Generic Validation Strategy for CDRs/ECVs

Deliverable De3.32

Issue 1.2

Due date of deliverable: 15 February 2014
Actual submission date: 18 February 2014

CORE-CLIMAX

Coordinating Earth observation data validation for RE-analysis for CLIMAte ServiceS

FP7 Grant agreement No 313085

Document History

Date	Update	Contributor	Organization
04/10/2013	Preliminary Structure Drafted	J. Schulz, Z. Su, W. Timmermans, Y. Zeng	EUMETSAT, ITC
14/10/2013	Comments on Structure	J. Schulz, K.-W. Andrea, T. Manninen	EUMETSAT, DWD, FMI
04/12/2013	Detailed Writing Plan Drafted	Y. Zeng, W. Timmermans, Z. Su	ITC
05/12/2013	Comments on Writing Plan	J. Schulz, E. Swinnen, T. Manninen	EUMETSAT, VITO, FMI
18/12/2013	Revised Writing Plan	Y. Zeng, W. Timmermans, Z. Su	ITC
19/12/2013	Telecom meeting on actions points, distributing writing tasks among MTF, ITC and FMI	Y. Zeng, J.-C. Calvet	ITC, MTF
10/01/2014	fAPAR input updated	E. Swinnen, C. Tote	VITO
11/01/2014	In-situ validation updated	A. Obregon, K.-W. Andrea, K. Frank	DWD
16/01/2014	Meteo-France input updated	J.-C. Calvet	MTF
17/01/2014	4 different Snow cover and one sea ice products added to chapter 2. Current Validation Practices	Cemal M. Tanis, Ali N. Arslan,	FMI
03/02/2014	Comments/Revisions	J.-C. Calvet	MTF
03/02/2014	Atmospheric composition	A. Riihelä, T. Manninen	FMI
05/02/2014	Updated Summary on atmospheric composition	A. Riihelä, T. Manninen	FMI
05/02/2014	Condensed input on snow cover and sea ice	Ali N. Arslan	FMI
05/02/2014	Condensed inputs on in-situ validation	A. Obregon, K.-W. Andrea, K. Frank	DWD
07/02/2014	Reanalysis Validation	Paul Poli, David Tan	ECMWF
11 /02/2014	First draft completed	Y. Zeng, J.-C. Calvet, W. Timmermans, Z. Su	ITC, MTF
12/02/2014	Comments	T. Manninen	FMI
12/02/2014	Revision updated	Y. Zeng, W. Timmermans, Z. Su	ITC
14/02/2014	Comments	T. Manninen	FMI
14/02/2014	Comments	E. Swinnen, C. Tote	VITO
14/02/2014	Revision Updated	Y. Zeng, W. Timmermans, Z. Su	ITC
17/02/2014	Comments	R. Roebeling, V. John, J. Schulz	EUMETSAT
17/02/2014	Comments	D. Tan	ECMWF
17/02/2014	Revision Updated	Y. Zeng, W. Timmermans, Z. Su	ITC
24/03/2014	Revision Updated	Y. Zeng, W. Timmermans, Z. Su	ITC

Document Release Sheet

Book captain:	Y. Zeng (ITC)	Sign Date
Approval (lead contractor for WP3):	Bob Su (ITC)	Sign Date
Endorsement:	Bob Su, Co-ordinator (ITC)	Sign Date

Dissemination level	
Public	**x**
Restricted to a group (including the Commission Services) specified by the consortium	
Confidential, only for members of the consortium and for the Commission Services	

Table of Contents

Acronym:

CDR: Climate Data Record

CEOS: Committee on Earth Observation Satellites

DG CLIMA: Directorate-General for Climate Action

ECV: Essential Climate Variable

ERA-Clim: European Reanalysis of Global Climate Observations

EURO4M: European Reanalysis and Observations for Monitoring

fAPAR: Fraction of Absorbed Photosynthetically Active Radiation

FCDR: Fundamental Climate Data Record

GMES: Global Monitoring for Environment and Security, now is called as Copernicus

GCOS: Global Climate Observing System

GSICS: Global Space-based Inter-Calibration System

GISC: GMES In-Situ Coordination

LAI: Leaf Area Index

MONARCH: Monitoring and Assessing Regional Climate Change in High Latitudes and the Arctic

