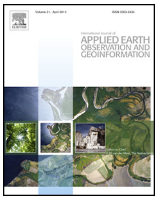
International Journal of Applied Earth Observation and Geoinformation 42 (2015) 150–161

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

journal homepage: www.elsevier.com/locate/jag



Analysis of current validation practices in Europe for space-based climate data records of essential climate variables

Y. Zeng[a,*], Z. Su[a], J.-C. Calvet[c], T. Manninen[e], E. Swinnen[f], J. Schulz[b], R. Roebeling[b], P. Poli[d], D. Tan[d], A. Riihelä[e], C.-M. Tanis[e], A.-N. Arslan[e], A. Obregon[g], A. Kaiser-Weiss[g], V.O. John[b], W. Timmermans[a], J. Timmermans[a], F. Kaspar[g], H. Gregow[e], A.-L. Barbu[c], D. Fairbairn[c], E. Gelati[c], C. Meurey[c]

[a] WRS Department, Faculty for Geo-information Science and Earth Observation, University of Twente, Netherlands
[b] European Organisation for the exploitation of meteorological satellites (EUMETSAT), Germany
[c] Centre National de Recherches Météorologiques (CNRM-GAME), UMR3589 (Meteo-France, CNRM), Toulouse, France
[d] European Centre for Medium-Range Weather Forecasts (ECMWF), UK
[e] Finnish Meteorological Institute (FMI), Finland
[f] Flemish Institute for Technological Research (VITO), Germany
[g] German Weather Service (DWD), Germany

ARTICLE INFO

Article history:
Received 24 March 2015
Accepted 15 June 2015
Available online 7 July 2015

ABSTRACT

The Climate Data Records (CDRs) of Essential Climate Variables (ECVs) that are based on satellite observations need to be precisely described. In particular, when these products are delivered to end-users, the error characteristics information and how this information is obtained (e.g., through a validation process) need to be documented. Such validation information is intended to help end-users understanding to what extent the product is suitable for their specific applications. Based on how different European initiatives approached the validation of CDR and ECV products, we reviewed several aspects of the current validation practices. Based on the analysis of current practices, essentials of validation are discussed. A generic validation process is subsequently proposed, together with a quality indicator.

1. Introduction

Both Global Climate Observation System (GCOS) climate monitoring principles (GCOS, 2003) and the USGCRP (U.S. Global Change Research Program) principles (USGCRP, 2003) highlight the important role of calibration and validation (hereafter, Cal/Val) in producing climate quality data from space. As most Cal/Val results concern short-term data records, it is important to define a realistic generic validation strategy for (long-term) climate data records (CDRs), derived from the existing validation practices. This is illustrated by the architecture for climate monitoring from space proposed by (Dowell et al., 2013): “Climate record processing requires a sustained expert understanding of both new and legacy climate sensors as well as a sustained web of support activities, including a significant effort on Cal/Val; research to reduce uncertainties, establish ‘community reference standards’; and collaborative product assessment and inter-comparison.”

Cal/Val activities are integral components of the Fundamental Climate Data Records (FCDRs) and Thematic Climate Data Records (TCDRs) processing chain, as illustrated in Fig. 1. The FCDR refers to a long-term data record of calibrated and quality-controlled sensor data designed to allow the generation of consistent products that are accurate and stable enough for climate monitoring (NRC, 2004). FCDRs are typically calibrated radiances, backscatter of active instruments or radio occultation bending angles. FCDRs also include the ancillary data used to calibrate them. The TCDR denotes a long-term data record of validated and quality-controlled geophysical variables derived from FCDRs (NRC, 2004).

Fig. 1 shows that processing starts with the availability of observations (e.g., raw sensor data). These observations are then calibrated, geolocated, and corrected for perturbing factors (e.g., atmospheric effects) to generate FCDR (or Level-1) products (e.g., radar backscatter or radiometer brightness temperature), which are then used to produce TCDRs (i.e., geophysical and biogeophysical variables) that are subsequently validated to check if GCOS requirements are met. Both TCDRs and FCDRs are then archived, together with relevant metadata. It is to note that the processes depicted in Fig. 1 are recursive. The observations are

---

* Corresponding author.

http://dx.doi.org/10.1016/j.jag.2015.06.006

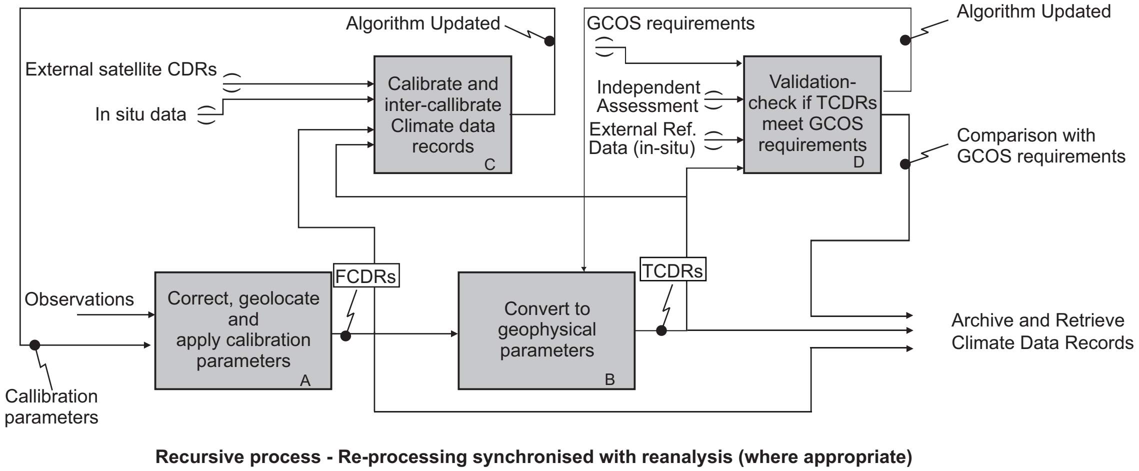


Fig. 1. Processing Chain of FCDRs and TCDRs. (Adapted from Fig 6.3 in Dowell et al., 2013).

reprocessed to generate improved FCDRs/TCDRs when improved information or algorithms become available. In fact, the complete chain of generating climate data records also includes the processes related to the peer-review of a new data record, assessments of data records, and the responses to user feedbacks (Dowell et al., 2013). The calibration ("C" in Fig. 1) is a process quantitatively defining the system response to known, controlled signal inputs (NRC, 2004; Xiong et al., 2010; Chander et al., 2013) have conducted a comprehensive overview of the current practice on the (inter) calibration of satellite sensors.

In this study, we focus on the validation ("D" in Fig. 1) of CDRs of ECVs (Essential Climate Variables). In practice, the validation approaches may vary from one application to another. For example, for weather forecast, the validation of certain variables does not need to consider time series of sufficient length, consistency, and continuity (e.g., the data collection and quality control approaches at different observation networks are not necessary coordinated and harmonized) (Estévez et al., 2011). On the other hand, such conditions are required to assess the climate variability and change. From discussions with TCDR users and data providers (Su et al., 2013a,b), a number of recommendations were derived for the implementation of validation: (1) a traceable validation documentation, (2) an independent review mechanism, (3) regular updates of validations, and (4) analysis of the factors generating uncertainties in CDRs.

The objective of this paper is to assess how different European initiatives/services approach the validation of ECV CDRs. The validation process will differ from ECV to ECV, and individual ECV production teams have already developed specific validation processes for their particular ECVs (ESA, 2010). It, therefore, comes to a point that a transparent, traceable validation process should be documented. In the following Section 2, the aspects of the validation process are discussed. In Section 3, the essentials of validation are discussed, after analyzing the current validation practice with some examples. In Section 4, a generic validation process is proposed. A set of quality indicators (a system maturity matrix) is introduced in Section 5, to facilitate the benchmarking of validation processes. A demonstration on how to assess a validation process using the quality indicator is presented. Conclusions and recommendations are drawn in Section 6.

2. Product validation

The Committee on Earth Observing Satellites (CEOS) working group for Cal/Val (WGCV) defines validation as the process of assessing, by independent means, the quality of the data products derived from satellite observations. This can be called product validation. The product validation ensures that the quality of the products is properly assessed, through quantification of the uncertainties in both the data itself and the measurement system deployed for generating the data. It includes a quantitative understanding and characterization of the measurement system and its bias in time and space. In this context, validation can be considered a process that encompasses the entire system, from sensor to product.

