriving common quality assessment metrics and plots. OLIVE quality assessment was focused on coarse resolution land products, mainly fAPAR and LAI. The validation was based on inter-comparison with reference satellite datasets over BELMANIP-2 (BEnchmark Land Multisite ANalysis and Intercomparison of Products) globally distributed sites [176] and direct validation against the DIRECT ground-based dataset [177]. The OLIVE tool is currently not maintained and needs to be fully redesigned with more advanced open-source language and state-of-the-art code architecture [150]. Furthermore, the BELMANIP-2 list of sites needs to be updated, using more recent land cover maps and optimising the sites to be better representative of S2's resolution [3].

A similar concept of a web-based validation platform was recently developed in the frame of CEOS-LPV, focusing on land surface albedo products [178]. Another successful example of an operational validation tool is the ESA Atmospheric Validation Data Centre (EVDC) [179], which follows the same philosophy of becoming a one-stop-shop, providing centralised open access to satellite sub-settings and spatiotemporal collocated ground-based reference data of atmospheric variables. The development and maintenance of similar tools is considered as the final step for enabling operational validation at global scale of satellite EO data, i.e., to reach stage 4 in CEOS-LPV terms, as recently pointed out in an overview paper [150].

Table 4. Needs and derived recommendation for the validation of land bio-geophysical products.

ID	Status and Needs	Recommendations
IDN-13	There is a clear geographical data gap, over the African, South American and Asian continents in the availability of Cal/Val data for vegetation parameters, notably fAPAR and LAI.	Space Agencies, in coordination with CEOS-LPV, should work towards filling the geographical gap for validation of vegetation parameters, mainly in Africa, Asia and South America.
IDN-14	For some variables, notably for vegetation parameters (FAPAR, LAI), current validation capacity in Europe solely relies on non-European networks, particularly NEON, which is representative of only North America biomes.	ESA should support and promote a European funded network with long-term perspective for validation of satellite-based vegetation parameters. Leveraging of existing facilities (ICOS) should be explored for setting up such network
IDN-15	Existing networks are not primarily designed for validation purposes; as a matter of fact, the lack of uncertainty information and the disparity of used protocols still limit their synergistic use for satellite-products validation.	Space Agencies, in coordination with CEOS-LPV, should foster adoption of community-endorsed Cal/Val best practices across existing networks and support the adaptation of these networks (upgrade protocols, improve uncertainty estimation) for meeting Cal/Val needs.
IDN-16	The heterogeneity in data format, documentation, quality information and metadata availability in existing networks is still limiting their integrated usage for satellite products Cal/Val	Space Agencies, in coordination with CEOS-LPV, should support on-going effort (GBOV) in requesting Cal/Val data providers to meet community agreed standards for quality information, metadata, and documentation.
IDN-17	Assessing maturity of existing networks to support Cal/Val needs is a laborious work. Moreover, this assessment should ideally be reported to networks' owners for helping them to converge to common quality standards, so that to facilitate inter-operability.	Space Agencies, in coordination with CEOS-LPV, should promote adoption of Maturity Matrix concept across existing and future ground-based networks for Cal/Val for enhancing interoperability.
IDN-18	Sustainability long-term still represents a major concern for existing and future Cal/Val networks. The lack of sustained funding is particularly urgent for operation and maintenance.	ESA should reinforce cooperation with national networks and promote within the EC the need for long-term sustainability. ESA to consider this need in the early phases of the mission design.
IDN-19	ESA, in collaboration with EC, is developing a suite of innovative space-borne sensors focusing on land, mainly: FLEX and CHIME. There is a need to enhance preparedness to support Cal/Val needs for these new missions.	ESA, in cooperation with CEOS-LPV, should improve readiness of ground-based infrastructure to support Cal/Val needs of current and future optical land missions. The strategy should exploit synergies across missions, such as using super-sites for measuring multiple variables, and work towards filling the gaps, such as for SIF measurements.
IDN-20	There is currently a lack of an integrated solution addressing Cal/Val needs of current Copernicus Sentinels. Synergies are under exploited with risk of duplicating the efforts and increasing cost in validation activities.	ESA should support on-going activities [180], working towards an integrated solution for sustaining Copernicus Sentinels Cal/Val needs. This includes promoting and exploiting synergies across existing and planned ESA and EC Cal/Val projects.

Table 4. *Cont.*

ID	Status and Needs	Recommendations
IDN-21	Provision of FRM in a sustainable way over a globally representative network of FRM sites is necessary for an operational land Cal/Val system (to reach stage 4 in CEOS-LPV terms) and could be built based on contribution from Agencies and EO programmes.	ESA should continue support FRM4VEG project, working in the transition from research to an operational system of permanent FRM sites and to further expand their geographical coverage in the longer term.
IDN-22	Community-agreed-upon protocols need to be developed for some terrestrial ECVs, notably for SR, fAPAR and phenology.	Space Agencies, in coordination with CEOS-LPV, should contribute in filling this gap in Cal/Val protocols and promote wide adoption by scientific community.
IDN-23	CEOS-LPV protocols are living documents, this entails periodic revision to keep pace with evolution of methodologies and technological solutions, notably UAV-based systems, automated sensors and LiDAR systems.	Space Agencies, in coordination with CEOS-LPV, should work towards revising existing protocols, in particular for LAI, to include recent technological advances, such as, UAV-based, automated sensors, and LiDAR systems.
IDN-24	Advanced technological solutions, such as UAV-based or automated systems, allow for addressing the issue of spatiotemporal upscaling. LiDAR systems allow for detailed characterisation of the site as well as for estimation of vegetation parameters. Yet, the accuracy of these systems needs to be carefully characterised and benchmarked against traditional methods.	Space Agencies, in coordination with CEOS-LPV, should promote inter-comparison exercises and evaluation studies to better characterise the accuracy of advanced technological solutions (UAV-based, low-cost automated sensors, LiDAR) against traditional methods, both for Cal/Val and for vegetation parameters estimation (fAPAR, LAI).
IDN-25	Concerning bio-geophysical retrievals, progresses still need to be made to better resolve the contribution of the different elements of the canopy, e.g., green and senescent parts, to the LAI, PAI and fAPAR estimation. Furthermore, there is often inconsistency in the variable definition (e.g., fAPAR/LAI green vs. total), and this should be carefully taken into account when inter-comparing different products.	Space Agencies, in coordination with CEOS-LPV, should sustain progress in vegetation variables retrieval and modelling approaches (LAI, fAPAR) and promote inter-comparison exercises to advance in understanding the discrepancies in the currently used retrieval algorithms.
IDN-26	The uptake of Cal/Val data from the EO satellite community is strongly limited by the difficulty in discovering, accessing, and using the available measurements, in particular for field campaigns. The adoption of FAIR (Findability, Accessibility, Interoperability, and Reusability) guiding principles [181] and the setting up of a centralised repository will greatly ease uptake of Cal/Val data.	Space Agencies, in coordination with CEOS-LPV, should sustain the effort in setting up a centralised repository of Cal/Val data for Land, following the FAIR guiding principles [181], to collect data acquired within current and future initiatives, such as FRM4VEG, the Joint Experiment for Crop Assessment and Monitoring (JECAM) [182], or FLEX Cal/Val campaigns.
IDN-27	Online validation tools based on community-agreed protocols are needed to allow transparent and standardised validation. The OLIVE tool [175] was a valuable example in this respect, but it is currently not maintained and needs to be updated.	Space Agencies, in coordination with CEOS-LPV, should support the upgrade and secure the maintenance of the OLIVE tool.
IDN-28	BELMANIP-2 collection of sites should be updated/improved. The updates should include using more recent land cover maps and optimising the sampling for S2 resolution.	Space Agencies, in coordination with CEOS-LPV, should support the evolution of BELMANIP-2 sites collection to optimise the sampling for S2 resolution.
IDN-29	A dedicated cloud-based platform with focus on Land Cal/Val would ease the validation process by providing a centralised access to Cal/Val reference data, satellite match-up and analytics to perform quality assessment.	Space Agencies, in coordination with CEOS-LPV, should prepare the ground for a Cal/Val platform for Land, providing centralised access to reference datasets, satellite match-up and tools to assess the uncertainties of the relevant EO products.

5. Overall Strategy and Readiness Level

5.1. Overall Land Cal/Val Strategy

In order to tackle the recommendations identified in the previous chapters (see Tables 2–4) with an integrated approach, ESA aims to implement a generic Cal/Val framework to be replicated and eventually fine-tuned to the specific needs of the geophysical variable under validation. The building blocks of this generic framework are the following (see Figure 2):