MyOcean-II: Ocean Monitoring and Forecasting

MACC-II: Monitoring atmospheric Composition & Climate

NOAA: National Oceanic and Atmospheric Administration

SCOPE-CM: Sustained and coordinated processing of Environmental Satellite data for Climate Monitoring

TCDR: Thematic Climate Data Record

WGCV: Work Group for Calibration and Validation

WMO: World Meteorological Organization

1. Introduction

1.1. Background

Considering the full set of existing GCOS ECV products[1], some are only based on in-situ data records or on retrievals from satellite observations, while others are derived from a combination of both or through data assimilation with a dynamical model. In general, in-situ data records have the advantage of spanning much longer time than satellite-based data records that currently reach a maximum length of about 30 years. The largest advantage of satellite-based data records is their spatial coverage and spatially relatively homogeneous measurement accuracy. In-situ and satellite-based data records can be combined to derive ECV products. The 2011 GMES Climate Service report has also highlighted that climate services should integrate observations from both satellite and in-situ measurements in order to develop value added climate monitoring products and climate impact indicators. **On the other hand, the development of climate products requires the implementation of validation and benchmarking methods. In particular, a generic validation strategy for CDRs/ECVs shall be proposed**.

1.2. Scope

The main objective of this report is to assess how different European initiatives/services approach the validation issues of ECVs, at both the technical and the strategy levels. The technical aspects range from the collection of the data used for validation, the design of instrumented validation sites, to the performance matrices to be used for the evaluation of the service. Although the validation process may differ from one ECV to another, generic validation methods can be defined using the current developments made by individual ECV production teams.

1.3. Focus

The focus of this report therefore comes to a point that a transparent traceable validation procedure should be documented, which includes:

1) The description of "best practice" protocols for the validation of an ECV product or group of ECV products. For example, a hierarchical approach to classify land product validation stages was adopted by the land product validation subgroup of CEOS Working Group for Calibration and Validation (WGCV)[2], identified as "best practice" protocols for land product, including validation data collection, analysis and accuracy reporting.

2) The description of guidelines on "best practice" for generating FCDRs[3] and TCDRs[4] for particular ECVs. For example, the generation of long-term climate data record needs to integrate different satellite measurements taken at different times and locations, by different instruments, operated by various satellite agencies. It is

[1] http://www.wmo.int/pages/prog/gcos/index.php?name=EssentialClimateVariables
[2] http://lpvs.gsfc.nasa.gov/
[3] Fundamental Climate Data Record (FCDR): a long-term data record of calibrated and quality-controlled sensor data designed to allow the generation of homogeneous products that are accurate and stable enough for climate monitoring.
[4] Thematic Climate Data Record (TCDR): a long-term data record of validated and quality-controlled geophysical variables derived from FCDRs.

important to make sure that measurements between the sensors are physically consistent. This may involve adjusting multiple instruments to a common baseline. However, it has become increasingly important (e.g. with more and more environmental satellites) and demanding, as the slightly different changes in satellite characteristics (e.g. orbital drift, sensor degradation and instrumental biases) will affect the consistency and accuracy of the long-term FCDRs. GSICS has identified a need for a full history of operating changes that affect the performance and calibration of each instrument, which should include the date and time of each operating changes that affect the performance and calibration of each instrument, a short summary of the change, and a quantitative assessment of the severity of the impact on the instrument's calibration. In addition, GSICS has developed common methodologies and implements operational procedures to establish stable common reference standards, which will enable improved and consistent accuracy among space-based observations worldwide for climate monitoring, weather forecasting, and environmental applications. It is therefore important to document how the generation of FCDRs follows such common reference standard.

3) The criteria for selection of datasets for validation. This may include (not limited to) the criteria for selection, quality control, spatial/temporal representativeness, masking and flagging (e.g. in terms of satellite data sets).

4) The criteria for defining the validation procedures. This may include (not limited to) validation at regional, continental to global scale, validation on long term trend/stability/consistency, evaluation methods and measures, independency, review mechanism, sustaining mechanism.

5) The criteria for Error Characterisation. This may include (not limited to) error budget analysis (e.g. error arising from simplification in radiative transfer, geolocation/interpolation error, error arising from the uncertainty in parameters used to derive the measurement), error propagation and the approach adopted to implement these error characterizations. A complete assessment of the error characteristics shall be required to allow the detection of climate change impacts over and above natural variability.