2.1. Validation concept

Fig. 2 shows two typical validation concepts: the scaling method and the direct comparison method. Fig. 2 can be regarded as an elaboration of the validation process given by component "D" in Fig. 1.

The scaling method is shown in the left panel of Fig. 2. The scaling method uses an intermediate Very High Resolution (VHR) satellite data layer (or airborne campaign data) to compare the ground measurements with products at coarser spatial resolution. This permits reducing the uncertainties and the difficulties during the integration of several punctual ground measurements over a common area (or an Elementary Surface Area, ESU) to be used for the validation of the product at a pixel level. This is valid especially for products around 100m of resolution or more, for which it is very difficult to integrate several measurements to reach an ESU of that dimension also taking into consideration the landscape heterogeneity. This is the case for most of the terrestrial ECV CDRs (e.g., land use, LAI).

The consolidated and qualified campaign data (yellow) boxes in Fig. 2 indicate the elements needed for validation, including the satellite data, the ancillary/auxiliary data and ground (reference) measurements. The left (green) boxes represent the FCDR processing and TCDR generation (the retrieved quantity or the retrieval) while the right (blue) ones are the processing of campaign data to produce VHR reference validation layers ("true quantity"). In the FCDR processing, raw satellite measurements are geolocated and atmospherically corrected first when relevant and then homogenized and inter-calibrated, to generate calibrated radiances, backscatter of active instruments or radio occultation bending angles. Afterwards, by means of data assimilation or modeling (e.g., radiative transfer model or specific retrieval algorithms), TCDR products are retrieved. In the right (blue) boxes, level 1 data are used to derive level 2 data products at very high resolution. Afterwards, the ground measurements are processed and "transferred" directly to the level 2 data product to represent the validation layers. Finally in the bottom (red) box, the TCDR products

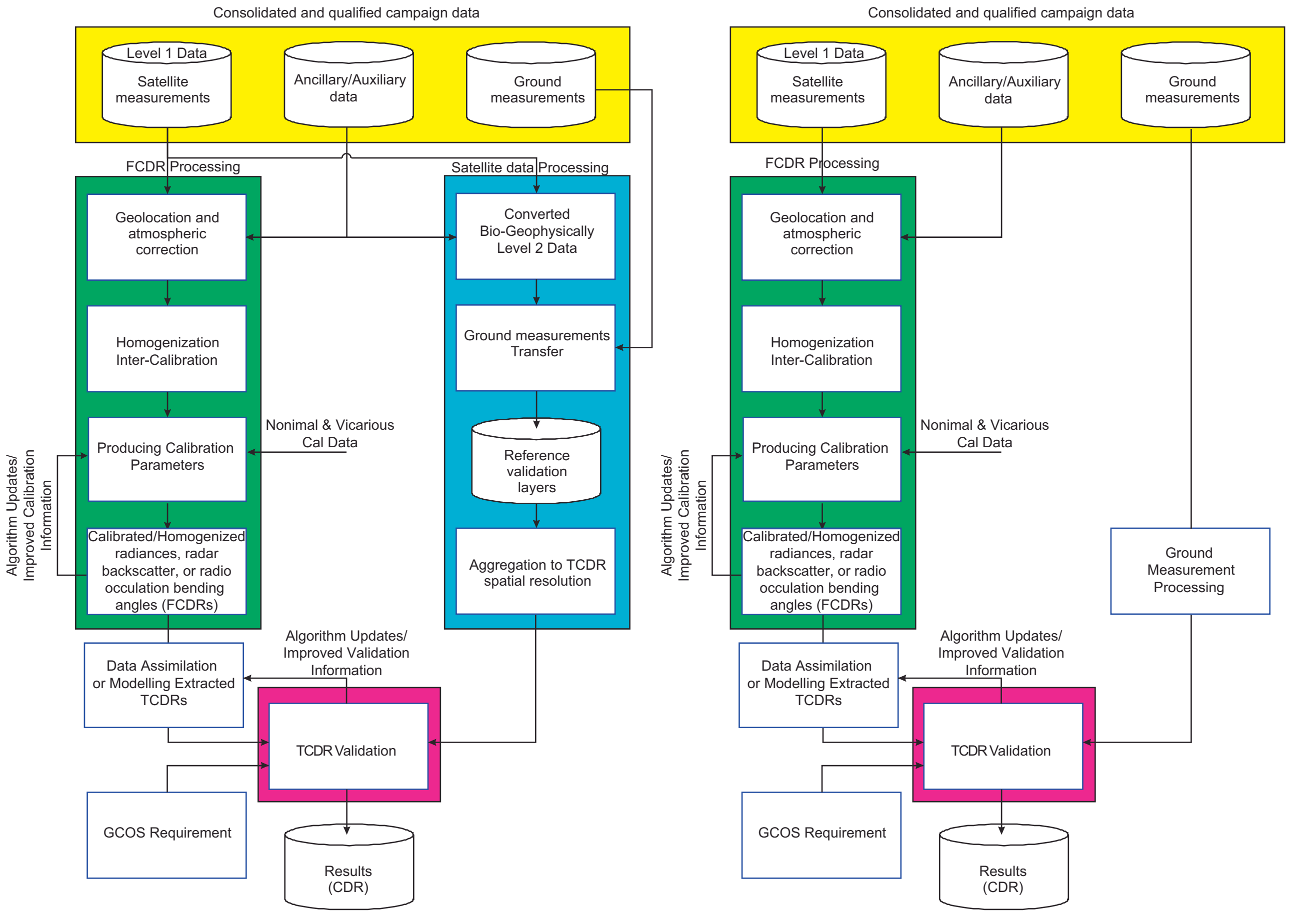


Fig. 2. Validation of ECV CDRs as an expansion of the validation function "D" in Fig. 1 showing two typical direct validation concepts, the scaling method (left) and the direct comparison method (right).

are validated by the use of the previously generated VHR reference validation layers.

It is noted that the generations of both FCDRs and TCDRs are recursive because when improved information becomes available (e.g., better algorithms or improved calibration/validation information) the observations are re-processed to generate improved CDRs. As understanding of sensor calibration issues and as the modeling of the radiative transfer from the Earth and atmosphere improves, products can be generated via reprocessing.

The main advantage of the scaling scheme is the fact that these intermediate layers are very close in terms of quality and resolution to the ground measurements (since they are obtained by VHR data), thus the uncertainties due to the ESU integration can be reduced because now the integration is applied to surfaces at the same spatial scale as the satellite pixels. One should realize that the optimal ESU size is determined by the level of within ESU variability that can be tolerated by the validation protocol and the effort available to conduct measures. The size of ESU within a reference region can vary with various factors (e.g. surface condition, instrument field of view and spatial sampling design etc.) (Fernandes et al., 2014). Nevertheless, as shown in the blue column of Fig. 2, the level 2 data retrieval, the transfer processing, and the aggregation introduce uncertainties that have to be monitored.

In the direct comparison method, the ground measurements are directly compared with the TCDR product retrieved for standard processing. The elaboration needed is the processing/integration of ground measurements to generate ESU comparable with the TCDR product. This method is more "direct" with respect to the other, but the ground measurement processing for the ESU generation can introduce additional uncertainties that again have to be estimated and monitored. Moreover, this processing is applicable if the campaign ground measurements have followed a specific protocol and the area to be covered by the ESU is comparable in size with the product pixel.

2.2. Aspects of validation

From the description of validation concept for ECV CDRs, there are at least three components needed for implementing a validation study (i.e., the purple–red highlighted TCDR box in Fig. 2): the validation requirements, the generated TCDR itself, and the established reference data. In addition, one needs to select methods to implement validation. In the following, each of these four components is briefly described.

2.2.1. Validation requirements and validation strategy

The GCOS, 2010 update (GCOS, 2010b) detailed the requirements for ECVs in atmospheric, oceanic and terrestrial domains. It is assumed that maximum benefit of the datasets and derived products for climate applications will be gained if the requirements are met. In this case, the validation requirements should be capable of testing products for compliance with GCOS requirements. Accordingly, the Land Product Validation (LPV) subgroup of CEOS WGCV identified a hierarchical approach to classifying land product validation stages (CEOS-WGCV-LPV, 2009). The Group for High-Resolution Satellite sea-surface temperature (SST) has developed a community-accepted validation strategy to validate the combined SST satellite products (ST-VAL 2008; Beggs et al., 2012). The validation requirements drive the definition of the validation strategy, which should define the process and activities at each stage of validation (Fetterolf, 2007; CEOS-WGCV-LPV, 2009).

