



Fiducial Reference Measurements (FRMs): What Are They?

Philippe Goryl ^{1,*}, Nigel Fox ², Craig Donlon ³ and Paolo Castracane ⁴

¹ European Space Agency ESA/ESRIN, 00044 Frascati, Italy

² Earth Observation, Climate and Optical Group, National Physical Laboratory, Teddington TW11 0LW, UK; nigel.fox@npl.co.uk

³ European Space Agency ESA/ESTEC, 2201 Noordwijk, The Netherlands; craig.donlon@esa.int

⁴ Rhea System S.p.A. for ESA/ESRIN, 00044 Frascati, Italy; paolo.castracane@ext.esa.int

* Correspondence: philippe.goryl@esa.int

Abstract: In recent years, the concept of a Fiducial Reference Measurement (FRM) has been developed to highlight the need for precise and well-characterised measurements tailored explicitly to the post-launch calibration and validation (Cal/Val) of Earth observation satellite missions. The confidence that stems from robust, unambiguous uncertainty assessment of space observations is fundamental to assessing the changes in the Earth system and climate model prediction and delivering the essential evidence-based input for policy makers and society striving to mitigate and adapt to climate change. The underlying concept of an FRM has long been a core element of a Cal/Val program, providing a ‘trustable’ reference against which performance can be anchored or assessed. The ‘FRM’ label was created to embody into such a reference a set of key criteria. These criteria included the establishment of documented evidence of uncertainty with respect to a community-agreed reference (ideally SI-traceable) and specific tailoring to the needs of a satellite mission. It therefore facilitates comparison and interoperability between products and missions in a cost-efficient manner. Committee on Earth Observation Satellites (CEOS) Working Group Cal/Val (WGCV) is now putting in place a framework to assess the maturity and compliance of a ‘Cal/Val reference measurement’ in terms of a set of community-agreed criteria which define it to be of CEOS-FRM quality. The assessment process is based on a maturity matrix that provides a visual assessment of the state of any FRM against each of a set of given criteria, making visible where it is mature and where evolution and effort are still needed. This paper provides the overarching definition of what constitutes an FRM and introduces the new CEOS-FRM assessment framework.

Keywords: earth observation; remote sensing; calibration; validation; fiducial reference measurements; CEOS



Citation: Goryl, P.; Fox, N.;

Donlon, C.; Castracane, P. Fiducial Reference Measurements (FRMs): What Are They? *Remote Sens.* **2023**, *15*, 5017. <https://doi.org/10.3390/rs15205017>

Academic Editor: Raquel De Los Reyes

Received: 17 August 2023

Revised: 12 October 2023

Accepted: 16 October 2023

Published: 18 October 2023



Copyright: © 2023 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

The need for post-launch calibration and validation (Cal/Val) of satellite sensors is well-established. For the primary level 1 observations, post-launch calibration against independent references external to the spacecraft ensures that unforeseen changes, and changes that cannot be corrected by any on-board calibration systems, are accounted for. Post-launch validation supports the accurate delivery of level 1 and higher-level products, confirming expected performance including processors and any retrieval algorithms. However, the cost, complexity and robustness of establishing mission-specific exercises have led to efforts by space agencies to improve coordination and generalisation of methods and infrastructure, through bodies like the Committee on Earth Observation Satellites (CEOS). This coordination not only reduces cost, but also leads to improved harmonisation and interoperability between sensors. Earth Observation data without proper calibration and uncertainty assessment has little value for most applications, because if the data cannot be trusted, no reliable information can be derived from them.

In parallel, driven in part by the critical needs of climate change, there has been a significant move towards a more coordinated and comprehensive assessment and reporting of the quality, biases and residual uncertainty in the observations made by different satellite sensors. This is illustrated by examples of infrastructure and methods on the CEOS calvalportal [1], most notably that of the Radiometric Calibration Network (RadCalNet) [2]. More recently, this need has been enhanced with the emergence of new commercial satellite operators ensuring that users have a transparent and fair means to judge the adequacy of data products to meet their needs.

All of these developments have led to the endorsement by The Group on Earth Observations (GEO) of a CEOS Working Group Cal/Val (WGCV) and World Meteorological Organization (WMO) Global Space-based Inter-Calibration System (GSICS)-led community initiative to create a Quality Assurance Framework for Earth Observation (QA4EO) [3]. QA4EO provides a set of principles and associated guidance including specifically tailored tools [4] to encourage the provision of internationally consistent quality indicators on their delivered data. QA4EO guidelines state that *it is critical that data and derived products are easily accessible in an open manner and have an associated indicator of quality traceable to reference standards (preferably SI) so users can assess suitability for their applications i.e., 'fitness for purpose'*.