In this report, an attempt is made to identify the "best practice" aiming to achieve an operational validation level, at which validation activities were regularly (e.g. 6-monthly or annually) implemented following the identified "best practice" and procedures. It is noted that short-term data records (e.g. <30yrs) will be taken into account for identifying the 'best practice' for validation. Although these data records cannot be defined as CDR at the current stage (e.g. due to the length of data record), they are ECV variables and hold the potentials to be CDR, by following the commonly identified 'best practice' protocol for validation.

Then, a quality indicator for assessing the completeness of validation shall be determined. From the raw sensor data to FCDR or TCDR, and subsequently to the ECV product, a number of validation steps are needed in order to meet GCOS requirements. The CORE-CLIMAX climate data record assessment tool (i.e. the system maturity matrix) will be used to assess how far the validation activities are approaching the completeness.

It is essential to recognize that the external reference[5] data (e.g. in-situ data) are required to assess the accuracy of space products and process, and calibrate and validate satellite data sets. At the meantime, the external reference data should be characterized in details on how they are established to ensure traceability. Furthermore, as indicated by GCOS, many ECV products[6] are directly derived from long-term in situ observations, which include air temperature, surface radiation budget, sea level, upper-air-temperature/upper-air wind speed/direction/upper-air water vapour, carbon dioxide, sea-surface temperature, river discharge, lake levels, snow water content, permafrost active layer thickness & soil temperature, soil carbon, soil moisture. Therefore, this report will also review the validation procedures for in situ products/services, focussing on:

- providing an overview of in-situ station networks and datasets;
- listing previous projects dealing with this;
- collecting the current community recommendations, and
- highlighting the missing issues.

In addition, this report will address the validation issues of reanalysis climate data records through identifying different components of assimilation-based reanalysis system.

1.4. Who are Addressed

Since CORE- CLIMAX is a coordinating project, the validation procedures should be intentionally high-level, conceptual and inclusive, in order to reach a broad consensus among all relevant communities.

At the European level there are major initiatives, which provide and intend to expand the provision of ECV products and services from satellites, including EU FP7 Climate projects (e.g. ERA-Clim, EURO4M, MONARCH), relevant Copernicus projects (e.g. MACC-II, MyOcean-II, Global Land Service), ESA CCI projects, EEA lead Projects (e.g. "Climate Change Impact and Vulnerability Indicator Report" & "Adaptation Report"), the EUMETSAT Central Facility and distributed Satellite Application Facilities such as that for Climate Monitoring (CM-SAF), and in-situ climate data records projects (e.g. E-OBS (KNMI), GPCC (DWD), HADSST (UKMO)). At the international level and national level, there are major coordination mechanisms and CDR programs existing, such as: Coordination Group for Meteorological Satellites (CGMS), CEOS Working Group on Climate, CEOS Working Group on Calibration/Validation (WGCV), CEOS WGCV on land products validation (CEOS WGCV-LPV), the WMO GSICS, SCOPE-CM programs, and NOAA CDR Programme. Except for the above mentioned international, national and European initiatives/services, stakeholders including WMO, GCOS, DG Climate and CEOS W0rking Group on Climate should be addressed as well.

1.5. Document Structure

In section 3, an overview of the current validation practices for satellite, in-situ and reanalysis data sets are summarized. Based on this review of current practices, the state-of-the-art is described together with the most recent innovations. In order to facilitate thorough description of the operationalization of validation, a number of essentials of validation

[5] In some cases, in-situ data are not the only source of reference data and the superior satellite or flight campaign based measurements can also serve as references.
[6] http://www.wmo.int/pages/prog/gcos/index.php?name=ObservingSystemsandData

procedure for CDRs/ECVs, and the climate data record assessment tool are summarized and introduced in section 4, respectively. The generic validation strategy is then proposed in section 5. The conclusions are presented in section 6.

2. Reference Documents

[R1] Global Land Service fAPAR Product Validation Report: http://land.copernicus.eu/global/sites/default/files/products/GIO-GL1_VR_FAPAR_I1.00.pdf

[R2] Global Land Service fAPAR Product ATBD: http://land.copernicus.eu/global/sites/default/files/products/GIO-GL1_ATBD_FAPAR_I1.00.pdf

[R3] Global Land Service fAPAR Product User Manual: http://land.copernicus.eu/global/sites/default/files/products/GIO-GL1_PUM_FAPAR_I1.00.pdf

[R4] MyOcean II Baltic Sea Regional NRT SAR Based Sea Ice Product User Manual: http://catalogue.myocean.eu.org/static/resources/myocean/pum/MYO2-SIW-PUM-011-004-and-011-v2.2.pdf