2.2.2. *TCDR Generation*

The generation of TCDRs needs to combine data from a variety of sources (space and in-situ) and emerging products from data assimilation. In order to assess the error or uncertainty that may arise from various sources, one should be able to trace back how the TCDR was generated and validated (i.e., traceability) (Dowell et al., 2013). Traceability is defined as the property of a measurement result which can be related to a reference through a documented unbroken chain of validations, each contributing to assessing the measurement uncertainty (Ellison et al., 2000).

2.2.3. *Selection of validation datasets*

Independent reference datasets are needed to validate the product. It is important to use quality criteria for the selection and production of reference datasets for validation, in order to document a traceable validation process. The choice of validation datasets is heavily dependent on the validation requirements and the existing validation capacities. For example, WMO has coordinated different types of in-situ networks for different purposes: (i) the global reference-observing networks, (ii) the global baseline observing networks, and (iii) the comprehensive observing networks. For each kind of network, the detailed processes, guidelines and manuals are documented to assist end-users to choose different networks for their particular purpose (GCOS, 2003; GCOS, 2010a). Accordingly, there are few key questions to be answered for this validation aspect: (a) how to get representative reference observations? (b) how representative is the reference data? (c) how to ascertain independence of the reference data?

2.2.4. *Validation method and uncertainties*

As discussed in Section 2.1, one can choose the scaling or direct method for implementing the validation, which depends on the validation requirement and the validation capacity. The selection of a validation method also depends on what are the parameters to be validated. For example, one may encounter question on how to validate products for which no reference observations are available or how to validate products for which multiple reference observations exist. Nevertheless, the validation method should be selected or designed to identify possible error sources and related uncertainties. It is expected that the validation can help investigating a number of quality parameters, including at least the random and systematic error and uncertainty, bias, and stability (Ellison et al., 2000; JGCM 2008; ESA, 2010). For the validation of climate data products, the questions that are relevant to this validation aspect include: (a) what is the required scale (regional, continental or global) ? (b) what is the primary aim (e.g., the long-term trend, stability or consistency) ? (c) are there review mechanisms and sustaining mechanisms for the validation?

3. Analysis of current validation practices

In the following, current validation practice examples from different European initiatives are introduced, following the validation aspects introduced in Section 2.2. For some initiatives or projects, a specific product is used as an example to demonstrate the validation practices. It is not intended to implement an exhaustive review of the current validation practices, but to show some typical examples to identify or abstract the essentials of validation (i.e., drawn in Section 3.3.).

3.1. *Contrasting validation strategies*

Table 1 lists three examples illustrating various validation aspects presented in Section 2.2. The three examples include the Fraction of Absorbed Photosynthetically Active Radiation (FAPAR) product of the Copernicus Global Land Service (CGLS), the Copernicus – MyOcean II Baltic sea regional Sea Ice (SI) thickness and concentration, and the ESA-CCI products. The first two examples show how different space-based products may have different validation processes while the last example shows how a generic validation strategy can be defined as implementing the validation processes for different products.

For the first example, the CGLS FAPAR product, the validation protocol includes in situ data collection and scaling, as well as indirect validation based on the inter-comparison against other available global FAPAR data sets and a cross-cutting quality monitoring technique based on data assimilation. The validation strategy for Baltic regional sea ice product is relying on the direct comparison with in situ data only, following WMO guidelines and standards (WMO, 2010; Smolyanitsky, 2012). The difference between the first two examples lies primarily in the richness of available validation data. The validation of FAPAR products relies on various sources (e.g., see the selection of validation datasets in Table 1). The spatial resolutions of FAPAR products are typically in the range of 1–2 km, and the temporal resolution ranges from daily to weekly, 10-daily or monthly. The currently existing discrepancies among different FAPAR products are mainly due to differences in concepts and definitions, retrieval algorithms or input data quality. The validation includes the definition of internationally accepted definitions and associated quantities, the use of standard spatial sampling and in situ measurement methods, as well as the use of standard reporting on statistical comparisons (GCOS, 2010b; Fernandes et al., 2014).

For Baltic regional sea ice, the archive of satellite data for sea-ice thickness available to date remains limited. Few in situ observations obtained from ships and a few coastal stations are routinely available for satellite sea ice charting, but the spatial coverage of precise ice thickness values needed for product validation (GCOS, 2010b) is limited. It is to note that the validation capacity for SI (especially for sea ice concentration) is now being extended to more planned dedicated satellite missions and the merge of historical satellite observations (Kwok 2010; NCAR, 2013). Challenges for the validation of SI products include the impact of snow cover on SI, as it influences the apparent ice emissivity and backscatter and cause errors in estimated ice concentrations. The lack of knowledge on snow load on the ice causes errors in estimated ice thicknesses from satellite observations. This partially constrains the validation activities for SI, which focus on meeting the validation requirements listed by GCOS, 2010 update (GCOS, 2010b).

For ESA-CCI, the main aspects of the validation process are kept as generic as possible across ECV products (ESA, 2010). The independent assessment and sustaining mechanisms (see validation aspects of ESA-CCI in Table 1) aim at ensuring the independence of the validation. In fact, it suggests all CCI projects should use the definition of validation approved by the CEOS-WGCV, and adhere to the independence requirement for validation. It is suggested to use established, community accepted, traceable validation protocols where they exist, select external datasets for validation and fully describe the validation process (ESA, 2010).

Although the validation strategy is driven by the validation requirements, it is also constrained by the validation capacity. While the validation of FAPAR in the CGLS is based on various reference datasets, the SI validation relies on few available reference datasets. The ESA-CCI example illustrates different validation aspects based on GCOS requirements (ESA, 2010; Bojinski, 2011). The ESA-CCI validation aims to reach transparency and traceability

Table 1
Examples of validation practices, analyzed with the five validation aspects.

Val. Aspects	1. Validation strategies	2. TCDR generation	3. Selection of validation datasets	4. Validation method and uncertainty
Examples				
Copernicus Global Land Service: FAPAR	–Protocol on the validation methods; –Protocol on in-situ data collection; –Guidelines for upscaling in-situ data; –Self-assessment validation; –Periodical independent assessment; –6-monthly quality monitoring reports	The FAPAR TCDR is based on normalized surface reflectance for a standard observational configuration.	Reference global biophysical products: –Inter-comparison: CYCLOPES (CYCV31)[a], MODIS (MODC5)[b], JRC FAPAR[c] and GEOLAND V0 (GEOV0)[d]; –Direct comparison: scaled local in-situ data from e.g. VALERI and DIRECT sites[e].	The validation is done relative to the user requirements of the Copernicus GLS[f]. Apart from direct validation, indirect validation is performed: –Inter-comparison: spatial consistency analysis, global statistical analysis, temporal consistency analysis; –Cross-cutting quality monitoring using a Land Data Assimilation System (LDAS); The FAPAR product is delivered with two quality indicator layers, i.e., the quality flag and the uncertainty estimate.
Copernicus-MyOcean II Baltic Sea Regional Sea Ice	Validation relies on the direct comparison with reference data for all three datasets: –ice concentration, –ice thickness; –ice drift.	–Ice thickness generation is based on the thickness history and SAR images. New ice concentration is produced as soon as new SAR image is available; –Ice drift is produced with using two SAR images with a time gap of less than three days, computing phase correlation if pairwise data windows sampled; –Ice concentration is based on the digitized ice charts produced by ice analysts.	–Drill measurements from icebreakers in Baltic Sea are used for ice thickness validation; –Ice drift reference data is collected with buoys in Gulf of Bothnia and Gulf of Finland. –Ice concentration reference data is produced with ASI algorithm by University of Bremeng[g] with SSMI/S data;	–For different products, the specified validation process and methods were predefined to do the statistical analysis of the comparison between the product and the reference data collected. –For ice thickness and concentration, difference values calculated and their presentations of distribution graphs and scatter plots are given. –Ice drift error is reported with the accuracy of estimated vectors and their presentation of scatter plots. Additional quality flags as indexes are included to datasets.
ESA-CCI	It includes two levels of activities: –Validation and error characterization internal to each CCI project; –Externally, ECV data product assessment through the Climate Modelers User Group (CMUG);	–Develop and validate algorithms to meet GCOS ECV requirements for (consistent, stable, error-characterized) global satellite data products from multi-sensor data archives; –Produce and validate, within an R&D context, the most complete and consistent possible time series of multi-sensor global satellite data products for climate research and modelling.	It includes requirements as listed below: –For validation, CCI project teams should use in-situ or suitable reference datasets that have not been used during the production of their CCI products; –The independence of the geophysical process should be considered.	–It should ensure that the validation is carried out (or at least verified) by staff not involved in the final algorithm selection; –It should use established, community accepted, traceable validation protocols where they exist; –A validation process should be an ongoing process that takes into account requirements and responses from users and should be fully documented in the Product Validation Plan. –To assure the quality of an ECV data product, and that the product specifications are reached, uncertainty of the product should be fully documented in the Comprehensive Error Characterization Report.