The QA4EO principles, and means to establish and demonstrate them, were tailored from those in common usage across many commercial and scientific sectors, ultimately derived from the best practices of the international metrology community. The latter is charged by intergovernmental treaties with ensuring international consistency and long-term reliability of measurements to support trade and legislation, through the international system of units (SI) [5]. The robustness of the SI and its dissemination over century-long timescales is underpinned by adherence to clear principles:

Unequivocal, documented, consensus definition of 'references' and the *mises-en-pratique* [6] to realise them.

Documented evidence of the degree of traceability (residual bias) and associated uncertainty of a measurement relative to the 'reference' through:

- regular, openly reported comparison with independent peers (where there is no independent representative measurement with lower uncertainties), or
- comparison (validation/calibration) against an independent representative measurement of the same measurement with similar or lower uncertainties than that being evidenced.
- detailed comprehensive assessment of uncertainty of the measurement and its traceability to the reference that is peer reviewed alongside comparison results.

One key element of QA4EO relates to post-launch Cal/Val of satellite-delivered data through an independently derived data set that can be correlated to that of the satellite observation. Of course, for this process to be robust, the independent data set itself must also be fully characterised in a manner consistent with QA4EO principles: i.e., that it has documented evidence of its level of consistency with a suitable reference (nominal 'truth') within an associated uncertainty. The uncertainty associated with the comparison process itself, which almost always involves a transformation of the 'reference' measurand to be comparable/representative of that observed by the satellite, must also be characterised and be sufficiently small. In practice, the 'reference' should be community agreed and ideally tied to SI. The need to evidence these core principles as a differentiator is the basis for creating the label Fiducial Reference Measurement (FRM).

In this paper, we provide some background as to the critical need for such a tailored set of validation data (FRM) (Section 2). This is followed in Sections 3.1 and 3.2 by a potted history of the creation of the core criteria that must be fulfilled to achieve the FRM label. In Section 3.3 we expand and further embellish the original criteria to ensure they are generic and 'fit for purpose' for a wide range of satellite sensors and applications, as a guide to users and developers of FRMs. Finally, in recognition of the need to ensure consistency in the application and interpretation of the FRM mandatory and defining criteria, we introduce,

in Section 3.4, a quality assurance and ‘endorsement’ process, under development within CEOS WGCV, utilizing a ‘maturity matrix’-based reporting mechanism [7]. This final step will aid space agencies and ‘new space’ providers to readily find suitable Cal/Val data that meet their needs and also potential funders of FRM developers to optimise their investments.

2. Post-Launch Calibration and Validation (Cal/Val)

In a generic manner, a post-launch Cal/Val program for an Earth Observation (EO) mission is typically composed of a set of different complementary activities bringing elements that need to be combined in order to produce consolidated and confident validation results to encompass the full scope of observations of a sensor. Note that the use of the word ‘comparison’ represents the situation where the nature of the products being compared is similar and of a similar uncertainty. ‘Validation’ is used where the link to the satellite product is more tenuous and/or the associated uncertainty of the non-satellite measurement is larger and cannot be considered similar, as in the later bullets. ‘Calibration’ is a special case of comparison where the uncertainty of the reference and the comparison process are sufficiently small that this allows the satellite sensor’s measurement to be assessed against it and potentially adjusted as part of the comparison process. Generically, the different components of a validation programme are the following:

- Comparison against tailored and accurate Fiducial Reference Measurements (FRMs) that provide few points but low uncertainty/high confidence;
- Comparisons against general in situ data that can provide more data over a wider set of conditions, but are generally individually less accurate;
- Comparison against other sources, including: inter-satellite comparison, satellites own Level 3 (including visual interpretation) (i.e., geo-referenced product of merged data derived from the sensor itself, and/or other sensors, e.g., seasonal trends, latitude variations) climatologies, etc.;
- Comparison against models: re-analysis data, data assimilation, integrated model analyses, forward modelling, etc.;
- Validation using monitoring tools: statistics, trends, systematic quality control, etc.;
- Comparison/calibration between operational satellites (e.g., using GSICS best practices).

All of these components are important and often necessary, but the first component (comparison to FRM) is of particular importance because it gives a reference properly characterised and traceable to a standard on which the validation results can be anchored and an uncertainty assessed. If of a sufficiently low uncertainty, the FRM can also be used to provide calibration information that can be used to update the sensor performance coefficients during operation.

As indicated above, satellite post-launch Cal/Val has traditionally relied on a wide range of data sets collated by different teams and based on, for example, in situ and aircraft campaigns, (semi)-permanently deployed instrumentation, comparisons with other satellite sensors, opportunistic measurements (e.g., from commercial shipping and aviation) and modelled natural phenomena (e.g., Rayleigh scattering of the atmosphere over dark ocean surfaces).