[R5] MyOcean II Baltic Sea Regional NRT SAR Based Sea Ice Product Quality Information Document: http://catalogue.myocean.eu.org/static/resources/myocean/quid/MYO2-OSI-QUID-011-ALL-V1.5.pdf

[R6] MyOcean II Baltic Sea Regional NRT SAR Based Sea Ice Product Validation Report: http://myocean.met.no/SIW-TAC/doc/myo-wp14-siw-fmi-bal-seaice_hr-obs-validation_latest.pdf

[R7] O3M SAF documents available at http://o3msaf.fmi.fi/documents.html

[R8] O3M SAF Validation Report for O3M-41 and O3M-42, doc. no. SAF/O3M/AUTH/VRR /O3, issue 1 .0

[R9] O3M SAF ORR VALIDATION REPORT for NTO/OTO, products O3M-50, -51, -52, and -53. version 1.1, 30/06/2013.

[R10] Algorithm Theoretical Basis Document for GOME-2 Total Column Products of Ozone, Tropospheric Ozone, NO2, Tropospheric NO2, BrO, SO2, H2O, HCHO, OClO and Cloud Properties. version 2/H, 21/05/2013.

[R11] Boersma, K. F., Eskes, H. J. and Brinksma, E. J.: Error analysis for tropospheric NO2 retrieval from space, J. Geophys. Res., 109(D4), doi:10.1029/2003JD003962, 2004.

[R12] Ma, J. Z., Beirle, S., Jin, J. L., Shaiganfar, R., Yan, P., and Wagner, T.: Tropospheric NO2 vertical column densities over Beijing: results of the first three years of ground-based MAX-DOAS measurements (2008–2011) and satellite validation, Atmos. Chem. Phys., 13, 1547-1567, doi:10.5194/acp-13-1547-2013, 2013.

[R13] O3M SAF ORR VALIDATION REPORT, O3M-58, issue/rev. 1 / 2

[R14] O3M SAF ORR VALIDATION REPORT, O3M-55, O3M-56, issue/rev. 1 / 1

[R15] O3M SAF ORR VALIDATION REPORT, O3M-82, issue/rev. 1 / 1

[R16] Yang, K., N. Krotkov, A. Krueger, S. Carn, P. K. Bhartia, and P. Levelt (2007), Retrieval of Large Volcanic SO_2 columns from the Aura Ozone Monitoring Instrument (OMI): Comparisons and limitations , JGR 112, D24S43, doi:10.1029/2007JD008825

[R17] Clarisse, L., Hurtmans, D., Clerbaux, C., Hadji-Lazaro, J., Ngadi, Y., and Coheur, P.-F.: Retrieval of sulphur dioxide from the infrared atmospheric sounding interferometer (IASI), Atmos. Meas. Tech., 5, 581-594, 2012.

[R18] Absorbing Aerosol Index Validation Report http://o3msaf.fmi.fi/docs/vr/Validation_Report_ARS_AAI_Jun_2013.pdf.

[R19] Absorbing Aerosol Index ATBD: http://o3msaf.fmi.fi/docs/atbd/Algorithm_Theoretical_Basis_Document_ARS_Jun_2013.pdf

[R20] Absorbing Aerosol Index Product User Manual: http://o3msaf.fmi.fi/docs/pum/Product_User_Manual_ARS_Sep_2013.pdf

[R21] LSA-SAF Snow Cover Product Webpage: http://landsaf.meteo.pt/algorithms.jsp;jsessionid=F05A2D5B0715F079381912AFCF3E07BA?seltab=4&starttab=4

[R22] LSA-SAF Snow Cover Product User Manual: http://landsaf.meteo.pt/GetDocument.do?id=280

[R23] LSA-SAF Snow Cover Product ATBD: http://landsaf.meteo.pt/GetDocument.do?id=279

[R24] LSA-SAF Snow Cover Product Validation Report: http://landsaf.meteo.pt/GetDocument.do?id=289

[R25] Siljamo, Niilo, Otto Hyvärinen, 2011: New Geostationary Satellite–Based Snow-Cover Algorithm. *J. Appl. Meteor. Climatol.*, **50**, 1275–1290.