(a) (Baret et al., 2007); (b) (Myneni et al., 2002); (c) (Pinty et al., 2011); (d) (Baret et al., 2013); (e) (Garrigues et al., 2008); (f) http://land.copernicus.eu/global/; (g) (Spreen et al., 2008) to enable scientific judgment and user acceptance (Bojinski, 2011), emphasizing the importance of independence in validation.

3.2. Indirect validation methods

The direct validation of ECV CDR is not straightforward, as in-situ observations are limited in space and for some parameters also in time, and as direct validation may require the scaling of the local in situ observations (Camacho et al., 2013). In addition, the locations of in situ data are often biased to easy accessibility or certain land cover type. In practice, indirect validation is used together with direct validation. Indirect validation consists in comparing the product with other products. In addition to the inter-comparison method (Camacho et al., 2013), the indirect validation may include triple collocation (Stoffelen 1998; Scipal et al., 2008; Yilmaz et al., 2012) and cross-cutting quality monitoring through data assimilation (Barbu et al., 2014).

3.2.1. Inter-comparison

The inter-comparison among several independent products (e.g., produced by different algorithms or observing systems) can increase confidence in the reliability of the identified statistical parameters. For example, the comparison of decadal global water vapor changes derived from two independent satellite time series shows that the trends of the two water vapor datasets are similar (Mieruch et al., 2014). Such inter-comparison indicates the sufficient stability in both satellite observations. The inter-comparison can also identify the systematic uncertainty. The comparison of monthly means of two global water vapor products shows that the systematic bias exists between the two datasets and the bias contains a pronounced seasonal component (Mieruch et al., 2010). It indicates that the accuracy and precision of a product are not static and vary seasonally. This is also true for other ECV CDRs, especially when their long term mean have seasonal variation (Entekhabi et al., 2010).

The thorough inter-comparison can also help to identify the gap in the current observation or modelling system. Through the inter-comparison of various global biophysical variable products, Camacho et al. (2013) identified the highest discrepancies and lowest correlation between products for the evergreen broadleaf forest and the needle-leaf forest where contamination by cloud or snow limits the reliability of the reflectance values used as inputs in the algorithms. Su et al. (2013a) inter-compared soil moisture analyses with in-situ observations and found that there is a seasonal shift of bias which may be induced by the decoupling between moisture and heat transport during freezing/thawing in land surface models.

Systematic errors can be generated by differences in sampling, retrieval algorithms, observation time and spatial resolution (Mieruch et al., 2010). In order to identify these errors, one may follow the approach that the CEOS WGCV land product validation subgroup has proposed for the validation of leaf area index (LAI) products (Fernandes et al., 2014). Validation statistics (e.g., measurement uncertainty, precision, and completeness) can be used to investigate issues related to the spatial and temporal continuity and consistency. In particular, temporal aspects include seasonality and temporal smoothness (Camacho et al., 2013; Fernandes et al., 2014). A concise overview of statistical methods commonly used in climate research can be found in (Hennemuth et al., 2013) and (WMO, 2011).

3.2.2. Triple collocation

The triple collocation (TC) technique has been developed to use three collocated data sets for jointly providing sufficient constraints determining the error variance to characterize the uncertainties (Stoffelen, 1998). It can help to identify individual relative error structure of in situ, remote sensing and reanalysis datasets (Su et al., 2014). In addition, more than three data sets can be used (multiple collocations). The TC technique has been now widely used in characterizing uncertainties in FAPAR (D'Odorico et al., 2014), sea surface temperature (Gentemann, 2014), soil moisture (Scipal et al., 2008; Yilmaz et al., 2012) and precipitation (Roebeling et al., 2012) products.

3.2.3. Cross-cutting quality monitoring

The most advanced cross-cutting quality monitoring technique consists in integrating satellite-derived products into a land surface model through a land data assimilation system (LDAS). The reanalysis produced by the LDAS accounts for the synergies of the various upstream products and provides statistics, which can be used to monitor the quality of the assimilated observations (Barbu et al., 2014) (e.g., provided that the input data, boundary conditions and model physics were kept unchanged). In the CGLS, the cross-cutting quality monitoring consists in assimilating the GEOV1 LAI and the ASCAT-derived surface soil moisture (SSM) products over France in the ISBA-A-gs land surface model (at 8 km × 8 km resolution). The resulting LDAS-France chain produces analyzed values of LAI and SSM, and permits the passive monitoring of other products of the CGLS such as FAPAR, surface albedo (SA), and land surface temperature (LST) (Calvet et al., 2014). Fig. 3 shows that the difference between the analyzed FAPAR from January to May 2014 was similar to the difference observed from 2007 to 2013. On the other hand, in June 2014, the analyzed FAPAR was higher than the observations, by about 0.5 on average while it was generally unbiased in June. This denoted a problem caused by the transition from SPOT-VGT (Raymaekers et al., 2014) to PROBA-V (Wolters et al., 2014) on 13 May 2014, and the CGLS PROBA-V processing chain had to be revised. Similar comparisons were implemented to monitor the quality of LAI, SSM, SA, and LST.

3.3. Essentials of validation

Based on the analysis of current validation practices (Sections 3.1 and 3.2), it appears that the validation of ECV CDRs implies the establishment of reference data, which should be kept as independent as possible. The first essential of a validation process is therefore:

A The characterization of the reference data

It is important to document how reference data is produced, especially for the satellite-based reference data. For example, for indirect validation, different satellite products will be compared with each other. The first thing to check is whether different satellite sensors measure the same physical quantity. Without the detailed documentation on how reference data are produced, it is difficult to investigate this point, which may seriously hamper inter-comparability among different satellite products.

On the other hand, although direct validation is the more desired approach, it is limited in space and time in case in situ observations are used as a reference. The typical disadvantage is the scale contrast between the satellite observations and in-situ measurements, in terms of its usage for validation. In addition, the operation of in-situ observation systems is affected by changes in various elements (e.g., changes in instrumentation, station moves, the local environment and observing practices), which affect the establishment of consistent, continuous and quality-controlled in situ observations (Peterson et al., 1998). Therefore, it is necessary to define:

A-1. Climate observation requirements for in situ sites. This implies the assessment of the performance of in situ sites in terms of representativeness, homogeneity, and long-term stability;

A-2. Guidelines for in situ data processing, especially for data quality control and for the up-scaling;

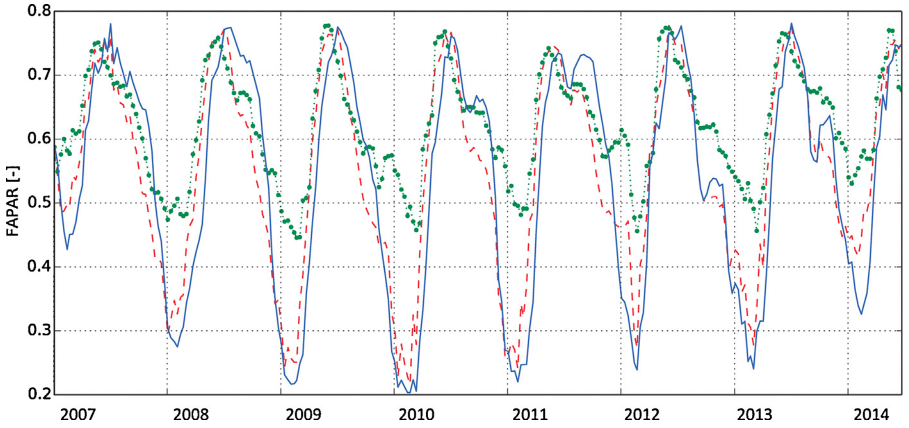


Fig. 3. Mean 10-daily values of FAPAR over France from January 2007 to June 2014 derived from: (green dots) the CGLS GEOV1 satellite-derived product at 1 km × 1 km spatial resolution, (blue solid line) the ISBA-A-gs land surface model at 8 km × 8 km spatial resolution, (red dashed line) the analysis resulting from the assimilation of the CGLS LAI and surface soil moisture products into the ISBA-A-gs land surface model at 8 km × 8 km spatial resolution. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

A-3. Protocols for in situ data collection.