To date, these different datasets have had varying levels of adherence to the critical QA4EO principles, and while all provide information, it is hard for satellite operators to distinguish high-quality data optimised for satellites from those that need more careful treatment. Consistency between datasets for comparison with a single satellite sensor is generally not sufficient to ensure the quality of all comparisons, for all sensors and all conditions. The creation and adherence to the FRM label is designed to address this issue.

3. Fiducial Reference Measurements (FRMs)

3.1. Need for FRMs

The terminology ‘Fiducial Reference Measurement’, with its initialism ‘FRM’, was first elaborated and used in the framework of the Sentinel-3 Validation Team (S3VT) meeting

in 2013 [8]. It was particularly relevant in the context of Copernicus for distinguishing between in situ measurements that are managed by the Copernicus in situ component [9], which are also used for other purposes in their own right and the Cal/Val measurements managed by the Copernicus Space Component that are used for satellite-specific Calibration or Validation purposes. As a result, the term FRM was initially defined as “the suite of independent tailored and fully characterised 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.” [8].

Figure 1 illustrates this principle: If we have no FRMs, then we cannot really use the mission for any quantitative purpose as we have no idea how accurate the data products are; if we have many FRMs, this is great scientifically (statistical significance, geographic/condition coverage, robust network, etc.) but incurs additional costs, with a reduced ROI. There is a balance between these two extremes to deliver a satellite mission with a known product quality that is “fit for purpose”.

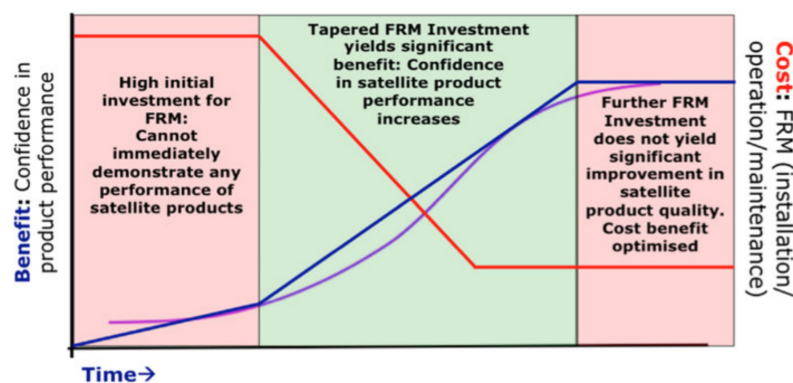


Figure 1. A simple cost/benefit analysis of the FRM concept. The red line indicates the investment profile and cost over time. The cost is high during initial investment to create the infrastructure and capability, then drops for operations. The blue line indicates benefits to the satellite mission (measured as ‘confidence in product performance’) over time. The purple curve indicates a hypothetical optimum cost/benefit curve. FRM: fiducial reference measurement. Note it could be argued that in the longer term, the ROI actually continues to increase as more satellite users benefit.

In an attempt to create a more robust validation framework and encourage, where appropriate, teams to specifically tailor their observations to address the needs of satellites, as well as helping funding bodies distinguish and allocate resources, the European Space Agency (ESA) expanded the concept to other applications and technology domains and actively encouraged the creation and utilization of the label ‘Fiducial Reference Measurement’ in the international arena through bodies such as CEOS. In creating the FRM label and associated criteria (see below), the objective has been to create a distinction and special case of observations that are optimised to meet the needs of the satellite community.

Note that the use of FRM labels for particular data sets does not imply that observations not classified as FRMs are of inferior quality. The FRM label is used where specific requirements (Section 3.2) are met for a specific satellite measurand/product and, in particular, where datasets are suitable for a defined class of satellite observations.

These principles subsequently formed the basis for an ongoing series of projects funded by European organisations, European Union (EU), ESA, European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) to explore, establish and evidence FRM quality observations and in many cases, networks of sites, around the globe that adhere to the FRM principles for a wide range of observation and sensor types spanning surface temperature and reflectance, atmospheric composition, sea and ice level height, etc.

3.2. What Is an FRM?

A Fiducial Reference Measurement (FRM) is:

a suite of independent, fully characterised, and traceable (to a community agreed reference, ideally SI) measurements of a satellite relevant measurand, tailored specifically to address the calibration/validation needs of a class of satellite borne sensor and that follow the guidelines outlined by the GEO/CEOS Quality Assurance framework for Earth Observation (QA4EO) [3].