- Metrology—The application of metrological best practices is the central pillar underpinning the overall ESA Land Cal/Val strategy, since metrology provides the generic guidelines and community-agreed standards to achieve full traceability of the measurements and detailed characterisation of the uncertainty budget. The metrological practices shall be used as underlying principles for developing all other elements of the Cal/Val strategy.
- RTM and Inter-comparison—The use of RTM simulations allows advancing our understanding of the uncertainty budget associated to the validation, since instrument-related error sources can be fully characterised and decoupled from modelling and retrieval errors. Inter-comparison exercises are key to understand the discrepancies between the various algorithms used within the community, to address the current limitations and, in the longer-term, to converge to a community-agreed solution. RTM simulations are often essential to support algorithm inter-comparison, notably for building reference datasets, as in the case of ACIX [118].
- FRM and supersites—In order to properly assess the accuracy of satellite EO products, we need to have an initial set of "*golden*" sites that represent our best knowledge of the geophysical parameter(s) under validation. These sites are super-characterised and the in situ measurements are carried out using community-endorsed best practices. Both the FRM and CEOS-LPV supersites share these basic underlying principles, although they are addressed with different, yet complementary viewpoints. While the CEOS-LPV supersites concept stems from the need to assess inter-variables' physical consistency, the FRM concept builds upon the fundamental requirement of adopting metrological practices in EO, ensuring traceability of the measurements and detailed uncertainty budget characterisation.
- Protocols—The availability of standard protocols is an essential element in establishling a common approach to the validation problem. When protocols are endorsed at CEOS-LPV level and widely accepted within the community, Cal/Val data from different networks and campaigns can be reliably merged, enhancing spatio-temporal sampling.
- Networks—Once the protocols and requirements for setting up a Cal/Val site are consolidated, they shall be replicated in a network of sites to enhance global coverage. The global network shall have sustained funding and ensure continuous operations, as well as possibly providing data in quasi-real time to support services, such as for GBOV [53], and guarantee that data are generated in a consistent manner across the sites with common approach for calibration and measurements and with harmonised data content, format definition and quality information.
- Ad-hoc campaigns—While networks are the foremost step for moving towards an operational Cal/Val system, the scientific and technological solutions will continue in parallel to evolve, for example, UAV-based and low-cost automated sensors. It is therefore critical to continue supporting ad-hoc Cal/Val campaigns for in-depth validation and for keeping pace with the evolutions in the science and technology domains. The outcomes of such campaigns are advanced protocols and devices, which can be injected into the overall Cal/Val cycle, quality certified following metrological practices and inter-comparison exercises. As these solutions will enhance their maturity level, they could be potentially considered for complementing and expanding the overall Cal/Val solution.
- Database and tools—The final stage of an operational Cal/Val system consists of providing the users with a consistent and fully automated tool, including satellite

sub-settings, and embedded analytics for satellite products comparison with reference data using standardised procedures. This is the framework originally developed by CEOS-LPV with the OLIVE tool [175], which is now being replicated to other terrestrial ECVs [178]. A centralised database with associated tools is key to enhance uptake of Cal/Val data within the community and for fostering the adoption of community-agreed standardised procedures.

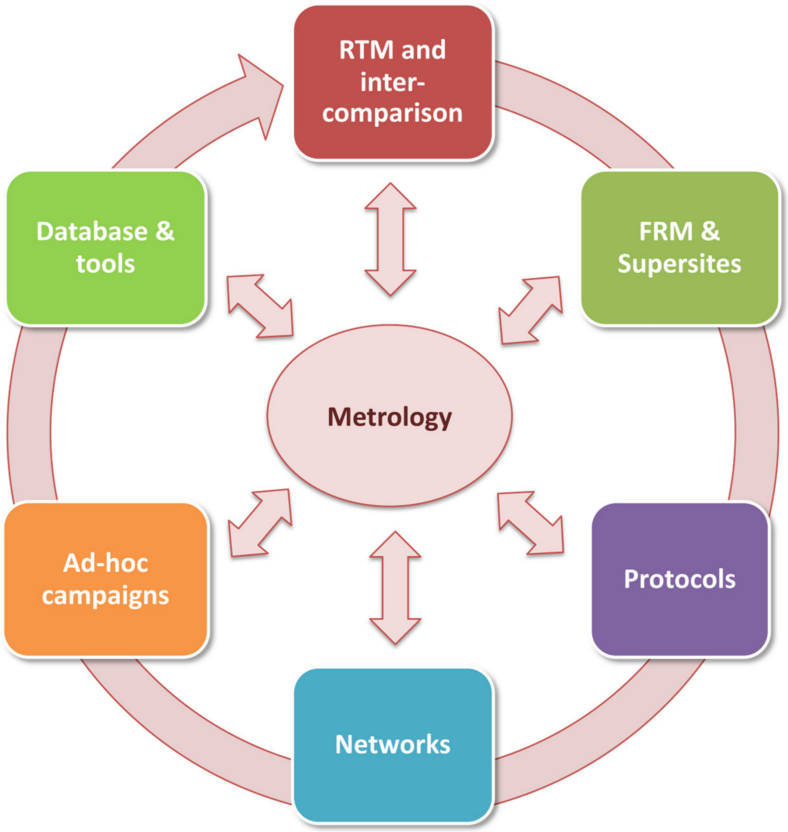


Figure 2. The figure presents the overall ESA generic framework for land Cal/Val, which is replicated and fine-tuned for the considered geophysical variable under validation.

5.2. Readiness Level

The application of the above-mentioned generic Cal/Val approach to ESA operational land products at different processing stages is here discussed and the readiness level for each element of the generic framework is assessed. The readiness levels, ranked with a score (from 0 to 5) are additionally reported in Tables 5–7 and illustrated in Figure 3. The objective of this review is to identify gaps in terms of infrastructure and methods. By addressing these gaps and increasing the readiness level of the various Cal/Val elements, we will enhance our capability to properly assess the accuracy of the relevant satellite EO products and verify whether the mission has effectively met the science objectives defined in the MRD. This is ultimately the main goal of the ESA Cal/Val strategy.

Overall, we see that we already reached a good level of readiness for Level 1 products, owing to the efforts spent in the past decades in developing and consolidating the protocols for calibration and intercalibration, as well as the availability of a ground-based network of SI-traceable measurements, such as RadCalNet [107]. New technological solutions are also being successfully tested, which can further complement the current Cal/Val facilities (see, in particular, the SPecular Array for Radiometric Calibration (SPARC) system [183]).

Conversely to Level 1 TOA products, significant gaps remain in the ability to assess radiometric quality at BOA level. As a matter of fact, our current capability in validating SR products is the weakest point in terms of readiness level over the whole processing

chain (see Figure 3). This is a major concern for reaching interoperability of current and future optical land imaging sensors and should be addressed with the highest priority in the coming years. The major gaps in this respect are the lack of a ground-based network of surface reflectance measurements and of community agreed protocols for SR validation. Significant discrepancies also remain in terms of RTM accuracies in the simulation of radiative interactions within the coupled surface–atmosphere system. Finally, crucial knowledge and data gaps were identified for cloud mask validation, which is an essential preprocessing step for the generation of SR products. ESA is currently tackling these gaps with high priority in the frame of the ACIX-CMIX [118], RAMI4ATM [125] and HYPERNETS [136] initiatives. However, much effort will be required in coordination with the other Space Agencies and the CEOS-LPV to fulfil these gaps with a sustained long-term solution.

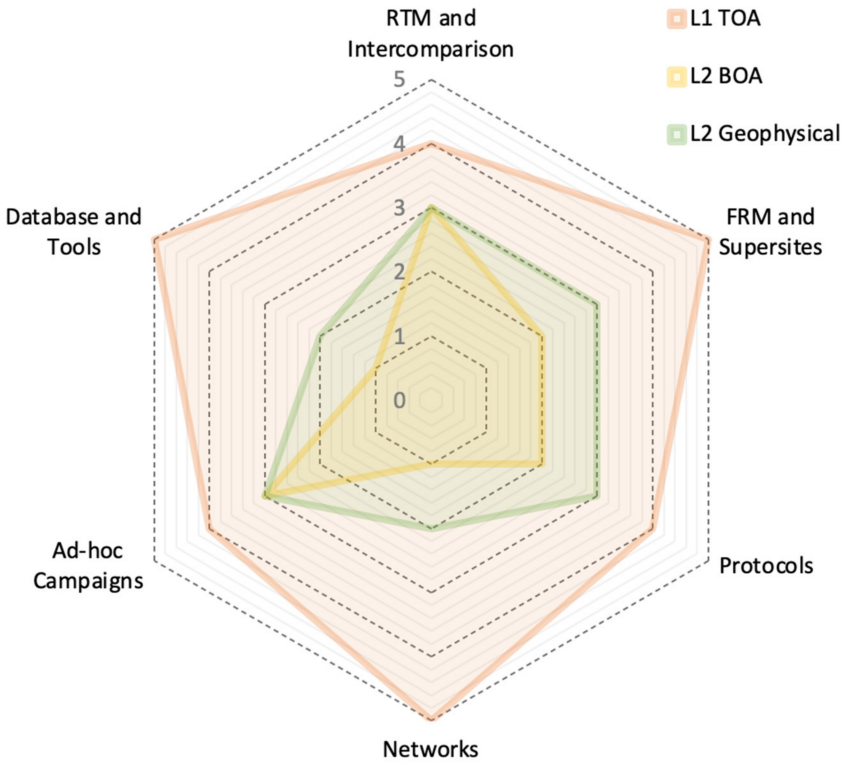


Figure 3. The figure presents the summary readiness level for L1 TOA, L2 BOA and L2 geophysical products.

When we consider Level 2 bio-geophysical products, some important needs were identified by the science community [3]. The CEOS-LPV sub-group was essential in addressing several of these gaps, notably with the release of new protocols, the definition of the super-sites concept and the work towards standardised validation frameworks. The ESA FRM4VEG project [152] further contributed to developing the advanced practices and upscaling procedures for SR, fAPAR and chlorophyll content validation. The GBOV initiative [53] additionally addressed the need of interoperability across networks and started filling the gaps in geographical coverage, particularly for vegetation parameters. Despite these on-going efforts, several challenges remain, owing to the disparity of the used practices, the lack of community-agreed protocols for some variables, the scarcity of Cal/Val data over some geographical areas, and the difficulty in discovering and accessing Cal/Val data. ESA in collaboration with the other Space Agencies and in coordination with CEOS-LPV sub-group, will contribute to filling these gaps with the aim of enhancing preparedness of infrastructure and methods for sustaining Cal/Val needs of current and future ESA/Copernicus land-focused optical missions.

Table 5. Readiness level for Level 1 TOA data with references to needs identified in Table 2.