Note that direct validation may involve high-resolution satellite data or flight campaign based observations. These observations can be used to solve the spatial coverage issue of in situ data. However, their temporal coverage is usually very limited.

After the establishment of reference data, the specific validation process should be identified, which needs:

- Definition of validation methods.

It is well accepted that different ECV CDRs have different validation processes and methods, as well as the establishment of reference data. The validation process should be practiced with essential strategies or mechanisms, including:

- Validation performed by the product developer;
- Independent assessment of products; and
- External review of the validation process.

As a final step in the validation process, an effort should be made to assess the consistency of the monitored climate data with multiple independent datasets of either the same variable or different (but physically interlinked) variables:

- Consistency check. One may implement an analysis of physical consistency among different variables that are independent of each other. One may also inter-compare multiple datasets for one common physical variable, to check spatial-temporal consistency between each other.

On top of the six essential points (A–F) listed above, validation facilities should be developed to automatize the validation process (i.e., operational validation level). The validation facility may include the development of dedicated validation tools, and infrastructures for data access and archiving (Weiss et al., 2014). The operational validation refers to the sustaining of the established validation processes and methods. In fact, the operational validation is implemented to perform automatic validation. In such way, the validation process becomes less and less time-consuming and more and more mature. Nevertheless, at operational validation level, the validation process should be reviewed externally regularly (e.g., annually to check the seasonal characteristics of errors and uncertainties).

The self-assessment serves as an internal validation process to ensure quality per product while the independent assessment is an unbiased validation process. The self-assessment refers to the validation implemented internally by the data producers while the independent assessment is done by an entity not involved in the making of the product. Such independent entities are entitled, at the same time, to implement an external review (e.g., evaluation) of the internal validation process. The external evaluation process should include a review of all validation-relevant documents.

It is well noted that the independence is an essential element of validation. For example, the reference data and the assessment of products should be as much independent as possible. It, therefore, comes to the point for understanding the degree of independence of the whole validation process. With respect to reference data, the potential independence levels (ESA, 2010) can vary among:

a Independent in situ data;
b Other in-situ data;
c Airborne campaign datasets for medium-scale comparisons;
d Other satellite datasets for large-scale comparisons;
e Historic datasets, trends, climatology for large-scale comparisons;
f Impact studies using other products (e.g., consistency among different variables);
g Cross-cutting quality monitoring using a data assimilation approach.

For an independent assessment and external review process, the independence is ensured by the entities such as:

a 'volunteer' external parties that have no connections to the making of products (e.g., pursuit of scientific excellence);
b contracted external parties;
c other external parties (e.g., end-users, stakeholders, commercial companies).

It must be realized that there is limited data available for validation for some variables. For example, the data developers may have used all available data sets to generate those variables. In that case, to ensure the independent validation, a completely new validation data set has to be produced. When the process of producing new data sets is not at the place, the validation might also involve comparisons with model data sets. This method should be used with cautiousness as a model itself suffers from limitations and uncertainties.

Table 2
Matching the generic validation process with the "Validation" and "Formal Validation Report".

Formal validation report	Validation	Generic validation process
Report on limited validation available from PI (Principal Investigator)	Validation using external reference data done for limited locations and times	1. The generation of independent reference data; 2. Assessing independence levels of reference data; 3. Self-assessment;
Report on comprehensive validation available from PI; Paper on product validation submitted	Validation using external reference data done for global and temporal representative locations and times	
Report on inter-comparison to other CDRs, etc. Available from PI and data Provider; Journal paper on product validation published	Score 3 + (Inter) comparison against corresponding CDRs (other methods, models, etc.)	7. Consistency check for inter-related CDRs;
Score 4 + Report on data assessment results exists	Score 4 + data provider participated in one inter-national data assessment	4. Independent assessment; 5. External review of self-assessment; 6. Assessing independence levels of point 4 & 5;
Score 5+ Journal papers more comprehensive validation, e.g., error covariance, validation of qualitative uncertainty estimates published		
	Score 4 + data provider participated in multiple inter-national data assessment and incorporating feedbacks into the product development cycle	8. Sustaining established processes & methods.

4. Proposed generic validation process

According to the analysis of the current practices, the proposed generic validation process for CDRs/ECVs may include:

a The generation of independent reference datasets;
b Assessing independence levels of reference datasets (Section 3.3);
c Self-assessment;
d Independent assessment;
e External review/evaluation of self-assessment validation practices;
f Assessing independence levels of point 4 and 5 (Section 3.3);
g Consistency check for inter-related CDRs/ECVs;
h Sustaining established processes and methods.

The above generic validation process differs from protocols (e.g., see A-1 to A-3 in Section 3.3) and is abstracted from the essentials of current validation practices (Section 3). It serves as a checklist for identifying how far the validation process has been approaching the current identified 'best practice' of validation.

For each point identified above, the specified protocol should be defined and followed. For example, for the first point, in the case of using in situ data as reference, the protocols listed in A-1 to A-3 should be considered. On the other hand, in the case of using satellite or flight campaign based observation as references, the protocol with a similar concept as for in-situ data should be defined. It is imperative to acknowledge that understanding how the reference data are established can ensure the traceability of a validation process. In fact, each point from the above should be made as transparent as possible to ensure traceability.

The first three points are internal validation processes. For the point 2, the independence scale of reference data should be defined. The point 3, self-assessment, requires the definition and documentation of the validation methods, validation plans (e.g., case studies, approaches) and error characterization (e.g., error propagation, error budget).

The external validation process consists of points 4–6. The independent assessment holds the same requirements (to be transparent) as the self-assessment does, which is to document the validation methods, validation plans, and error characterization. One step further is that it necessitates the following point 6 to assess the independence level of the external assessments. Additionally, one of the external validation process (i.e., the point 5) facilitates the independent review on the self-assessment validation, by checking the transparency and traceability of the internal validation activities. The independence level of point 5 should be checked as well by point 6, in which the independence level of external assessments can be represented by the entities implementing assessment and reviewing.

Except for the internal and external validation process, points 7 and 8 serve as the synergic components of the generic validation process. The consistency check can help identify gaps in gridded information in the interactions and exchanges between the atmosphere, ocean, and land domains. It is to check the physical consistency of the monitored climate variables with other climate variables. For example, the closing of the hydrological cycle can be checked by investigating the physical consistency between the runoff, groundwater, soil moisture, snow, precipitation, and evapotranspiration data. The consistency check is also referred to the (inter) comparison among multiple independent datasets of the same physical variables, e.g., produced by different algorithms. For point 8, the aim is to achieve an operational validation level, at which validation activities and data release are regularly implemented. Both points 7 and 8 should be documented to enable traceability.

5. Quality indicators for assessing the validation process

As discussed before, Cal/Val is the indispensable parts of the CDR processing chain (Fig. 1). It is recognized that different Cal/Val processes applied to the same raw sensor data may lead to the generation of data product for different purposes. For example, the validation may be implemented to meet a short-term weather purpose or a long-term climate purpose. Therefore, a yardstick is needed to distinguish how far the current validation practice is approaching the "best practice" of a validation process for the generation of climate data records. Arguably, this corresponds to the concept of "maturity of a validation process". The sub-thematic parts "user documentation" and "uncertainty characterization" of a SMM (System Maturity Matrix) (EUMETSAT, 2014a) can be used to assess the maturity of a validation process.

The "user documentation" part of the SMM is deemed as essential to the effective use and understanding of a data record. There are four sub-thematic areas to assess the completeness of user documentation, with six maturity levels. The sub-thematic areas include (1) formal description of scientific methodology; (2) formal validation report; 3) formal product user guide, and 4) formal

Table 3
Assessment of ESA-CCI soil moisture with "User Documentation" and "Validation" of SMM.