In many cases, FRMs may be a subset of ‘in situ’ measurements, originally designed to make observations in their own right at a specific geo-location, both individually and/or as part of a network, but they could also include measurements from a campaign, from airborne or seaborne platforms, or even from another satellite, providing it meets the FRM criteria. The newly emerging SITSats (SI Traceable Satellites) [10] are a good example of a potential satellite-based FRM. These FRMs provide the maximum Return On Investment (ROI) for a satellite mission by delivering, to satellite operators and their 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 defining mandatory characteristics for an FRM are:

1. **Traceability**—FRM measurements should have documented evidence of their traceability (including assessment of any bias and its associated uncertainty) to a community agreed reference ideally tied to the International System of units, SI, (e.g. via a comparison ‘round robin’ or other with peers and/or a metrology institute together with regular pre- and post-deployment calibration of instruments). This should be carried out using SI-traceable ‘metrology’ standards and/or community recognised best practices, for both instrumentation and observations;
2. **Independence of satellite-under-test**—FRM measurements are independent from the satellite (under test) geophysical retrieval process;
3. **Uncertainty budget**—A comprehensive uncertainty budget for all FRM instruments and for the derived (processed) measurements (including any transformation of the measurand to match the measurand of the satellite product) is available and maintained;
4. **Documented protocols**—FRM measurement protocols, procedures and community-wide quality management practices (measurement, processing, archive, documents, etc.) are defined, published and adhered to by FRM instrument deployments;
5. **Accessibility**—FRM datasets, including metadata and reports documenting processing, are accessible to other researchers, allowing independent verification of processing systems;
6. **Representativeness**—FRM datasets are required to determine the on-orbit uncertainty characteristics of satellite geophysical measurements via independent validation activities. The degree of representativeness of the FRM to that of the satellite observation as well as the satellite to FRM comparison process are thus required to be documented and the uncertainty assessed. Note for any individual satellite sensor, the exact sampling and elements of the comparison process may differ, even within a generic sensor class, but the documentation and evidence to support the uncertainty analysis must be presented in a manner that can be readily interpreted by a user. The data and associated uncertainty information should be available in a timely manner and in a form that is readily utilizable by a satellite operator.
7. **Adequacy of Uncertainty**—The uncertainty of the FRM, including the comparison process, must be commensurate with the requirements of the class of satellite sensor they are specified to support.
8. **Utility**—FRM datasets are designed to apply to a class of satellite missions. They are not mission-specific. For example, in the spectral domain, it is expected that different

spectral bands of different sensors can be addressed and not only tailored exclusively to the exact bands of a single sensor such as Sentinel 2.

These principles describe a generic approach which can be adapted to suit the requirements of specific satellite measurements and/or validation methodology but simplistically, FRMs should all, where practical, seek to achieve the following:

- Validation/measurement instruments are characterised and calibrated in laboratories where the reference instrument is characterised and traceable to standards (from metrology institutes);
- Measurement protocols are established and agreed within the community;
- Field in situ measurements (on sea, land, etc.) follow protocols and uncertainty budgets are derived;
- Analysis and intercomparison between measurements are carried out;
- Open-access websites/data portals and publications are provided to disseminate results and best practices.

The first bullet of the mandatory characteristics, i.e., traceability, is one of the most critical and relates to evidencing the metrological quality and rigour of the FRM. It is thus important to be clear on the definition of the terminology. Here, 'traceability' and all other related terms such as 'uncertainty' are as defined in the International Organization for Standardization (ISO) guide on vocabulary for international metrology [11]:

Traceability is property of a measurement result whereby the result can be related to a reference through a documented unbroken chain of calibrations, each contributing to the measurement uncertainty.

Note that it is not sufficient to simply purchase an instrument, even with a valid calibration certificate that robustly can evidence traceability to SI through some independent means, but that the use of that instrument in making the measurement can also be fully evidenced. The latter usually involves detailed documentation of the process and some form of independent comparison or calibration.

It should also be noted that in some cases traceability to SI is not necessarily easy to achieve or demonstrate, and in those cases, a community agreed reference or protocol can be used. For example, the use of a Langley plot for calibrating a sun-photometer used for Aerosol Optical Depth (AOD) measurements [12] or the Amazon rain forest for calibration assessment of Synthetic Aperture Radar (SAR) imagers [13] would all be acceptable.

FRMs, whilst ideally developed to accommodate as wide a range of sensors and applications as possible, will always be limited to a particular scope. It is important that the scope (parameter and type of satellite) to which the FRM is applicable is clearly defined and also that any procedure/tailoring of the data needed to match a particular sensor is also well described; for example, an FRM for top-of-atmosphere radiance for a high-resolution optical imager will need to define the spectral and spatial resolution of the FRM data and how it should be convolved to a satellite sensor, including the assessment of the uncertainty associated with that convolution. In many situations the specific details of how this may be done would be on a case-by-case basis, but the limitations and considerations should be transparent to the user.