Element	Readiness	Comment
RTM and Intercomparison	4/5	Good understanding of radiometric uncertainties, though progress should be made to embed metrological practices in the mission design phase [IDN-01], to enhance discoverability of pre-launch characterisation data [IDN-02] and to improve characterisation of geometric mismatch errors [IDN-03].
FRM and supersites	5/5	RadCalNet [107] sites are compliant to FRM guidelines since SI-traceability and metrological practices are fully embedded in the design. The TRUTHS mission [114] will allow reaching full SI-traceability in space.
Protocols	4/5	Community-agreed protocols have been developed in the past decades and consolidated in the frame of CEOS-IVOS and GSICS. Yet, there is a need to provide uncertainty per-pixel at Level 1 [IDN-04] and to enhance the accuracy of RTMs [IDN-05] and of the Lunar irradiance models [IDN-06].
Networks	5/5	The RadCalNet network [107] is fully operational and will be expanded with additional new sites to increase geographical coverage. New technological solutions are also being successfully tested [183].
Ad-hoc campaigns	4/5	Several ad-hoc campaigns and inter-comparison exercises were run in the past and there is a very good understanding of the required methodologies.
Database and Tools	5/5	Various databases are available, on ESA side, the DIMITRI tool, operationally used for S2 and S3 radiometric assessment [74].

Table 6. Readiness level for L2 BOA data with references to needs identified in Table 3.

Element	Readiness	Comments
RTM and Intercomparison	3/5	Significant progresses were made thanks to ACIX [118], although there are still gaps in fully understanding the discrepancies between the AC algorithms and in characterising SR products' uncertainties [IDN-12]. These challenges will be addressed within RAMI4ATM [125]. Similar knowledge and data gaps were identified for cloud mask validation [IDN-09, IDN-10, IDN-11].
FRM and supersites	3/5	An initial set of super-sites has been defined as part of FRM4VEG following metrological practices. The SRIX4VEG [154] will build on that by consolidating best-practices with a focus on the use of UAV-based hyperspectral measurements. The readiness level for this element is therefore expected to rise in the next 2 years as a result of these activities.
Protocols	2/5	There are no community-endorsed protocols for SR validation, this is currently on-going work in the frame of FRM4VEG [152] and SRIX4VEG [154]. The readiness level is therefore expected to rise in the next 2 years.
Networks	1/5	No network is currently available for validating SR [IDN-07]. HYPERNETS project [136] aims to fulfil this need. Likewise, there is a lack of a global network for cloud mask validation [IDN-08].
Ad-hoc campaigns	3/5	Several campaigns have been carried out in the past years using field spectrometry, airborne, and recently UAV-based spectrometry. The use of UAV platforms is today considered a very flexible and cost-effective solution.
Database and Tools	1/5	No database and tools are currently available for SR validation.

Table 7. Readiness level for L2 bio-geophysical data with references to needs identified in Table 4.

Element	Readiness	Comments
RTM and Intercomparison	3/5	Inter-comparisons of retrieval algorithms are recommended for quantifying cross-mission biases as well as for better assessing the discrepancies due to inconsistent variable definition (e.g., green Vs. total fAPAR) [IDN-25].
FRM and supersites	3/5	FRM4VEG and CEOS-LPV supersites are the essential first step. Yet, there is a need to move towards sustained operations of permanent FRM sites [IDN-21]. In the Copernicus domain, there is a lack of an integrated solution to support Cal/Val needs of current and future missions [IDN-19, IDN-20].
Protocols	3/5	Community-endorsed protocols have recently been published by CEOS-LPV for key ECVs, e.g., LST, Albedo, Soil moisture. However, there is a lack of protocols for SR, fAPAR and phenology [IDN-22] and a need to update and maintain protocols to keep pace with science and technological evolutions [IDN-23]. This is currently on-going work within CEOS-LPV.
Networks	2/5	A large number of regional and continental networks are currently available. However, the majority of them were not primarily designed for Cal/Val purposes. There is a need to embed Cal/Val practices across these networks and work towards harmonisation [IDN-15, IDN-16, IDN-17]. The MM concept, as in GAIA-CLIM, can facilitate network interoperability. Though, some gaps remain in terms of geographical coverage [IDN-13, IDN-14], and there is a need to ensure sustainability in the long-term [IDN-18].
Ad-hoc campaigns	3/5	There is a deluge of Cal/Val data acquired as part of ad-hoc campaigns and new devices are being tested, in particular UAV-based. Yet, the accuracy of these new methods needs to be properly assessed [IDN-24].
Database and Tools	2/5	There is an urgent need to upgrade and secure maintenance of the OLIVE tool [175] and update the used reference database [IDN-27, IDN-28]. Overall, there is an urgent need to enhance the discoverability and accessibility of Cal/Val data following FAIR guiding principles [IDN-26] and of a dedicated cloud-based platform for land Cal/Val [IDN-29].

6. Conclusions and Outlook

The ESA strategy for Land products validation was presented and discussed within this paper. The overall validation problem is addressed from an end-to-end system perspective, which includes rigorous verification of the satellite system output starting from the calibration at the TOA level up to the retrieval of geophysical parameters. The main pillars underpinning this strategy are a reinforced focus on metrological practices to ensure traceability along the full processing chain and the need of providing uncertainty per-measurand, ideally at pixel-level.

The survey of the state-of-the-art in Cal/Val, which was the outcome of a series of ESA Cal/Val Workshops, allows the identification of concrete needs and data gaps to be addressed in the coming years in the different domains. While for the assessment of TOA radiometry there is a long-lasting experience and a very good level of readiness in terms of infrastructure and practices, important gaps appear at the Level 2 BOA stage.

Among the most urgent gaps, there is an obvious need for setting up and maintaining a global network of surface reflectance ground-based measurements. This is the main priority for ESA, since it is currently the weakest point in terms of readiness level along the whole processing chain, while it is an essential element to facilitate the interoperability of current and future land imaging sensors.

In terms of ground-based facilities, there is a large number of potential national and continental networks, although they significantly differ with respect to the used practices and there is little attention in providing uncertainty in Cal/Val data. The objective for the coming years will be to contribute to harmonising best practices across existing networks to enhance their interoperability moving, in the longer term, to a sustained operational global network of FRM sites.

The general approach to Cal/Val was finally presented, which is based on a set of building blocks. The application of this framework and the detailed recommendations gathered during past Cal/Val Workshops will allow to shape ESA Cal/Val strategy for land in the coming years, addressing the gaps with an integrated solution. The longer-term objective is to pave the way for enhanced interoperability among current and future ESA/Copernicus land imaging sensors for the benefit of land applications.

Author Contributions: The concept, methodology and design originate from F.N. and P.G. The writing of the paper was led by F.N. All authors contributed in the writing and revision, specifically for their relevant area of expertise and responsibility. P.G. is the Head of the ESA Sensor Performance, Products and Algorithm Section (SPPA), V.B. and S.D. are the ESA Data Quality and Cal/Val managers respectively of the S2 and S3 missions, F.G. is the ESA S2 Mission Manager, F.N. is leading the IDEAS-QA4EO Cal/Val framework and oversees the Proba-V Data Quality and Cal/Val activities, S.S. and B.T. support the Cal/Val activities of S2 and S3 missions, J.A. contributes to the CHIME mission science and Cal/Val requirements definition, G.D. coordinates the ACIX and CMIX activities. All authors have read and agreed on the published version of the manuscript.

Funding: The writing and editing of this overview paper was carried out as part of the ESA-funded Quality Assurance for Earth Observation (IDEAS-QA4EO) framework contract. The various projects mentioned within the manuscript received funding mainly from ESA and the EC under different frameworks; we refer to the relevant project's web sites and reference papers for detailed information about each project.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Acknowledgments: The authors greatly acknowledge the three anonymous reviewers for the insightful comments, which were instrumental in improving the quality of the manuscript. The authors wish to thank the participants of the "ESA Workshop on Land Validation Strategy (30 Novemver–1 December 2020)" for their feedback on the current status and needs in Land Cal/Val. This feedback was essential to derive the recommendations reported in this paper and for shaping the ESA Cal/Val Strategy for Land. In addition, Fernando Camacho and Jadu Dash, are acknowledged for supporting the submission of this manuscript to their MDPI Special Issue on "Recent Advances in Satellite Derived Global Land Product Validation".