	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6
User documentation						
√Formal description of scientific methodology	√	√	√	√	√	
Formal Validation Report	√	√	√	√		
Formal Product User Guide	√	√	√	√	√	√
Formal description of operations concept	√	√	√	√		
Uncertainty characterization						
Standards	√	√	√	√	√	
Validation	√	√	√	√	√	
Uncertainty quantification	√	√	√	√	√	
Automated quality monitoring	√	√	√	√	√	

description of operations concept. It emphasizes the traceability of each step in climate data generation process (e.g., see Fig. 1), by documenting those steps. The "uncertainty characterization" of the SMM assesses (with six maturity levels) the practices used to characterize and represent uncertainty in the data record. It includes four sub-thematic areas trying to encompass (1) the validation standards used, (2) the validation process, (3) how uncertainty is quantified and (4) if an automated quality monitoring is implemented. This category emphasizes the traceability of each step of a validation process to characterize uncertainty.

5.1. Mapping the validation process using a SMM

As discussed before, the "User Documentation" category focuses on documenting each step for generating a climate data record, to assess the traceability of the CDR generation process. Meanwhile, the "Uncertainty Characterization" category tries to assess the traceability of the uncertainty characterization process. The sub-thematic components of these two categories, "Formal Validation Report" and "Validation" correspond directly to the assessment of a validation process. Other sub-thematic components are also relevant. For example, the "formal description of scientific methodology" is relevant to the second validation aspect (i.e., TCDR generation) as identified in Section 2.2. The "automated quality monitoring" is relevant to sustaining the validation process. It is emphasized that the validation process assessed by the "validation" component should be documented to enable traceability. The maturity level of such documenting process can be assessed by the "Formal Validation Report" component.

The four validation aspects as introduced in Section 2 are generalized from the generation chain of CDR shown in Fig. 1. It helps to understand how the current validation practices are documented, in a way enabling the analysis of the validation processes. Based on the analysis, a generic validation process can be proposed as discussed in section 4. We can see that the second validation aspect, "TCDR generation", is not included in the generic validation process. It is however covered by "Formal description of scientific methodology" in the "User Documentation" of the SMM. The SMM spares the need to consider this aspect of the assessment of the validation process.

Table 2 shows how the eight generic validation process points were mapped to the "validation" and "Formal Validation Report". The first three points of the generic validation process can be benchmarked by the maturity level 2 & 3. These two maturity levels refer to validation implemented using external reference data over various representative locations and times. This means the validation has been implemented by using the independent reference data and may involve already the self-assessment. Therefore, the first three generic points are corresponding to these two maturity levels. For the maturity level 4, the (inter) comparison is emphasized. This level includes the comparison of multiple independent datasets (e.g., derived from different methods & models) for the same physical variable or different variables that are physically inter-related. At this sense, the generic point 7 can be matched with the maturity level 4. When comes to the maturity level 5, it tries to seek the community acceptance through participating international data assessment. The generic points 4, 5 & 6 are tackling issues relevant to external assessments, and are matched with the maturity level 5 of "Validation", and the maturity levels 5 & 6 of "Formal Validation Report" (Table 2). The maturity level 6 of "Validation" is not corresponding to that of "Formal Validation Report". It is because the level 6 of "Validation" includes the concept of sustaining the validation process (e.g., the incorporation of user feedbacks into the product development cycle), while the level 6 of "Formal Validation Report" does not. The generic point 8 is about sustaining the validation process and matches with the description of the maturity level 6 of "Validation". Nevertheless, it is to note that for the "User Documentation" in the SMM, the sustaining concept or mechanism of a validation process is partially covered with the sub-thematic column of "Formal Description of operations concept".

5.2. Demonstration example

Table 3 shows an example of how the ESA-CCI soil moisture product (Liu et al., 2011; Dorigo et al., 2012; Liu et al., 2012; Chuang et al., 2014; Dorigo et al., 2014) can be assessed with the SMM on "User Documentation" and "Uncertainty Characterization". The ticked cells indicate the maturity level being reached for ESA CCI soil moisture product. Level 5 is reached by "formal description of scientific methodology" as the algorithm has been updated corresponding to the updates of the data record and has proper document version numbering. There is a maturity level of 4 for "Formal validation report". Although there are detailed reports and published papers on validation using in situ reference data and reanalysis data, and the inter-comparison result to other ECV CDRs, there is not yet a report on climate data assessment. As for the "Formal Product User Guide", the updated user guide is available from data provider's web page and is corresponding to the updated dataset. Therefore, it is at a matured level. For the time being, the ESA-CCI soil moisture is in a transition phase from phase 1 to phase 2. For example, it is transitioning from the scientific consultation to the detailed specification for system development and data product generation. Although there is a comprehensive description of the operation concept, there is limited practical implementation available. As a result, the "Formal description of operations concept" is at the maturity level 4.

A similar explanation for how the maturity level being scored for "Uncertainty Characterization" can be given. For "Standard" column (scored level 5), although the existing document describes how traceable comparison was chained to a specified reference, it has not yet been fully established. For "validation" column (scored level 5), although the data provider has participated in international data assessment workshops, the feedbacks from end-users have not yet been fully and operationally considered into the product development. It is realized that meaningful user feedbacks and requirements can only be formulated after many years of research and development efforts (Mittelbach et al., 2013). For "Uncertainty Quantification" column (scored level 5), the quantitative estimates of uncertainty have been provided, and the temporal and spatial error covariance quantified. However, the uncertainty estimate has not been validated using superior quality data sets. For "Automated Quality Monitoring" (scored level 5), the automated monitoring has been partially implemented but the resulting feedback has not yet being incorporated into metadata or documentation. If one checks how close the ESA-CCI Soil Moisture approaches the generic validation process as identified in Table 2, the first 7 generic validation process points have been reached. Nevertheless, it is not yet fully reaching the point 8, which is relevant to sustaining a validation process.

6. Conclusions

6.1. A generic validation process

Cal/Val is the integral component of the Fundamental Climate Data Records (FCDRs) and Thematic Climate Data Records (TCDRs) processing chains, as illustrated in Fig. 1. Particularly, the validation concept was discussed and the validation aspects were identified. The validation aspects help to understand how the current validation practices are documented, in a way enabling the analysis of the validation processes of some typical examples. The examples with different validation capacities indicated that the validation strategy is driven by validation requirement and also constrained by the existing validation capacities. The examples with different comparison methods show that the indirect comparison method can help to identify systematic uncertainties and errors, and to identify the gap in the current observation or modelling systems.

Based on the analysis of current practices, the essentials of validation are identified and a generic validation process is proposed. The generic validation process consists of internal validation, external validation and synergic components. The internal validation is usually implemented before and after the release of data products by using external reference data. The external validation brings the released data products (and the associated self-assessment approaches) to a broader platform for evaluation. The synergic component addresses the consistency of the data product and the sustaining mechanisms of the validation process (e.g., to enable an operational/regular validation practice).

The identified generic process starts from the generation of independent reference data, which represents the existing validation capacity. This capacity is the fundamental factor affecting the validation process. The validation requirement is another prerequisite to define a validation process. The consideration of both validation capacity and validation requirements drives the final definition of a validation process. By the nature of validation, independence is an essential element. Based on this nature, the proposed generic validation process is to emphasize independence at each step. It is recommended to assess the degree of independence of the reference datasets and the external reviewers. Such assessment can be subsequently implemented by using community-accepted independence levels (in Section 3.3).

Assessing the consistency of products highlights the possible synergies between various products, which in the end may help improve the overall quality of the ECV products. The use of data assimilation, i.e., the integration of various ECV products into models for reanalysis is a key component of the validation. Data assimilation relies on a thorough analysis of the errors and permits the assessment of the consistency between ECV products, provided that the input data, the boundary condition and the model physics are kept unchanged. There is a strong heritage related to the data assimilation techniques for the atmospheric and oceanic variables. The importance of such cross-cutting quality monitoring, i.e., check the consistency between distinct ECV products (e.g., LAI and surface albedo over land) or products across domains (e.g. soil moisture and precipitation), has to be emphasized. The data assimilation systems used to produce reanalysis can be used to monitor the consistency of ECV products of a given domain (e.g. the LDAS used in Copernicus Global Land Service). It is important to have users involved in the validation. The impact of using the products in applications is a key information for further improvement of the products.