3.3. FRMs in Practice

One of the first projects that could be considered an FRM (although the terminology did not exist at the time) was dedicated to Sea Surface Temperature (SST) validation; it was performed in Miami, hosted by Miami University, more than 20 years ago. The objective of the campaign was to calibrate and compare infrared radiometers used in the validation of the different sea and land surface temperature products derived from satellites. The objective included an assessment of the relative performance of each instrument as well as ensuring that surface measurements used in satellite product validation were traceable to SI units at the highest possible confidence level. This inter-comparison exercise ensured a coherence between the validation measurements used in various international Cal/Val

programs and provided increased confidence in the results used for validation of satellite-derived surface temperatures.

This first inter-comparison of infrared radiometers was held at Rosenstiel School of Marine, Atmospheric, and Earth Science (RSMAS) in Miami during March 1998. This comparison involved several high-quality radiometers and some off-the-shelf devices. The National Institute of Standards and Technology (NIST) provided their standard blackbody target [14] for calibration of each radiometer. The first exercise was followed by a second in 2001 prior to the launch of Aqua Modis and Envisat, and the third one took place in May 2009, under the auspices of the CEOS WGCV Infrared and Visible Optical Sensors Subgroup (IVOS), prior to the launch of Suomi-NPP. This third comparison followed the newly established QA4EO guidelines, with protocols developed to ensure a robust traceability to SI (via NIST and National Physical Laboratory—NPL) and establish the degree of equivalence between participants. The results were then published [15,16].

Following the creation of the FRM label and criteria, an enhanced version of the previous ‘Miami series’ became FRM4STS—Fiducial Reference Measurements for Surface Temperature Satellites [17]. Started in 2014, the aim was to establish and maintain SI traceability of global FRMs for satellite-derived surface temperature product validation (water and land) and to help develop a case for their long-term sustainability. It required:

- Comparisons to ensure consistency between measurement teams with predefined protocols;
- Accessible common descriptions and evaluation of uncertainties;
- Robust links to SI;
- Experiments to evaluate sources of bias/uncertainty under different operational conditions;
- Guidance for prospective validation teams on efforts needed to establish and maintain their measurement instrumentation in the laboratory and on campaign;
- Traceability of their instruments and also on their use in the laboratory and field.

The CEOS plenary endorsed the project in 2014 with the objective to facilitate international harmonisation.

With the success of FRM4STS [17] as an example, the European Space Agency developed a series of projects, following the same logic, aimed at improving the validation and qualifying the validation measurements as fiducial in various domains (see e.g. [18] in the context of optical remote sensing applied for the generation of marine climate-quality products). Table 1 lists (not exhaustive) some ESA projects with the relevant parameters and satellite missions, and in most cases, their websites provide access to the results and documentation.

In addition to these FRM labelled projects, there are initiatives such as CEOS WGCV RadCalNet [19], and also Hypernets [20], that can both justify the FRM label not only for an individual site but also for a network of sites. In both of these examples, each individual site contributing to the network must adhere to FRM criteria to provide data to a central processing hub. However, in addition, the hub serving as the interface to the community, delivering commonly processed data, is itself also of FRM quality, with documented procedures, uncertainties and comparisons to underpin the evidence. There will be many more examples of networks and/or data sets that meet the FRM criteria, and so some effort is being established to provide an endorsement and cataloguing process.

Although the majority of FRM data sets are likely to be sub-orbital in origin, this is not a prerequisite requirement. Targets such as the moon using models such as GIRO [21], the GSICS Implementation of the Robotic Lunar Observatory (ROLO) model, and LIME [22], the Lunar Irradiance Model of ESA (where the uncertainties and methods are well-established and documented), can be considered FRMs for specific applications, particularly stability monitoring. Similarly, the future SITSat missions such as TRUTHS [23] and Clarreo [24] will provide Top of Atmosphere (ToA), and in the case of TRUTHS, Bottom of Atmosphere (BoA), FRM data of a quality to not only validate but also calibrate other satellite sensors.

Table 1. ESA FRM projects, the parameters they consider and the primary satellite sensor considered. The FRMs are considered to apply to other similar sensors in this class.

FRM Project	Parameter	Indicative ESA Satellite Mission Sensor
FRM4SM	Soil moisture	SMOS/CIMR
FRM4STS	SST, LST	Sentinel-3 SLSTR/CIMR
FRM4SOC	Ocean colour	Sentinel-3 OLCI
FRM4ALT	Altimetry	Sentinel-3 topography, Sentinel-6
St3TART	Water surface height, land and sea ice thickness	Sentinel-3 topography (CRISTAL)
FRM4RADAR	Cloud products	EarthCARE
FRM4DOAS	Tropospheric & stratospheric NO ₂ vertical profiles, total O ₃ column, tropospheric HCHO profile	Sentinel-5P/Sentinel-5/Sentinel-4
FRM4GHG	Under development: aerosol product; advanced urban tropospheric NO ₂ product CO, CH ₄ Characterisation and comparison of small portable FTIR instrument products with established TCCON; CO ₂	Sentinel-5P, Sentinel-5, CO2M
FRM4Veg	Surface reflectance, FAPAR	Sentinel-2, Sentinel-3 OLCI
FRM4AQ (Pandonia FRM)	O ₃ , NO ₂ , SO ₂ , H ₂ O, CH ₄ , CO, HCHO and aerosols	Sentinel-5P/Sentinel-5/Sentinel-4