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Quality Assurance Framework for Earth Observation (QA4EO) Web Portal. Available online: http://www.qa4eo.org/ (accessed on 31 May 2021).
2. Global Earth Observation System of Systems (GEOSS) Web Portal. Available online: https://www.earthobservations.org/geoss.php (accessed on 31 May 2021).
3. ESA GMQ Session, "ESA Workshop on Land Validation Strategy (30/11–1/12/2020): Highlights and Recommendations"; Report from Land Cal/Val WS 2020. Available online: https://earth.esa.int/eogateway/documents/20142/37627/LPV2020_Report_v1.3.pdf (accessed on 31 May 2021).
4. European Copernicus Programme Web Portal. Available online: https://www.copernicus.eu/en (accessed on 31 May 2021).
5. Drusch, M.; Del Bello, U.; Carlier, S.; Colin, O.; Fernandez, V.; Gascon, F.; Hoersch, B.; Isola, C.; Laberinti, P.; Martimort, P.; et al. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sens. Environ.* **2012**, *120*, 25–36. [CrossRef]
6. Donlon, C.; Berruti, B.; Buongiorno, A.; Ferreira, M.H.; Féménias, P.; Frerick, J.; Goryl, P.; Klein, U.; Laur, H.; Mavrocordatos, C.; et al. The global monitoring for environment and security (GMES) sentinel-3 mission. *Remote Sens. Environ.* **2012**, *120*, 37–57. [CrossRef]
7. Dierckx, W.; Sterckx, S.; Benhadj, I.; Livens, S.; Duhoux, G.; Van Achteren, T.; Francois, M.; Mellab, K.; Saint, G. PROBA-V mission for global vegetation monitoring: Standard products and image quality. *Int. J. Remote Sens.* **2014**, *35*, 2589–2614. [CrossRef]
8. Copernicus Global Land Service (CGLS) Web Portal. Available online: https://land.copernicus.eu/global/index.html (accessed on 31 May 2021).
9. Copernicus Climate Change Service (C3S) Web Portal. Available online: https://climate.copernicus.eu/ (accessed on 31 May 2021).
10. Niro, F. Evaluation of orbital drift effect on Proba-V surface reflectances time series. *Remote Sens.* **2021**, *13*, 2250. [CrossRef]

11. ESA Web Page. Available online: https://www.esa.int/Enabling_Support/Space_Engineering_Technology/Proba_Missions/Ending_global_plant_tracking_Proba-V_assigned_new_focus (accessed on 31 May 2021).
12. Saunier, S.; Louis, J.; Debaecker, V.; Beaton, T.; Cadau, E.G.; Boccia, V.; Gascon, F. Sen2like, a tool to generate sentinel-2 harmonised surface reflectance products-first results with Landsat-8. In Proceedings of the IGARSS 2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 28 July 28–2 August 2019; pp. 5650–5653.
13. Moreno, J.F.; Goulas, Y.; Huth, A.; Middleton, E.; Miglietta, F.; Mohammed, G.; Nedbal, L.; Rascher, U.; Verhoef, W.; Drusch, M. Very high spectral resolution imaging spectroscopy: The Fluorescence Explorer (FLEX) mission. In Proceedings of the 2016 IEEE International Geoscience and Remote Sensing Symposium, Beijing, China, 10–15 July 2016; pp. 264–267.
14. Drusch, M.; Moreno, J.; Del Bello, U.; Franco, R.; Goulas, Y.; Huth, A.; Kraft, S.; Middleton, E.M.; Miglietta, F.; Mohammed, G.; et al. The fluorescence explorer mission concept—ESA's earth explorer 8. *IEEE Trans. Geosci. Remote Sens.* **2016**, *55*, 1273–1284. [CrossRef]
15. Nieke, J.; Rast, M. Towards the copernicus hyperspectral imaging mission for the environment (CHIME). In Proceedings of the IGARSS 2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018; pp. 157–159.
16. Koetz, B.; Bastiaanssen, W.; Berger, M.; Defourney, P.; Del Bello, U.; Drusch, M.; Drinkwater, M.; Duca, R.; Fernandez, V.; Ghent, D.; et al. High spatio-temporal resolution land surface temperature mission—A Copernicus candidate mission in support of agricultural monitoring. In Proceedings of the IGARSS 2018–2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018; pp. 8160–8162.
17. Sentinel-2 Product Specifications Document. Available online: https://sentinel.esa.int/documents/247904/685211/Sentinel-2-Products-Specification-Document (accessed on 31 May 2021).
18. Sentinel-2 MSI-Level-2A Algorithm Theoretical Basis Document (ATBD). Available online: https://earth.esa.int/c/document_library/get_file?folderId=349490&name=DLFE-4518.pdf (accessed on 31 May 2021).
19. Sentinel-3 OLCI Level 1 Products Specifications. Available online: https://sentinels.copernicus.eu/web/sentinel/user-guides/sentinel-3-olci/processing-levels/level-1 (accessed on 31 May 2021).
20. Sentinel-3 SLSTR Level 1 Products Specifications. Available online: https://sentinels.copernicus.eu/web/sentinel/user-guides/Sentinel-3-slstr/product-types/level-1b (accessed on 31 May 2021).
21. Gobron, N.; Pinty, B.; Mélin, F.; Taberner, M.; Verstraete, M.M.; Robustelli, M.; Widlowski, J.L. Evaluation of the MERIS/ENVISAT FAPAR product. *Adv. Space Res.* **2007**, *39*, 105–115. [CrossRef]
22. Gobron, N.; Pinty, B.; Aussedat, O.; Taberner, M.; Faber, O.; Mélin, F.; Lavergne, T.; Robustelli, M.; Snoeij, P. Uncertainty estimates for the FAPAR operational products derived from MERIS—Impact of top-of-atmosphere radiance uncertainties and validation with field data. *Remote Sens. Environ.* **2008**, *112*, 1871–1883. [CrossRef]
23. Dash, J.; Curran, J. The MERIS Terrestrial Chlorophyll Index. *Int. J. Remote Sens.* **2004**, *25*, 5403–5413. [CrossRef]
24. Pastor-Guzman, J.; Brown, L.; Morris, H.; Bourg, L.; Goryl, P.; Dransfeld, S.; Dash, J. The Sentinel-3 OLCI Terrestrial Chlorophyll Index (OTCI): Algorithm Improvements, Spatiotemporal Consistency and Continuity with the MERIS Archive. *Remote Sens.* **2020**, *12*, 2652. [CrossRef]
25. Sentinel-3 SYN Algorithm Theoretical Basis Document (ATBD). Available online: https://earth.esa.int/documents/247904/349589/SYN_L2-3_ATBD.pdf (accessed on 31 May 2021).
26. Wolters, E.; Dierckx, W.; Iordache, M.D.; Swinnen, E. PROBA-V Products User Manual. VITO, 2014. Available online: http://proba-v.vgt.vito.be/sites/default/files/Product_User_Manual.pdf (accessed on 31 May 2021).
27. Schaepman-Strub, G.; Schaepman, M.E.; Painter, T.H.; Dangel, S.; Martonchik, J.V. Reflectance quantities in optical remote sensing—Definitions and case studies. *Remote Sens. Environ.* **2006**, *103*, 27–42. [CrossRef]
28. GSICS Web Site. Available online: https://gsics.wmo.int (accessed on 31 May 2021).
29. CEOS Infrared and Visible Optical Sensors Subgroup (CEOS-IVOS) Web Site. Available online: http://ceos.org/ourwork/workinggroups/wgcv/subgroups/ivos/ (accessed on 31 May 2021).
30. CEOS Land Surface Imaging Virtual Constellation (LSI-VC) Web Site Site. Available online: http://ceos.org/ourwork/virtual-constellations/lsi/ (accessed on 31 May 2021).
31. Lewis, A.; Lacey, J.; Mecklenburg, S.; Ross, J.; Siqueira, A.; Killough, B.; Szantoi, Z.; Tadono, T.; Rosenavist, A.; Goryl, P.; et al. CEOS analysis ready data for Land (CARD4L) overview. In Proceedings of the IGARSS 2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018; pp. 7407–7410.
32. CEOS Land Products Validation (LPV) Group Web Site. Available online: https://lpvs.gsfc.nasa.gov/ (accessed on 31 May 2021).
33. Land Product Validation Workshop 2018 Web Site. Available online: https://earth.esa.int/eogateway/events/lpve-2018-land-product-validation-and-evolution-workshop?text=LPV+2018 (accessed on 31 May 2021).
34. ESA Sentinel-2 Validation Team Meeting 4th. Available online: https://nikal.eventsair.com/NikalWebsitePortal/4th-sentinel-2-validation-team-meeting/esa (accessed on 31 May 2021).
35. ESA-EUMETSAT Sentinel-3 Validation Team Web Site. Available online: https://www.s3vt.org/QuickEventWebsitePortal/sentinel-3-validation-team-s3vt/esa (accessed on 31 May 2021).
36. CEOS Terms and Definition Web Page. Available online: http://calvalportal.ceos.org/t-d_wiki (accessed on 31 May 2021).
37. Justice, C.; Belward, A.; Morisette, J.; Lewis, P.; Privette, J.; Baret, F. Developments in the 'validation' of satellite sensor products for the study of the land surface. *Int. J. Remote Sens.* **2000**, *21*, 3383–3390. [CrossRef]