6.2. Quality indicators for assessing the validation process

The thematic categories of "User Documentation" and "Uncertainty Characterization" in the SMM (System Maturity Matrix) were referenced for the assessment of the validation process. The most relevant components of these two thematic categories, "Formal Validation Report" and "Validation" (six maturity levels), were adopted as a quality indicators to assess the validation process. The generic validation process was mapped to these indicators and a demonstration on how to use these indicators to assess the validation process was shown. This is an analytic approach to the very complex question of validation. One should remember that these quality indicators reflect the quality of the process, not the quality of the product.

The quality indicators allows one to comprehensively check the validation. The ESA CCI soil moisture project was used as an example. Through the assessment, the validation process of the ESA CCI soil moisture project phase one has reached the maturity level 6 of "Formal Validation Report", and level 5 for "validation". From the assessment, it is clear that the documentation of the validation process is matured (i.e., level 6 of "Formal Validation Report"). However, it does not mean that the validation process has been matured as a whole. For instance, it is not yet reaching the sustaining level of the validation process (i.e., level 5 for "validation"). It suggests that the ESA-CCI soil moisture project should develop the relevant validation facility to automatize its validation process, in order to achieve higher maturity level.

It is to note that the examples shown in this study are for satellite products. In the case of in situ data sets or reanalysis, the assessment of "user documentation" and "uncertainty characterization" can be quite different. For in situ data, the quality control process is corresponding to the validation process. For reanalysis data, the monitoring of different standard metrics is referred to the validation practice (Zeng et al., 2014). Therefore, the mapping of the validation process for in situ and reanalysis to the quality indicators requires different considerations. The recently organized CORE-CLIMAX Capacity Assessment Workshop (EUMETSAT, 2014b) has paved the way to incorporate the smooth mapping of the validation process for satellite, in situ and reanalysis datasets.

Acknowledgements

This work is a contribution to the CORE-CLIMAX coordination action, co-funded by the European Commission within the Copernicus initiative in FP7 (grant agreement No 313085).

References

Barbu, A.L., Calvet, J.C., Mahfouf, J.F., Lafont, S., 2014. Integrating ASCAT surface soil moisture and GEOV1 leaf area index into the SURFEX modelling platform: a land data assimilation application over France. Hydrol. Earth Syst. Sci. 18 (1), 173–192.

Baret, F., Hagolle, O., Geiger, B., Bicheron, P., Miras, B., Huc, M., Berthelot, B., Niño, F., Weiss, M., Samain, O., Roujean, J.L., Leroy, M., 2007. LAI, fAPAR and fCover

CYCLOPES global products derived from VEGETATION: Part 1: Principles of the algorithm. Remote Sens. Environ. 110 (3), 275–286.

Baret, F., H. Makhmara, R. Lacaze, B. Smets (2013). BioPar Product User Manual: LAI, FAPAR, FCover, NDVI version 1 from SPOT/BEGETATION data: FAPAR. GIO Global Land Component -Lot I Operation of the Global Land Component(Issue I1.00): pp 39.

Beggs, H., Verein, G., Kippo, H., Underwood, M., 2012. Enhancing ship of opportunity sea surface temperature observations in the Australian region. J. Oper. Oceanogr. 5 (1), 59–73.

Bojinski, S., 2011. Response by ESA to GCOS – results of the climate change initiative requirements analysis. Eur. Space Agency, ESRIN, DG-H/2011/3007/ECO/dms/kw: pp72.

Calvet, J. C., Barbu, A. L., Fairbairn, D., Gelati, E., Meurey, C. (2014). Cross-cutting quality monitoring report. GIO Global Land Component -Lot I Operation of the Global Land Component(I1.00): 1–45.

Camacho, F., Cernicharo, J., Lacaze, R., Baret, F., Weiss, M., 2013. 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. 137 (0), 310–329.

CEOS-WGCV-LPV (2009). A hierarchical approach to classify land product validation stages. http://lpvs.gsfc.nasa.gov/

Chander, G., Hewison, T., Fox, N., Wu, X., Xiong, X., Blackwell, W., 2013. Overview of intercalibration of satellite instruments. IEEE Trans. Geosci. Remote Sens. 51 (3), 1056–1080.

Chuang D., Dorigo W., Hahn S., Melzer T., Paulik C., Reimer C., Vreugdenhil M., Wagner W., Kidd R. (2014). ECV Production, Fusion of Soil Moisture Products: Algorithms Theoretical Baseline Document, Version 2.0. ESA CCI Phase I Soil Moisture Deliverables: 1–32.

D'Odorico, P., Gonsamo, A., Pinty, B., Gobron, N., Coops, N., Mendez, E., Schaepman, M.E., 2014. Intercomparison of fraction of absorbed photosynthetically active radiation products derived from satellite data over Europe. Remote Sens. Environ. 142, 141–154.

Dorigo, W., de Jeu, R., Chung, D., Parinussa, R., Liu, Y., Wagner, W., Fernandez-Prieto, D., 2012. Evaluating global trends (1988–2010) in harmonized multi-satellite surface soil moisture. Geophys. Res. Lett., 39.

Dorigo, W.A., Gruber, A., De Jeu, R.A.M., Wagner, W., Stacke, T., Loew, A., Albergel, C., Brocca, L., Chung, D., Parinussa, R.M., Kidd, R., 2014. Evaluation of the ESA CCI soil moisture product using ground-based observations. Remote Sens. Environ. 0.

Dowell, M., Lecomte, P., Husband, R., Schulz, J., Mohr, Y., Tahara, R., Eckman, E., Lindstrom, E., Wooldridge, C., Hilding, S., Bates, J., Ryan, B., Lafeuille, J., Bojinski, S., (2013). Strategy towards an architecture for climate monitoring from space This report is available from: <www.ceos.org/>; <www.wmo.int/sat/>; http://<www.cgms-info.org/>: pp. 39.

Ellison, S. L. R., Rosslein, M., Williaams, A. (2000). Quantifying uncertainty in analytical measurement. EURACHEM/CITAC Guide, ISBN 0 948,926 15 5.

Entekhabi, D., Reichle, R.H., Koster, R.D., Crow, W.T., 2010. Performance metrics for soil moisture retrievals and application requirements. J. Hydrometeorol. 11 (3), 832–840.

ESA, 2010. ESA climate change initiative: CCI project guidelines. Eur. Space Agency, ESRIN, EOP-DTEX-EOPS-SW-10-0002: pp 52.

Estévez, J., Gavilán, P., Giráldez, J.V., 2011. Guidelines on validation procedures for meteorological data from automatic weather stations. J. Hydrol. 402 (1–2), 144–154.

EUMETSAT (2014a). CORE-CLIMAX System Maturity Matrix Instruction Manual. Deliverable D222, CC/EUM/MAN/13/002, http://www.coreclimax.eu/?q=Reports: pp.41.

EUMETSAT (2014b). Preparing the European Climate Data Records for the Copernicus Climate Change Monitoring Service. Copernicus Observer, In Motion: http://newsletter.gmes.info/issue-08/article/preparing-european-climate-data-records-copernicus-climate-change-monitoring

R. Fernandes, S. Plummer, J. Nightingale, F. Baret, F. Camacho, H. Fang, S. Garrigues, N. Gobron, M. Lang, R. Lacaze, S. LeBlanc, M. Meroni, B. Martinez, T. Nilson, B. Pinty, J. Pisek, O. Sonnentag, A.Verger, J. Welles, M. Weiss, J.L. Widlowski (2014). Global Leaf Area Index Product Validation Good Practices Version 2.0. in G. Schaepman-Strub, M. Roman, & J. Nickeson (Eds.) Best Practice for Satellite-Derived Land Product Validation (p.76): Land Product Validation Subgroup, doi:10.5067/doc/ceoswgcv/lpv/lai.002: 1–76.

Fetterolf, D.M., 2007. Developing a sound process validation strategy. Int. BioPharm. 20 (12), 6.

Garrigues, S., Lacaze, R., Baret, F., Morisette, J.T., Weiss, M., Nickeson, J.E., Fernandes, R., Plummer, S., Shabanov, N.V., Myneni, R.B., Knyazikhin, Y., Yang, W., 2008. Validation and intercomparison of global leaft area index products derived from remote sensing data. J. Geophys. Res. 113, G02028, http://dx.doi.org/10.01029/02007JG000635

GCOS (2003). GCOS Climate Monitoring Principles. http://www.wmo.int/pages/prog/gcos/index.php?name=ClimateMonitoringPrinciples

GCOS, 2010a. Guideline for the Generation of Datasets and Products Meeting GCOS Requirements. GCOS-143, Geneva, Switzerland, pp. 12.