3.4. Achieving and Evidencing FRM Status

The FRM label is now becoming used internationally, including within the recently published Global Climate Observing System (GCOS) implementation plan [25], and many data providers are seeking to claim that they are FRM compliant. Similarly, satellite operators are requesting that validation data come from an FRM-compliant source. However, whilst the FRM principles are clear, they still necessarily leave room for interpretation, and there is no real authority to stop the label being used in an unjustified or misleading manner even by those using it with the best intentions.

CEOS WGCV has thus recognised the need to establish a means to enable Cal/Val data providers to robustly evidence that they are FRM compliant or, at least, show their progress towards being fully FRM compliant to allow their users (satellite operators and data users) to assess and weight the utility of the Cal/Val data in validating the performance of a particular satellite. To this end, they have added the prefix CEOS-FRM, allowing some ability to control the usage of the label. Recognising that the number of potential CEOS-FRM data providers may become very large, it is necessary that any ‘endorsement’ process needs to be time efficient and low cost but also that it maintains integrity. It also needs to be flexible to address different levels of maturity in the methodologies and products that are being validated as well as technology domains.

It is critical that the characteristics and performance of any CEOS-FRM can be unequivocally considered consistent and trustworthy (within their self-declared metrics) and readily interpretable so that any information derived from it can be reliably utilized by a user. This is not to say that FRM for any particular measurement of a measurand or application must all meet the same level of performance or indeed necessarily fully meet all the CEOS-FRM criteria (at least whilst developing a capability) but that any potential user can readily identify and assess suitability for their application. Ensuring that the CEOS-FRM classification has meaning requires not only a set of common criteria but also obliges some form of independent assessment against them, or at least the potential for such assessment. This could be achieved through the creation of a formal process implemented

though a body such as ISO, where independent assessment/audit against a set of criteria is undertaken by an approved body funded by the entity seeking accreditation.

Whilst aspects of what is entailed within meeting FRM compliance could benefit from and indeed may be most easily achieved through use of services adhering to ISO standards, e.g. ISO 17025 for calibration of instruments, it is not anticipated that the formality and inherent cost associated with such an endorsement process will be necessary or even practical at this time. However, the increased investment required to meet the full criteria of CEOS-FRM needs to be recognised and an appropriate scientific value attributed to it, by funding agencies and in particular those running satellite programs.

The CEOS WGCV framework takes a more pragmatic approach; relying on self-assessment and transparency/accessibility of evidence against a set of criteria where the evidence is then subject to peer review through a board of experts led by CEOS WGCV.

In order to be flexible, to maximise inclusivity and to encourage the development and evolution of CEOS-FRM from new or existing teams, compliance with criteria will be based on a gradation scaling rather than a simple pass/fail. The degree of compliance and associated gradation can then be presented in a Maturity Matrix (MM) model, similar to that of EDAP (Earthnet Data Assessment Project) [7], to allow intended users of the FRM to assess suitability for their particular application and indeed funders to decide on where and what aspects to focus any investment. The matrix model provides a visual 'simple' assessment of the state of any FRM for all given criteria, making visible where it is mature and where evolution and effort need to be expended.

The CEOS-FRM Maturity Matrix provides a high-level colour-coded summary of the characteristics of the FRM under analysis against specific criteria. Although the criteria are intended to be generic, how they are interpreted and assessed in detail may vary with the nature of the FRM and sensor class. It is intended that this same concept can be used for an individual measurement/method at a geo-located site, e.g., a test site, as well as at a network of such similar sites. The matrix contains a column for each 'high level' category/criterion, i.e., the row: Nature of FRM, FRM Instrumentation etc., and cells for each aspect (subsection of category) in a column below each of these categories. Grades are indicated by the colour of the respective grid cell, which are defined in the key (the grade criteria). The CEOS-FRM Maturity Matrix can be further subdivided into two parts as follows:

- Self assessment CEOS-FRM MM
- Verification CEOS-FRM MM

An example of such a maturity matrix is provided in Tables 2 and 3 for a fictitious FRM provider; note that the exact labels for each category in the MM is subject to change at the time of publication due to ongoing review and should be considered indicative.