38. Joint Committee for Guides in Metrology (JCGM). International Vocabulary of Metrology—Basic and General Concepts and Associated Terms, BIPM, Sèvres, France. Available online: https://www.bipm.org/utils/common/documents/jcgm/JCGM_200_2012.pdf (accessed on 31 May 2021).
39. Bojinski, S.; Verstraete, M.; Peterson, T.C.; Richter, C.; Simmons, A.; Zemp, M. The concept of essential climate variables in support of climate research, applications, and policy. *Bull. Am. Meteorol. Soc.* **2014**, *95*, 1431–1443. [CrossRef]
40. Nightingale, J.; Boersma, K.F.; Muller, J.; Compernolle, S.; Lambert, J.C.; Blessing, S.; Giering, R.; Gobron, N.; De Smedt, I.; Coheur, P.; et al. Quality assurance framework development based on six new ECV data products to enhance user confidence for climate applications. *Remote Sens.* **2018**, *10*, 1254. [CrossRef]
41. Nightingale, J.; Mittaz, J.; Douglas, S.; Dee, D.; Ryder, J.; Taylor, M.; Old, C.; Dieval, C.; Fouron, C.; Duveau, G.; et al. Ten priority science gaps in assessing climate data record quality. *Remote Sens.* **2019**, *11*, 986. [CrossRef]
42. JCGM. Evaluation of Measurement Data—Guide to the Expression of Uncertainty in Measurement JCGM 100:2008, GUM 1995 with Minor Corrections, First Edition September 2008. Available online: http://www.bipm.org/utils/common/documents/jcgm/JCGM_100_2008_E.pdf (accessed on 31 May 2021).
43. Gorroño, J.; Fomferra, N.; Peters, M.; Gascon, F.; Underwood, C.I.; Fox, N.; Kirches, G.; Brockmann, C. A radiometric uncertainty tool for the Sentinel 2 mission. *Remote Sens.* **2017**, *9*, 178. [CrossRef]
44. Gorroño, J.; Hunt, S.; Scanlon, T.; Banks, A.; Fox, N.; Woolliams, E.; Underwood, C.; Gascon, F.; Peters, M.; Fomferra, N.; et al. Providing uncertainty estimates of the Sentinel-2 top-of-atmosphere measurements for radiometric validation activities. *Eur. J. Remote Sens.* **2018**, *51*, 650–666. [CrossRef]
45. S2 RUT GitHub Tool Repository. Available online: https://github.com/senbox-org/snap-rut (accessed on 31 May 2021).
46. Smith, D.; Hunt, S.E.; Etxaluze, M.; Peters, D.; Nightingale, T.; Mittaz, J.; Woolliams, E.R.; Polehampton, E. Traceability of the Sentinel-3 SLSTR Level-1 Infrared Radiometric Processing. *Remote Sens.* **2021**, *13*, 374. [CrossRef]
47. Immler, F.J.; Dykema, J.; Gardiner, T.; Whiteman, D.N.; Thorne, W.; Vömel, H. Reference quality upper-air measurements: Guidance for developing GRUAN data products. *Atmos. Meas. Tech.* **2010**, *3*, 1217–1231. [CrossRef]
48. Loew, A.; Bell, W.; Brocca, L.; Bulgin, C.E.; Burdanowitz, J.; Calbet, X.; Donner, R.V.; Ghent, D.; Gruber, A.; Kaminski, T.; et al. Validation practices for satellite-based Earth observation data across communities. *Rev. Geophys.* **2017**, *55*, 779–817. [CrossRef]
49. Vermote, E.F.; Kotchenova, S. Atmospheric correction for the monitoring of land surfaces. *J. Geophys. Res. Atmos.* **2008**, *113*, D23S90. [CrossRef]
50. Nappo, C.J.; Caneill, J.Y.; Furman, R.W.; Gifford, F.A.; Kaimal, J.C.; Kramer, M.L.; Lockhart, T.J.; Pendergast, M.M.; Pielke, R.A.; Randerson, D.; et al. The Workshop on the Representativeness of Meteorological-Observations, June 1981, Boulder, Colorado. *Bull. Am. Meteorol. Soc.* **1982**, *63*, 761–764.
51. Baret, F.; Weiss, M.; Allard, D.; Garrigue, S.; Leroy, M.; Jeanjean, H.; Fernandes, R.; Myneni, R.; Privette, J.; Morisette, J. VALERI: A Network of Sites and a Methodology for the Validation of Medium Spatial Resolution Land Satellite Products. *Remote Sens. Environ.* **2005**, *76*, 36–39.
52. Holben, B.N.; Eck, T.F.; Slutsker, I.A.; Tanre, D.; Buis, J.; Setzer, A.; Vermote, E.; Reagan, J.A.; Kaufman, Y.J.; Nakajima, T.; et al. AERONET—A federated instrument network and data archive for aerosol characterization. *Remote Sens. Environ.* **1998**, *66*, 1–16. [CrossRef]
53. Ground-Based Observations for Validation (GBOV) web site. Available online: https://gbov.acri.fr/ (accessed on 31 May 2021).
54. Hollmann, R.; Merchan, C.J.; Saunders, R.; Downy, C.; Buchwitz, M.; Cazenave, A.; Chuvieco, E.; Defourny, P.; de Leeuw, G.; Forsberg, R.; et al. The ESA climate change initiative: Satellite data records for essential climate variables. *Bull. Amer. Meteor. Soc.* **2013**, *94*, 1541–1552. [CrossRef]
55. Merchant, C.J.; Paul, F.; Popp, T.; Ablain, M.; Bontemps, S.; Defourny, P.; Hollmann, R.; Lavergne, T.; Laeng, A.; Leeuw, G.D.; et al. Uncertainty information in climate data records from Earth observation. *Earth Syst. Sci. Data* **2017**, *9*, 511–527. [CrossRef]
56. Donlon, C.J.; Wimmer, W.; Robinson, I.; Fisher, G.; Ferlet, M.; Nightingale, T.; Bras, B. A second-generation blackbody system for the calibration and verification of seagoing infrared radiometers. *J. Atmos. Ocean. Technol.* **2014**, *31*, 1104–1127. [CrossRef]
57. Theocharous, E.; Usadi, E.; Fox, N. CEOS comparison of IR brightness temperature measurements in support of satellite validation. In *Part I: Laboratory and Ocean Surface Temperature Comparison of Radiation Thermometers*; National Physical Laboratory: Teddington, UK, 2010; p. 130.
58. Fox, N.; Greening, M.C. A Guide to Comparisons-Organisation, Operation and Analysis to Establish Measurement Equivalence to Underpin the Quality Assurance Requirements of Geo, Versio-4, QA4EO-QAEO-GEN-DQK-004. 2010. Available online: http://qa4eo.org/docs/QA4EO-QAEO-GEN-DQK-004_v4.0.pdf (accessed on 31 May 2021).
59. Nightingale, J.; Schaepman-Strub, G.; Nickeson, J.; Baret, F.; Herold, M. Assessing Satellite-Derived Land Product Quality for Earth System Science Applications: Results from the CEOS-LPV Sub-Group. In Proceedings of the International Symposium on Remote Sensing of Environment, Sydney, Australia, 10–15 April 2011.
60. CCI Holtzer-Popp Maturity Matrix. Available online: http://cci.esa.int/sites/default/files/content/docs/Holzer-Popp_maturity_matrix_022014.pdf (accessed on 31 May 2021).
61. ESA Climate Change Initiative Web Site. Available online: https://climate.esa.int/en/ (accessed on 31 May 2021).
62. GAIA-CLIM Maturity Matrix. Available online: http://www.gaia-clim.eu/page/maturity-matrix-assessment (accessed on 31 May 2021).

63. ESA Earthnet Data Assessment Pilot (EDAP) Project's Web-Site. Available online: https://earth.esa.int/eogateway/activities/edap (accessed on 31 May 2021).
64. ESA VH-RODA Workshop 2021. Available online: https://earth.esa.int/eogateway/events/vh-roda-workshop-2021 (accessed on 31 May 2021).
65. CEOS Analysis Ready Data Web Site. Available online: http://ceos.org/ard/ (accessed on 31 May 2021).
66. USGS Landsat Collection 2 Web Site. Available online: https://www.usgs.gov/core-science-systems/nli/landsat/landsat-collection-2?qt-science_support_page_related_con=1#qt-science_support_page_related_con (accessed on 31 May 2021).
67. Ralf, G.; Quast, R.; Mittaz, J.D.; Hunt, S.E.; Harris, M.; Woolliams, E.R.; Merchant, C.J. A novel framework to harmonise satellite data series for climate applications. *Remote Sens.* **2019**, *11*, 1002. [CrossRef]
68. EC Fiduceo Project's Web Site. Available online: https://cordis.europa.eu/project/id/638822 (accessed on 31 May 2021).
69. Claverie, M.; Ju, J.; Masek, J.G.; Dungan, J.L.; Vermote, E.F.; Roger, J.C.; Skakun, S.V.; Justice, C. The Harmonized Landsat and Sentinel-2 surface reflectance data set. *Remote Sens. Environ.* **2018**, *219*, 145–161. [CrossRef]
70. Delwart, S.; Preusker, R.; Bourg, L.; Santer, R.; Ramon, D.; Fischer, J. MERIS in-flight spectral calibration. *Int. J. Remote Sens.* **2007**, *28*, 479–496. [CrossRef]
71. XIONG, X.J.; Angal, A.; Sun, J.; Choi, T.J.; Johnson, E. On-orbit performance of MODIS solar diffuser stability monitor. *J. Appl. Remote Sens.* **2014**, *8*, 083514. [CrossRef]
72. Sterckx, S.; Adriaensen, S.; Dierckx, W.; Bouvet, M. In-orbit radiometric calibration and stability monitoring of the PROBA-V instrument. *Remote Sensing* **2016**, *8*, 546. [CrossRef]
73. Chander, G.; Hewison, T.J.; Fox, N.; Wu, X.; Xiong, X.; Blackwell, W.J. Overview of intercalibration of satellite instruments. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 1056–1080. [CrossRef]
74. Database for Imaging Multi-spectral Instruments and Tools for Radiometric Intercomparison (DIMITRI) Web Portal. Available online: https://dimitri.argans.co.uk/ (accessed on 31 May 2021).
75. Cosnefroy, H.; Leroy, M.; Briottet, X. Selection and characterization of Saharan and Arabian desert sites for the calibration of optical satellite sensors. *Remote Sens. Environ.* **1996**, *58*, 101–114. [CrossRef]
76. CEOS Endorsed PICS Sites List Provided within the CEOS Cal/Val Portal. Available online: http://calvalportal.ceos.org/ca/pics_sites;jsessionid=8199EA6A87B6423AC217D19B9E2F2347 (accessed on 31 May 2021).
77. Govaerts, Y.M. Sand dune ridge alignment effects on surface BRF over the Libya-4 CEOS calibration site. *Sensors* **2015**, *15*, 3453–3470. [CrossRef] [PubMed]
78. Cedric, B.; Briottet, X.; Bréon, F.-M.; Viallefont-Robinet, F.; Bouvet, M. Revisiting pseudo invariant calibration sites (PICS) over sand deserts for vicarious calibration of optical imagers at 20 km and 100 km scales. *Remote Sens.* **2019**, *11*, 1166.
79. PICSCAR Web Portal. Available online: https://picscar.magellium.com/ (accessed on 31 May 2021).
80. Govaerts, Y.M. "Estimating the Accuracy of 1D Radiative Transfer Models over the Libya-4 Site". Available online: https://www.eradiate.eu/site/post/20191210-rtm_accuracy/ (accessed on 31 May 2021).
81. Eradiate Web Site. Available online: http://eradiate.eu/site/ (accessed on 31 May 2021).
82. Doelling, D.R.; Morstad, D.; Scarino, B.R.; Bhatt, R.; Gopalan, A. The characterization of deep convective clouds as an invariant calibration target and as a visible calibration technique. *IEEE Trans. Geosci. Remote Sens.* **2012**, *51*, 1147–1159. [CrossRef]
83. Doelling, D.; Haney, C.; Bhatt, R.; Scarino, B.; Gopalan, A. Geostationary visible imager calibration for the CERES SYN1deg edition 4 product. *Remote Sens.* **2018**, *10*, 288. [CrossRef]
84. Vermote, E.; Kaufman, Y.J. Absolute calibration of AVHRR visible and near infrared channels using ocean and cloud views. *Int. J. Remote Sens.* **1995**, *16*, 2317–2340. [CrossRef]
85. Lamquin, N.; Bruniquel, V.; Gascon, F. Sentinel-2 L1C radiometric validation using deep convective clouds observations. *Eur. J. Remote Sens.* **2018**, *51*, 11–27. [CrossRef]
86. Lamquin, N.; Bourg, L.; Clerc, S.; Donlon, C. OLCI A/B tandem phase analysis, part 3: Post-tandem monitoring of cross-calibration from statistics of Deep Convective Clouds observations. *Remote Sens.* **2020**, *12*, 3105. [CrossRef]
87. Stone, T.C. Radiometric calibration stability and inter-calibration of solar-band instruments in orbit using the moon. In Proceedings of the SPIE Earth Observing Systems XIII, San Diego, CA, USA, 11–13 August 2008.
88. Meygret, A.; Latry, C.; Kelbert, A.; Blanchet, G. PLEIADES high resolution optical sensors radiometric and spatial calibration based on stars. In Earth Observing Systems XXIV. *Int. Soc. Opt. Photonics* **2019**, *11127*, 1112720.
89. Sun, J.; Angal, A.; Xiong, X.; Chen, H.; Geng, X.; Wu, A.; Choi, T.; Chu, M. MODIS reflective solar bands calibration improvements in Collection 6. In Earth Observing Missions and Sensors: Development, Implementation, and Characterization II. *Int. Soc. Opt. Photonics* **2012**, *8528*, 85280N.
90. Morfitt, R.; Barsi, J.; Levy, R.; Markham, B.; Micijevic, E.; Ong, L.; Scaramuzza, P.; Vanderwerff, K. Landsat-8 Operational Land Imager (OLI) radiometric performance on-orbit. *Remote Sens.* **2015**, *7*, 2208–2237. [CrossRef]
91. Wilson, N.; Greenberg, J.; Jumpasut, A.; Collison, A.; Weichelt, H. *Absolute Radiometric Calibration of Planet Dove Satellites, Flocks 2p & 2e*; Planet: San Francisco, CA, USA, 2017; 2p.
92. Stone, T.C.; Kieffer, H.H. Assessment of uncertainty in ROLO lunar irradiance for on-orbit calibration. In Earth Observing Systems IX. *Int. Soc. Opt. Photonics* **2004**, *5542*, 300–310.
93. Stone, T.C.; Kieffer, H.; Lukashin, C.; Turpie, K. The moon as a climate-quality radiometric calibration reference. *Remote Sens.* **2020**, *12*, 1837. [CrossRef]