GCOS, 2010b. Implementation Plan for the Global Observing System for Climate in Support of the UNFCCC (2010 UPDATE). GCOS-138, Geneva, Switzerland, pp. 180.

Gentemann, 2014. Three way validation of MODIS and AMSR-E sea surface temperatures. J. Geophys. Res. 119, 2583–2598, http://dx.doi.org/10.1002/2013JC009716

Hennemuth, B., Bender, S., Bulow, K., Dreier, N., Keup-Thiel, E., Kruger, O., Mudersbach, C., Radermacher, C., Schoetter R.,(2013). Statistical methods for the analysis of simulated and observed climate data, applied in projects and institutions dealing with climate change impact and adaptation. CSC report 13, Climate Service Center, Germany pp.135.

JGCM (2008). Evaluation of measurement data Guide to the expression of uncertainty in measurement. JGCM 100:2008: Copyright of this document is shared jointly by the JGCM member organizations (BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML).

Kwok, R., 2010. Satellite remote sensing of sea-ice thickness and kinematics: a review. J. Glaciol. 56 (200), 1129–1140.

Liu, Y.Y., Dorigo, W.A., Parinussa, R.M., de Jeu, R.A.M., Wagner, W., McCabe, M.F., Evans, J.P., van Dijk, A., 2012. Trend-preserving blending of passive and active microwave soil moisture retrievals. Remote Sens. Environ. 123, 280–297.

Liu, Y.Y., Parinussa, R.M., Dorigo, W.A., De Jeu, R.A.M., Wagner, W., van Dijk, A., McCabe, M.F., Evans, J.P., 2011. Developing an improved soil moisture dataset by blending passive and active microwave satellite-based retrievals. Hydrol. Earth Syst. Sci. 15 (2), 425–436.

Mieruch, S., Schröder, M., Noël, S., Schulz, J., 2010. Comparison of monthly means of global total column water vapor retrieved from independent satellite observations. J. Geophys. Res. D23, 115.

Mieruch, S., Schröder, M., Noël, S., Schulz, J., 2014. Comparison of decadal global water vapor changes derived from independent satellite time series. J. Geophys. Res. Atmos. 119, 489–412, 499.

Mittelbach, H., Kidd, R., Dorigo,W., De Jeu, R., Wagner, W., Seneviratne S., Haas E. (2013). ESA CCI Soil Moisture: User Requirements Document (URD). ESA CCI Phase I Soil Moisture Deliverables D1.1(2) 1-49.

Myneni, R.B., Hoffman, S., Knyazikhin, Y., Privette, J.L., Glassy, J., Tian, Y., Wang, Y., Song, X., Zhang, Y., Smith, G.R., Lotsch, A., Friedl, M., Morisette, J.T., Votava, P., Nemani, R.R., Running, S.W., 2002. Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data. Remote Sens. Environ. 83, 214–231.

The climate data guide: sea ice concentration data overview, comparison table and graphs. Retrieved from https://climatedataguide.ucar.edu/climate-data/sea-ice-concentration-data-overview-comparison-table-and-graphs 2015.

NRC, 2004. Climate Data Records from Environmental Satellites: Interim Report. The National Academies Press, Washington, D.CPP136.

Peterson, T.C., Easterling, D.R., Karl, T.R., Groisman, P., Nicholls, N., Plummer, N., Torok, S., Auer, I., Boehm, R., Gullett, D., Vincent, L., Heino, R., Tuomenvirta, H., Mestre, O., Szentimrey, T., Salinger, J., Førland, E.J., Hanssen-Bauer, I., Alexandersson, H., Jones, P., Parker, D., 1998. Homogeneity adjustments of in situ atmospheric climate data: a review. Int. J. Climatol. 18 (13), 1493–1517.

Pinty, B., Clerici, M., Andredakis, I., Kaminski, T., Taberner, M., Verstraete, M.M., Gobron, N., Plummer, S., Widlowski, J.L., 2011. Exploiting the MODIS albedos with the two-stream inversion package (JRC-TIP): 2. Fractions of transmitted and absorbed fluxes in the vegetation and soil layers. J. Geophys. Res. Atmos. 116 (D9), D09106.

Raymaekers, D., Garcia, A., Di Bella, C., Beget, M.E., Llavallol, C., Oricchio, P., Straschnoy, J., Weiss, M., Baret, F., 2014. SPOT-VEGETATION GEOV1 biophysical parameters in semi-arid agro–ecosystems. Int. J. Remote Sens. 35 (7), 2534–2547.

Roebeling, R.A., Wolters, E.L.A., Meirink, J.F., Leijnse, H. (2012). Triple Collocation of Summer Precipitation Retrievals from SEVIRI over Europe with Gridded Rain Gauge and Weather Radar Data. Journal of Hydrometeorology, 13, (5) 1552–1566.

Scipal, K., Holmes, T., de Jeu, R., Naeimi, V., Wagner, W., 2008. A possible solution for the problem of estimating the error structure of global soil moisture data sets. Geophys. Res. Lett. 35 (24), L24403.

Smolyanitsky, V., (2012). WMO standards and guidelines for sea ice: ice charting and observations. A presentation for the GCW CryoNet Implementation Meeting: Vienna, Austria, 20-22 November 2012.

Spreen, G., Kaleschke, L., Heygster, G., 2008. Sea ice remote sensing using AMSR-E 89-GHz channels. J. Geophys. Res. Oceans 13 (C2), C02S03.

ST-VAL (2008). Satellite sea surface temperature validation (ST-VAL). https://www.ghrsst.org/ghrsst/tags-and-wgs/stval-wg/

Stoffelen, A., 1998. Toward the true near-surface wind speed: Error modeling and calibration using triple collocation. Journal of Geophysical Research: Oceans 103 (C4), 7755–7766.

Su, C.-H., Ryu, D., Crow, W.T., Western, A.W., 2014. Beyond triple collocation: applications to soil moisture monitoring. J. Geophys. Res. Atmos. 119 (11), 6419–6439.

Su, Z., de Rosnay, P., Wen, J., Wang, L., Zeng, Y., 2013a. Evaluation of ECMWF's soil moisture analyses using observations on the Tibetan plateau. J. Geophys. Res. Atmos. 118, 5304–5318.

Z. Su, W. J. Timmermans, Y. Zeng (2013b). 2nd collocation meeting of EC FG7, ESA, and EUMETSAT ECV projects, 15–16 May 2013, Brussels, Belgium: Meeting Report. GEWEX News 23(3): 10–11.

USGCRP (2003). Strategic plan for the U.S. climate change science program, July 2003. Chapter 12, Observing and monitoring the climate system,. Washington DC 142 pp.

Weiss, M., Baret, F., Block, T., Koetz, B., Burini, A., Scholze, B., Lecharpentier, P., Brockmann, C., Fernandes, R., Plummer, S., Myneni, R., Gobron, N., Nightingale, J., Schaepman-Strub, G., Camacho, F., Sanchez-Azofeifa, A., 2014. On line validation exercise (OLIVE): a web based service for the validation of medium resolution land products. Application to FAPAR products. Remote Sens. 6 (5), 4190–4216.

WMO, 2010. Sea-Ice Information Services in the World. WMO-574, Geneva Switzerlands, pp. 92.
WMO, 2011. Guide to Climatological Practices. WMO-100, Geneva Switzerlands, pp. 117.
E. Wolters, W. Dierckx, J. Dries, E. Swinnen (2014). Proba-v products user manual v1.1. http://proba-v.vgt.vito.be/sites/default/files/Product_User_Manual.pdf1–74.
Xiong, X., Cao, C., Chander, G., 2010. an overview of sensor calibration inter-comparison and applications. Front. Earth Sci. China 4 (2), 237–252.
Yilmaz, M.T., Crow, W.T., Anderson, M.C., Hain, C., 2012. An objective methodology for merging satellite- and model-based soil moisture products. Water Resour. Res. 48 (11), W11502.
Zeng, Y., Z. Su, J., Schulz, J.-C. Calvet, P., Poli, D., Tang, T., Manninen, A. Riihelä, C.-M. Tanis, A.-N. Arslan, E., Swinnen, C., Tote, A., Obregon, K.-W. Andre, R., Roebeling, V., John, K. Kaspar,W. J. Timmermans and J. Timmermans (2014). Generic Validation Strategy for CDRs/ECVs. CORE-CLIMAX Deliverables (De3.32): pp. 1-31.