In addition to this broad-based summary, an overall classification of the degree of compliance will be provided based on meeting specific gradations for particular criteria. At present there are four classifications (note the number and definitions for these classifications is subject to change as CEOS finalizes the framework):

- Class A—Where the FRM fully meets all the criteria necessary to be considered a CEOS-FRM for a particular class of instrument and measurand. It should achieve a grade of 'Ideal' in the 'guidelines adherence' in the verification section of the MM and green (at least 'excellent') for all other verification categories where these have been carried out.
- Class B—Where the FRM meets many of the key criteria and has a path towards meeting the Class A status in the near term. It should achieve at least 'Excellent' in guideline adherence in the verification section of the MM and green (at least 'excellent') for all other verification aspects where these have been carried out. Ideally, it should have a roadmap/timeline of activities being undertaken to reach the higher class.
- Class C—Meets or has some clear path towards achieving the criteria needed to reach a higher class and provides some demonstrable clear value to the validation of a class of satellite instruments/measurands. It should achieve at least 'Good' in the guideline adherence in the verification section of the MM and at least 'good' for all

- other verification categories where these have been carried out. Ideally, it should have a roadmap/timeline of activities being undertaken to reach the higher class.
- Class D—Is a relatively basic adherence to the CEOS-FRM criteria. Needs to have a clear strategy and aspiration to progress towards a higher class. This can be considered an entry level class for those starting out on developing an FRM. It should achieve at least ‘Basic’ in the guideline adherence in the verification section of the MM and at least ‘Good’ for all other verification categories where these have been carried out. FRM owners/developers must have a roadmap/timeline of activities being undertaken to reach the higher class.

Table 2. Suggested format of a maturity matrix for an FRM product for a fictitious FRM provider. Colours are defined in Table 3.



Self-Assessment					Independent Assessor
Nature of FRM	FRM Instrumentation	Operations/Sampling	Data	Metrology	Verification
Descriptor	Instrument documentation	Automation level	Data completeness	Uncertainty characterisation	Guidelines adherence
Location/availability of FRM	Evidence of traceable calibration	Measurand sampling/representativeness	Availability and usability	Traceability Documentation	Utilization/feedback
Range of instruments	Maintenance plan	ATBDs on processing: algorithms/software	Data Format	Comparison/calibration of FRM	Metrology verification
Complementary observations	Operator expertise	Guidelines on transformation to satellite pixel	Ancillary Data 	Adequacy for intended class of instrument/measurand	Independent verification
FRM CLASSIFICATION					A B C D (to be selected)

Table 3. Grading criteria for the maturity matrix in Table 2.

Grade
Not assessed
Not assessable
Basic
Good
Excellent
Ideal
Not public 

4. Conclusions

In this paper, we define criteria and describe examples of a ‘new’ category, at least in terms of formalized labels, of Earth observation data expressly designed and characterised to meet the Cal/Val needs of the world’s satellite sensors in a coordinated cost-efficient manner. These FRM data sets are typically derived as a subset of a broad-based set of similar data which have been tailored so that they can provide data representative of that measured by a particular class of satellite sensors and that have well-defined and documented estimates of their uncertainty to international references such that they can be unequivocally relied upon for satellite Cal/Val.

FRMs, and in particular CEOS-FRMs, will provide the means to facilitate confidence and harmonisation not only in the data delivered from national and international space agency missions but also ‘new space’ commercial providers seeking to evidence the quality of their data. The criteria needed to fully achieve FRM status are detailed, as is the outline of a new endorsement process currently being developed for implementation by CEOS. It is anticipated that many post-launch Cal/Val data providers will seek to achieve full CEOS-FRM status and, through the different classes, allow an evolutionary development with evidence of progress towards CEOS-FRM to avoid barriers to uptake and encourage the development of new FRMs.

Author Contributions: Conceptualization, C.D., P.G. and N.F.; methodology, N.F., P.C. and P.G.; writing—original draft preparation, N.F., P.G. and C.D.; writing—review and editing, N.F., P.C. and P.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: No new data were created for this article.

Acknowledgments: The FRM principles formed the basis for an on-going series of projects funded by European organisations: EU, ESA, EUMETSAT. We acknowledge the support from these organisations that led to a “joint vision by ESA and EUMETSAT for Fiducial Reference Measurements” that was presented at European Commission—ESA—EUMETSAT Trilateral meeting on 5 June 2020. The authors would also like to thank Emma Woolliams for useful discussion and comments on the drafting of this paper.