94. Wagner, S.C.; Hewison, T.; Stone, T.; Lacherade, S.; Fougnie, B.; Xiong, X. A summary of the joint GSICS–CEOS/IVOS lunar calibration workshop: Moving towards intercalibration using the Moon as a transfer target. In Sensors, Systems, and Next-Generation Satellites XIX. *Int. Soc. Opt. Photonics* **2015**, *9639*, 96390Z.
95. EUMETSAT Lunar Model. Available online: https://www.eumetsat.int/lunar-calibration (accessed on 31 May 2021).
96. Choi, T.; Shao, X.; Cao, C. On-orbit radiometric calibration of Suomi NPP VIIRS reflective solar bands using the Moon and solar diffuser. *Appl. Opt.* **2018**, *57*, 9533–9542. [CrossRef] [PubMed]
97. Lunar Irradiance Model ESA (LIME). Available online: http://calvalportal.ceos.org/lime;jsessionid=D4529A9DCB79ABBDD2CE4187D45557CB (accessed on 31 May 2021).
98. Taylor, S.; Adriaensen, S.; Toledano, C.; Barreto, Á.; Woolliams, E.; Bouvet, M. LIME: The Lunar Irradiance Model of the European Space Agency. In Proceedings of the EGU General Assembly 2021, Online Event, 19–30 April 2021.
99. Neneman, M.; Wagner, S.; Bourg, L.; Blanot, L.; Bouvet, M.; Adriaensen, S.; Nieke, J. Use of Moon Observations for Characterization of Sentinel-3B Ocean and Land Color Instrument. *Remote Sens.* **2020**, *12*, 2543. [CrossRef]
100. Presentation of the Lunar Irradiance Model Inter-Comparison Exercise Made by T. Stone (USGS) during GSICS 2021 Annual Meeting, See Lunar Session within the GSICS Web Portal. Available online: http://gsics.atmos.umd.edu/bin/view/Development/Annual2021 (accessed on 30 June 2021).
101. Chander, G.; Mishra, N.; Helder, D.L.; Aaron, D.B.; Angal, A.; Choi, T.; Xiong, X.; Doelling, D.R. Applications of spectral band adjustment factors (SBAF) for cross-calibration. *IEEE Trans. Geosci. Remote Sens.* **2012**, *51*, 1267–1281. [CrossRef]
102. Cao, C.; Xiong, X.; Wu, A.; Wu, X. Assessing the consistency of AVHRR and MODIS L1B reflectance for generating fundamental climate data records. *J. Geophys. Res. Atmos.* **2008**, *113*. [CrossRef]
103. Clerc, S.; Donlon, C.; Borde, F.; Lamquin, N.; Hunt, S.E.; Smith, D.; McMillan, M.; Mittaz, J.; Woolliams, E.; Hammond, M.; et al. Benefits and lessons learned from the Sentinel-3 tandem phase. *Remote Sens.* **2020**, *12*, 2668. [CrossRef]
104. Lamquin, N.; Clerc, S.; Bourg, L.; Donlon, C. OLCI A/B tandem phase analysis, part 1: Level 1 homogenisation and harmonisation. *Remote Sens.* **2020**, *12*, 1804. [CrossRef]
105. Lamquin, N.; Déru, A.; Clerc, S.; Bourg, L.; Donlon, C. OLCI A/B tandem phase analysis, part 2: Benefits of sensors harmonisation for Level 2 products. *Remote Sens* **2020**, *12*, 2702. [CrossRef]
106. Hunt, S.E.; Mittaz, J.; Smith, D.; Polehampton, E.; Yemelyanova, R.; Woolliams, E.R.; Donlon, C. Comparison of the Sentinel-3A and B SLSTR Tandem Phase Data using metrological principles. *Remote Sens.* **2020**, *12*, 2893. [CrossRef]
107. Bouvet, M.; Thome, K.; Berthelot, B.; Bialek, A.; Czapla-Myers, J.; Fox, N.; Goryl, P.; Henry, P.; Ma, L.; Marcq, S.; et al. RadCalNet: A radiometric calibration network for Earth observing imagers operating in the visible to shortwave infrared spectral range. *Remote Sens.* **2019**, *11*, 2401. [CrossRef]
108. RadCalNet Web Portal. Available online: https://www.radcalnet.org/ (accessed on 31 May 2021).
109. Sterckx, S.; Wolters, E. Radiometric top-of-atmosphere reflectance consistency assessment for landsat 8/OLI, Sentinel-2/MSI, PROBA-V, and DEIMOS-1 over Libya-4 and RadCalNet calibration sites. *Remote Sens.* **2019**, *11*, 2253. [CrossRef]
110. Banks, A.C.; Hunt, S.E.; Gorroño, J.; Scanlon, T.; Woolliams, E.R.; Fox, N. A comparison of validation and vicarious calibration of high and medium resolution satellite-borne sensors using RadCalNet. In Sensors, Systems, and Next-Generation Satellites XXI. *Int. Soc. Opt. Photonics* **2017**, *10423*, 10423A.
111. Jing, X.; Leigh, L.; Teixeira Pinto, C.; Helder, D. Evaluation of RadCalNet Output Data Using Landsat 7, Landsat 8, Sentinel 2A, and Sentinel 2B Sensors. *Remote Sens.* **2019**, *11*, 541. [CrossRef]
112. Alhammoud, B.; Jackson, J.; Clerc, S.; Arias, M.; Bouzinac, C.; Gascon, F.; Cadau, E.G.; Iannone, R.Q.; Boccia, V. Sentinel-2 Level-1 Radiometry Assessment Using Vicarious Methods from DIMITRI Toolbox and Field Measurements From RadCalNet Database. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2019**, *12*, 3470–3479. [CrossRef]
113. TRUTHS Web Portal at NPL. Available online: https://www.npl.co.uk/earth-observation/truths (accessed on 31 May 2021).
114. Fox, N.; Green, P. Traceable radiometry underpinning terrestrial-and helio-studies (TRUTHS): An element of a space-based climate and calibration observatory. *Remote Sens.* **2020**, *12*, 2400. [CrossRef]
115. Wielicki, B.A.; Young, D.F.; Mlynczak, M.G.; Thome, K.J.; Leroy, S.; Corliss, J.; Anderson, J.G.; Ao, C.O.; Bantges, R.; Best, F.; et al. Achieving climate change absolute accuracy in orbit. *Bull. Am. Meteorol. Soc.* **2013**, *94*, 1519–1539. [CrossRef]
116. Shea, Y.; Fleming, G.; Kopp, G.; Lukashin, C.; Pilewskie, P.; Smith, P.; Thome, K.; Wielicki, B.; Liu, X.; Wu, W. Clarreo Pathfinder: Mission Overview and Current Status. In Proceedings of the IGARSS 2020 IEEE International Geoscience and Remote Sensing Symposium, Virtual Symposium, 26 September–2 October 2020; pp. 3286–3289.
117. Bréon, F.M.; Vermote, E. Correction of MODIS surface reflectance time series for BRDF effects. *Remote Sens. Environ.* **2012**, *125*, 1–9. [CrossRef]
118. Doxani, G.; Vermote, E.; Roger, J.C.; Gascon, F.; Adriaensen, S.; Frantz, D.; Hagolle, O.; Hollstein, A.; Kirches, G.; Li, F.; et al. Atmospheric correction inter-comparison exercise. *Remote Sens.* **2018**, *10*, 352. [CrossRef]
119. Kotchenova, S.Y.; Vermote, E.F.; Matarrese, R.; Klemm, F.J., Jr. Validation of a vector version of the 6S radiative transfer code for atmospheric correction of satellite data. Part I: Path radiance. *Opt. Soc. Am. Appl. Opt.* **2006**, *45*, 6762–6774.
120. Vermote, E.; Doxani, G.; Roger, J.C.; Gascon, F.; Skakun, S. Atmospheric Correction Inter-Comparison Exercise II. In Proceedings of the American Geophysical Union (AGU) Fall Meeting 2019, San Francisco, CA, USA, 9–13 December 2019.

121. Pahlevan, N.; Mangin, A.; Balasubramanian, S.V.; Smith, B.; Alikas, K.; Arai, K.; Barbosa, C.; Bélanger, S.; Binding, C.; Bresciani, M.; et al. ACIX-Aqua: A global assessment of atmospheric correction methods for Landsat-8 and Sentinel-2 over lakes, rivers, and coastal waters. *Remote Sens. Environ.* **2021**, *258*, 112366. [CrossRef]
122. Emde, C.; Buras-Schnell, R.; Kylling, A.; Mayer, B.; Gasteiger, J.; Hamann, U.; Kylling, J.; Richter, B.; Pause, C.; Dowling, T.; et al. The libradtran software package for radiative transfer calculations. *Geosci. Model Dev.* **2016**, *9*, 1647–1672. [CrossRef]
123. Berk, A.; Conforti, P.; Kennett, R.; Perkins, T.; Hawes, F.; van den Bosch, J. MODTRAN®6: A major upgrade of the MODTRAN®radiative transfer code. In Proceedings of the 2014 6th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), Lausanne, Switzerland, 24–27 June 2014; pp. 1–4. [CrossRef]
124. Kotchenova, S.Y.; Vermote, E.F.; Levy, R.; Lyapustin, A. Radiative transfer codes for atmospheric correction and aerosol retrieval: Intercomparison study. *Appl. Opt.* **2008**, *47*, 2215–2226. [CrossRef] [PubMed]
125. RAdiation Transfer Model Intercomparison for Atmosphere (RAMI4ATM) Web Portal. Available online: https://rami-benchmark.jrc.ec.europa.eu/_www/RAMI4ATM.php (accessed on 31 May 2021).
126. RAdiation transfer Model Intercomparison-V (RAMI-V) Web Portal. Available online: https://rami-benchmark.jrc.ec.europa.eu/_www/index.php (accessed on 31 May 2021).
127. Web portal of the ACIX II—CMIX 2Nd WS. Available online: https://earth.esa.int/eogateway/events/ceos-wgcv-acix-ii-cmix-atmospheric-correction-inter-comparison-exercise-cloud-masking-inter-comparison-exercise-2nd-workshop?text=CMIX (accessed on 31 May 2021).
128. Tarrio, K.; Tang, X.; Masek, J.G.; Claverie, M.; Ju, J.; Qiu, S.; Zhu, Z.; Woodcock, C.E. Comparison of cloud detection algorithms for Sentinel-2 imagery. *Sci. Remote Sens.* **2020**, *2*, 100010. [CrossRef]
129. Zekoll, V.; Main-Knorn, M.; Louis, J.; Frantz, D.; Richter, R.; Pflug, B. Comparison of masking algorithms for sentinel-2 imagery. *Remote Sens.* **2021**, *13*, 137. [CrossRef]
130. Baetens, L.; Desjardins, C.; Hagolle, O. Validation of copernicus Sentinel-2 cloud masks obtained from MAJA, Sen2Cor, and FMask processors using reference cloud masks generated with a supervised active learning procedure. *Remote Sens.* **2019**, *11*, 433. [CrossRef]
131. Frantz, D. FORCE—Landsat+ Sentinel-2 analysis ready data and beyond. *Remote Sens.* **2019**, *11*, 1124. [CrossRef]
132. Glossary of NOAA Weather Service. Available online: https://forecast.weather.gov/glossary.php?word=CLD (accessed on 31 May 2021).
133. Hollstein, A.; Segl, K.; Guanter, L.; Brell, M.; Enesco, M. Ready-to-use methods for the detection of clouds, cirrus, snow, shadow, water and clear sky pixels in Sentinel-2 MSI images. *Remote Sens.* **2016**, *8*, 666. [CrossRef]
134. Foga, S.; Scaramuzza, L.; Guo, S.; Zhu, Z.; Dilley, R.D., Jr.; Beckmann, T.; Schmidt, G.L.; Dwyer, J.L.; Hughes, M.J.; Laue, B. Cloud detection algorithm comparison and validation for operational Landsat data products. *Remote Sens. Environ.* **2017**, *194*, 379–390. [CrossRef]
135. Skakun, S.; Vermote, E.F.; Artigas, A.E.S.; Rountree, W.H.; Roger, J.C. An experimental sky-image-derived cloud validation dataset for Sentinel-2 and Landsat 8 satellites over NASA GSFC. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *95*, 102253. [CrossRef]
136. HYPERNETS Web Portal. Available online: https://www.hypernets.eu/ (accessed on 31 May 2021).
137. Presentation of Hypernets Project Made by K. Ruddick (RBINS) during ESA Very-High-Resolution (VHR-RODA) WS. Available online: https://earth.esa.int/eogateway/documents/20142/37627/1B3+HYPERNETS%40ESA-VHRODA_2021-04-21-FINAL-Distribution.pdf/dba957ee-dc39-e90d-4bdf-f73100078c6e (accessed on 31 May 2021).
138. Zibordi, G.; Mélin, F.; Berthon, J.F.; Holben, B.; Slutsker, I.; Giles, D.; D'Alimonte, D.; Vandemark, D.; Feng, H.; Schuster, G.; et al. AERONET-OC: A network for the validation of ocean color primary products. *J. Atmos. Ocean. Technol.* **2009**, *26*, 1634–1651. [CrossRef]
139. Mélin, F.; Sclep, G. Band shifting for ocean color multi-spectral reflectance data. *Optics Express* **2015**, *23*, 2262–2279. [CrossRef] [PubMed]
140. Pahlevan, N.; Smith, B.; Binding, C.; O'Donnell, D.M. Spectral band adjustments for remote sensing reflectance spectra in coastal/inland waters. *Opt. Express* **2017**, *25*, 28650–28667. [CrossRef]
141. Vanhellemont, Q. Sensitivity analysis of the dark spectrum fitting atmospheric correction for metre-and decametre-scale satellite imagery using autonomous hyperspectral radiometry. *Opt. Express* **2020**, *28*, 29948–29965. [CrossRef]
142. Sánchez-Zapero, J.; Camacho, F.; Martínez-Sánchez, E.; Lacaze, R.; Carrer, D.; Pinault, F.; Benhadj, I.; Muñoz-Sabater, J. Quality Assessment of PROBA-V Surface Albedo V1 for the Continuity of the Copernicus Climate Change Service. *Remote Sens.* **2020**, *12*, 2596. [CrossRef]
143. D'Odorico, P.; Gonsamo, A.; Pinty, B.; Gobron, N.; Coops, N.; Mendez, E.; Schaepman, M.E. Intercomparison of fraction of absorbed photosynthetically active radiation products derived from satellite data over Europe. *Remote Sens. Environ.* **2014**, *142*, 141–154. [CrossRef]
144. Fuster, B.; Sánchez-Zapero, J.; Camacho, F.; García-Santos, V.; Verger, A.; Lacaze, R.; Weiss, M.; Baret, F.; Smets, B. Quality assessment of PROBA-V LAI, fAPAR and fCOVER collection 300 m products of copernicus global land service. *Remote Sens.* **2020**, *12*, 1017. [CrossRef]
145. Cendrero-Mateo, M.; Wieneke, S.; Damm, A.; Alonso, L.; Pinto, F.; Moreno, J.; Guanter, L.; Celesti, M.; Rossini, M.; Sabater, N.; et al. Sun-induced chlorophyll fluorescence III: Benchmarking retrieval methods and sensor characteristics for proximal sensing. *Remote Sens.* **2019**, *11*, 962. [CrossRef]