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

References

1. CEOS Cal/Val Portal. Available online: <https://calvalportal.ceos.org/> (accessed on 16 October 2023).
2. Radiometric Calibration Network (RadCalNet) Portal. Available online: <https://www.radcalnet.org/> (accessed on 16 October 2023).
3. Quality Assurance Framework for Earth Observation. Available online: <https://qa4eo.org/> (accessed on 16 October 2023).
4. Quality Assurance Framework for Earth Observation—QA4EO Tools. Available online: <https://qa4eo.org/tools.php> (accessed on 16 October 2023).
5. The International System of Units (SI). Available online: <https://www.bipm.org/en/measurement-units> (accessed on 16 October 2023).
6. Mises-en-Pratique. Available online: <https://www.bipm.org/en/publications/mises-en-pratique> (accessed on 16 October 2023).
7. The Earthnet Data Assessment Project (EDAP+). Available online: <https://earth.esa.int/eogateway/activities/edap> (accessed on 16 October 2023).
8. Donlon, C.; Goryl, P. Fiducial Reference Measurements (FRM) for Sentinel-3. In Proceedings of the Sentinel-3 Validation Team (S3VT) Meeting, ESA/ESRIN, Frascati, Italy, 26–29 November 2013.
9. The Copernicus In-Situ Component. Available online: <https://insitu.copernicus.eu/> (accessed on 16 October 2023).
10. Boesch, H.; Brindley, H.; Carminati, F.; Fox, N.; Helder, D.; Hewison, T.; Houtz, D.; Hunt, S.; Kopp, G.; Mlynyczak, M.; et al. *SI-Traceable Space-Based Climate Observation System: A CEOS and GSICS Workshop, National Physical Laboratory, London, UK, 9–11 September 2019*; NPL: Teddington, UK, 2019. [CrossRef]
11. BIPM. *International Vocabulary of Metrology*, 3rd ed.; VIM Definitions with Informative Annotations; BIPM: Paris, France, 2017; Available online: <https://www.bipm.org> (accessed on 16 October 2023).
12. Holben, B.N.; Eck, T.F.; Slutsker, I.; Tanre, D.; Buis, J.P.; Setzer, A.; Vermote, E.; Reagan, J.A.; Kaufman, Y.; Nakajima, T.; et al. AERONET—A federated instrument network and data archive for aerosol characterization. *Remote Sens. Environ.* **1998**, *66*, 1–16. [CrossRef]
13. Amazon Rain Forest Site for SAR Calibration. Available online: <https://calvalportal.ceos.org/web/guest/amazon-rain-forest-sites> (accessed on 16 October 2023).
14. Fowler, J.B. A Third Generation Water Bath Blackbody Source. *J. Res. Natl. Inst. Stand. Technol.* **1995**, *5*, 591–599. [CrossRef] [PubMed]
15. Rice, J.P.; Butler, J.J.; Johnson, B.C.; Minnett, P.J.; Maillet, K.A.; Nightingale, T.J.; Hook, S.J.; Abtahi, A.; Donlon, C.J.; Barton, I.J. The Miami, 2001 Infrared Radiometer Calibration and Intercomparison: 1. Laboratory Characterization of Blackbody Targets. *J. Atmos. Ocean. Technol.* **2004**, *21*, 258–267. [CrossRef]
16. Theocharous, E.; Fox, N.P. CEOS Comparison of IR Brightness Temperature Measurements in Support of Satellite Validation. Part II: Laboratory Comparison of the Brightness Temperature of Blackbodies. Available online: <https://eprintspublications.npl.co.uk/4759/1/OP4.pdf> (accessed on 16 October 2023).
17. Fiducial Reference Measurements for Satellite Temperature Product Validation. Available online: <https://www.frm4sts.org/> (accessed on 16 October 2023).
18. Zibordi, G.; Donlon, C.; Albert, P. *Optical Radiometry for Ocean Climate Measurements*, 1st ed.; Elsevier: Amsterdam, The Netherlands, 2014.
19. Bouvet, M.; Thome, K.; Berthelot, B.; Bialek, A.; Czaplak-Myers, J.; Fox, N.P.; 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]
20. Goyens, C.; De Vis, P.; Hunt, S. Automated Generation of Hyperspectral Fiducial Reference Measurements of Water and Land Surface Reflectance for the Hypernets Networks. In Proceedings of the 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, Belgium, Brussels, 11–16 July 2021. [CrossRef]
21. Lunar Calibration Algorithm Working Area. Available online: <http://gsics.atmos.umd.edu/bin/view/Development/LunarWorkArea> (accessed on 16 October 2023).
22. Lunar Irradiance Model ESA: LIME. Available online: <https://calvalportal.ceos.org/lime> (accessed on 16 October 2023).

23. 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](#)]
24. Wielicki, B.A.; Young, D.F.; Mlynchak, 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. In *BAMS (Bulletin of the American Meteorological Society)*; American Meteorological Society: Boston, MA, USA, 2013; Volume 94, pp. 1519–1539. [[CrossRef](#)]
25. The 2022 GCOS Implementation Plan (GCOS-244). Available online: https://library.wmo.int/index.php?lvl=notice_display&id=22134 (accessed on 16 October 2023).

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.