146. Porcar-Castell, A.; Tyystjärvi, E.; Atherton, J.; Van der Tol, C.; Flexas, J.; Pfündel, E.E.; Moreno, J.; Frankenberg, C.; Berry, J.A. Linking chlorophyll a fluorescence to photosynthesis for remote sensing applications: Mechanisms and challenges. *J. Exp. Bot.* **2014**, *65*, 4065–4095. [CrossRef]
147. Schüttemeyer, D.; Burba, M.; Drusch, M.; Elfving, A.; Mecklenburg, S. ESA's campaign activities in support of the FLEX mission. In Proceedings of the IGARSS 2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018; pp. 3924–3926.
148. Fluorescence Box Web Portal. Available online: https://www.jb-hyperspectral.com/products/flox/ (accessed on 31 May 2021).
149. Fernandes, R.; Plummer, S.; Nightingale, J.; Frederic, B.; Camacho de Coca, F.; Fang, H.; Garrigues, S.; Gobron, N.; Lang, M.; Lacaze, R.; et al. *CEOS Global LAI Product Validation Good Practices*; CEOS: Washington, DC, USA, 2014.
150. CEOS-LPV Document Repository. Available online: https://lpvs.gsfc.nasa.gov/documents.html (accessed on 31 May 2021).
151. Bayat, B.; Camacho, F.; Nickeson, J.; Cosh, M.; Bolten, J.; Vereecken, H.; Montzka, C. Toward operational validation systems for global satellite-based terrestrial essential climate variables. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *95*, 102240. [CrossRef]
152. ESA FRM4VEG Web Portal. Available online: https://frm4veg.org/ (accessed on 31 May 2021).
153. Origo, N.; Gorroño, J.; Ryder, J.; Nightingale, J.; Bialek, A. Fiducial Reference Measurements for validation of Sentinel-2 and Proba-V surface reflectance products. *Remote Sens. Environ.* **2020**, *241*, 111690. [CrossRef]
154. ESA SRIX4VEG Web Portal. Available online: https://frm4veg.org/srix4veg/ (accessed on 31 May 2021).
155. CEOS LPV Supersites. Available online: https://lpvs.gsfc.nasa.gov/LPV_Supersites/LPVsites.html (accessed on 31 May 2021).
156. National Science Foundation's National Ecological Observatory Network (NEON) Web Site. Available online: https://neonscience.org (accessed on 31 May 2021).
157. Kampe, T.U.; Johnson, B.R.; Kuester, M.A.; Keller, M. NEON: The first continental-scale ecological observatory with airborne remote sensing of vegetation canopy biochemistry and structure. *J. Appl. Remote Sens.* **2010**, *4*, 043510. [CrossRef]
158. Surface Radiation Budget Network (SURFRAD) Web Site. Available online: https://www.esrl.noaa.gov/gmd/grad/surfrad/index.html (accessed on 31 May 2021).
159. Augustine, J.A.; DeLuisi, J.J.; Long, C.N. SURFRAD–A national surface radiation budget network for atmospheric research. *Bull. Am. Meteorol. Soc.* **2000**, *81*, 2341–2358. [CrossRef]
160. Integrated Carbon Observation System (ICOS) Web Site. Available online: https://www.icos-ri.eu/icos-research-infrastructure (accessed on 31 May 2021).
161. Gielen, B.; de Beeck, M.O.; Loustau, D.; Ceulemans, R.; Jordan, A.; Papale, D. Integrated carbon observation system (icos): An infrastructure to monitor the european greenhouse gas balance. In *Terrestrial Ecosystem Research Infrastructures: Challenges and Opportunities*; Chabbi, A., Loescher, H.W., Eds.; CRC Press, Taylor & Francis Group: Boca Raton, FL, USA, 2017; pp. 506–520. ISBN 9781498751315.
162. Baldocchi, D.; Falge, E.; Gu, L.; Olson, R.; Hollinger, D.; Running, S.; Anthoni, P.; Bernhofer, C.; Davis, K.; Evans, R.; et al. FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water.vapor, and energy flux densities. *Bull. Am. Meteorol. Soc.* **2001**, *82*, 2415–2434. [CrossRef]
163. FLUXNET Web-Site. Available online: https://fluxnet.fluxdata.org/ (accessed on 31 May 2021).
164. Terrestrial Ecosystem Research Network (TERN). Available online: https://www.tern.org.au/ (accessed on 31 May 2021).
165. Cleverly, J.; Eamus, D.; Edwards, W.; Grant, M.; Grundy, M.J.; Held, A.; Karan, M.; Lowe, A.J.; Prober, S.M.; Sparrow, B.; et al. TERN, Australia's land observatory: Addressing the global challenge of forecasting ecosystem responses to climate variability and change. *Environ. Res. Lett.* **2019**, *14*, 095004. [CrossRef]
166. Roosjen, P.P.; Brede, B.; Suomalainen, J.M.; Bartholomeus, H.M.; Kooistra, L.; Clevers, J.G. Improved estimation of leaf area index and leaf chlorophyll content of a potato crop using multi-angle spectral data–potential of unmanned aerial vehicle imagery. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *66*, 14–26. [CrossRef]
167. Arroyo-Mora, J.; Kalacska, M.; Inamdar, D.; Soffer, R.; Lucanus, O.; Gorman, J.; Naprstek, T.; Schaaf, E.S.; Ifimov, G.; Elmer, K.; et al. Implementation of a UAV–hyperspectral pushbroom imager for ecological monitoring. *Drones* **2019**, *3*, 12. [CrossRef]
168. Arroyo-Mora, J.; Kalacska, M.; Løke, T.; Schläpfer, D.; Coops, N.C.; Lucanus, O.; Leblanc, G. Assessing the impact of illumination on UAV pushbroom hyperspectral imagery collected under various cloud cover conditions. *Remote Sens. Environ.* **2021**, *258*, 112396. [CrossRef]
169. Brown, L.A.; Ogutu, B.O.; Dash, J. Tracking forest biophysical properties with automated digital repeat photography: A fisheye perspective using digital hemispherical photography from below the canopy. *Agric. For. Meteorol.* **2020**, *287*, 107944. [CrossRef]
170. Brown, L.A.; Meier, C.; Morris, H.; Pastor-Guzman, J.; Bai, G.; Lerebourg, C.; Gobron, N.; Lanconelli, C.; Clerici, M.; Dash, J. Evaluation of global leaf area index and fraction of absorbed photosynthetically active radiation products over North America using Copernicus Ground Based Observations for Validation data. *Remote Sens. Environ.* **2020**, *247*, 111935. [CrossRef]
171. Wilkes, P.; Lau, A.; Disney, M.; Calders, K.; Burt, A.; de Tanago, J.G.; Bartholomeus, H.; Brede, B.; Herold, M. Data acquisition considerations for terrestrial laser scanning of forest plots. *Remote Sens. Environ.* **2017**, *196*, 140–153. [CrossRef]
172. Lovell, J.L.; Jupp, D.L.; Culvenor, D.S.; Coops, N.C. Using airborne and ground-based ranging lidar to measure canopy structure in Australian forests. *Can. J. Remote Sens.* **2003**, *29*, 607–622. [CrossRef]
173. Pimont, F.; Soma, M.; Dupuy, J.L. Accounting for wood, foliage properties, and laser effective footprint in estimations of leaf area density from multiview-LiDAR data. *Remote Sens.* **2019**, *11*, 1580. [CrossRef]

174. Brede, B.; Calders, K.; Lau, A.; Raumonen, P.; Bartholomeus, H.M.; Herold, M.; Kooistra, L. Non-destructive tree volume estimation through quantitative structure modelling: Comparing UAV laser scanning with terrestrial LIDAR. *Remote Sens. Environ.* **2019**, *233*, 111355. [CrossRef]
175. Weiss, M.; Baret, F.; Block, T.; Koetz, B.; Burini, A.; Scholze, B.; Lecharpentier, P.; Brockmann, C.; Fernandes, R.; Plummer, S.; et al. On Line Validation Exercise (OLIVE): A web based service for the validation of medium resolution land products. *Appl. FAPAR Products Remote Sens.* **2014**, *6*, 4190–4216. [CrossRef]
176. Baret, F.; Morissette, J.; Fernandes, R.; Champeaux, J.L.; Myneni, R.; Chen, J.; Plummer, S.; Weiss, M.; Bacour, C.; Garrigues, S.; et al. Evaluation of the representativeness of networks of sites for the validation and inter-comparison of land biophysical products. *IEEE Trans. Geosc. Remote Sens.* **2006**, *44*, 1794–1803. [CrossRef]
177. Camacho, F.; Cernicharo, J.; Lacaze, R.; Baret, F.; Weiss, M. GEOV1: LAI, FAPAR essential climate variables and FCOVER global time series capitalizing over existing products. Part 2: Validation and intercomparison with reference products. *Remote Sens. Environ.* **2013**, *137*, 310–329. [CrossRef]
178. Sanchez-Zapero, J.; Madrid, L.; dela Camacho, F. SALVAL: A semi-automatic surface albedo validation tool. In Proceedings of the V International Symposium on Recent Advances in Quantitative Remote Sensing (RAQRS), Valencia, Spain, 18–22 September 2017.
179. ESA EVDC Web Portal. Available online: https://evdc.esa.int/ (accessed on 31 May 2021).
180. Copernicus Calibration and Validation Solution Web Portal. Available online: https://ccvs.eu/ (accessed on 31 May 2021).
181. Wilkinson, M.D.; Dumontier, M.; Aalbersberg, I.J.; Appleton, G.; Axton, M.; Baak, A.; Blomberg, N.; Boiten, J.W.; da Silva Santos, L.B.; Bourne, E.; et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* **2016**, *3*, 1–9. [CrossRef]
182. Joint Experiment for Crop Assessment and Monitoring (JECAM) Web Portal. Available online: http://jecam.org/ (accessed on 31 May 2021).
183. The FLARE Network Web Portal. Available online: https://flare-network.com/ (accessed on 31 May 2021).