[R26] H-SAF H10 (SN-OBS-1) Snow Mask Product Webpage: http://hsaf.meteoam.it/description-sn-obs-1.php

[R27] H-SAF H10 (SN-OBS-1) Snow Mask Product User Manual: http://hsaf.meteoam.it/documents/PUM/SAF_HSAF_PUM-10_1_1.pdf

[R28] H-SAF H10 (SN-OBS-1) Snow Mask Product ATBD: http://hsaf.meteoam.it/documents/ATDD/SAF_HSAF_ATBD-10_1_1.pdf

[R29] H-SAF H10 (SN-OBS-1) Snow Mask Product Validation Report: http://hsaf.meteoam.it/documents/PVR/SAF_HSAF_PVR-10_1_1.pdf

[R30] Surer, S. and Akyurek, Z., 2012. Evaluating the utility of the EUMETSAT HSAF snow recognition product over mountainous areas of eastern Turkey. *Hydrological Sciences Journal*, 57 (8), 1–11.

[R31] Aerosol CCI product ATBD: http://www.esa-aerosol-cci.org/?q=webfm_send/226

[R32] Aerosol CCI product Validation Report: http://www.esa-aerosol-cci.org/?q=webfm_send/477

[R33] GlobSnow Snow Extent Product Webpage: http://www.globsnow.info/index.php?page=Snow_Extent

[R34] GlobSnow Snow Extent Product User Manual: http://www.globsnow.info/se/GlobSnow_SE_product_readme_v1.2.pdf

[R35] Snow Extent Algorithm Development Document 2: (Internal validation report) http://www.globsnow.info/swe/Aux_data/Deliverables_GS2/GS2_DEL_11_WP2.2_v1_r2_final.pdf

[R36] Sari Metsämäki, Olli-Pekka Mattila, Jouni Pulliainen, Kirsikka Niemi, Kari Luojus, Kristin Böttcher, An optical reflectance model-based method for fractional snow cover mapping applicable to continental scale, Remote Sensing of Environment, Volume 123, August 2012, Pages 508-521.

[R37] GHG-CCI documents available from http://www.esa-ghg-cci.org/?q=node/95

[R38] GHG-CCI Algorithm selection report (ASR) of GHG-CCI, final version.

[R39] Product validation and Intercomparison report (PVIR) of GHG-CCI, version 2.0.

[R40] B. Dils, M. Buchwitz, M. Reuter, O. Schneising, H. Boesch, R. Parker, S. Guerlet, I. Aben, T. Blumenstock, J. P. Burrows, A. Butz, N. M. Deutscher, C. Frankenberg, F. Hase, O. P. Hasekamp, J. Heymann, M. De Mazière, J. Notholt, R. Sussmann, T. Warneke, D. Griffith, V. Sherlock, D. Wunch: The Greenhouse Gas Climate Change Initiative (GHG-CCI): Comparative validation of GHG-CCI SCIAMACHY/ENVISAT and TANSOFTS/ GOSAT CO2 and CH4 retrieval algorithm products with measurements from the TCCON network, submitted to AMTD, 2013

[R41] Ozone_CCI documents available at http://www.esa-ozone-cci.org/?q=documents

[R42] Ozone_CCI Product Specification Document (PSD), ref: Ozone_cci_ PSD_3.1, version 3.0

[R43] Ozone_CCI Product Validation and Selection Report (PVASR), ref: Ozone_cci_PVASR, issue 1.5

[R44] Brühl, C., S. R. Drayson, J. M. Russell III, P. J. Crutzen, J. McInerney, P. N. Purcell, H. Claude, H. Gernand, T. McGee, I. McDermid, and M. R. Gunson: HALOE Ozone Channel Validation, J. Geophys. Res., Vol. 101, No. D6, pp. 10,217-10,240, April 1996.

[R45] Nazaryan, H., M.P. McCormick, and J.M. Russell III: New studies of SAGE II and HALOE ozone profile and long-term change comparisons, J.Geophys. Res., 110 (D9), doi: 10.1029/2004JD005425, 2005.

[R46] Wang, H. J., D.M. Cunnold, L.W. Thomason, J.M. Zawodny, and G.E. Bodeker, "Assessment of SAGE version 6.1 ozone data quality", J.Geophys. Res., 107, doi:10.1029/2002JD002418, 2002.

3. Examples of Current Validation Practices

In the following, the current validation practices of different initiatives and programmes are introduced with the focus mentioned in section 1.3. Attention should be paid to those initiatives/programs from which details for each focus point is lacking. In that case, the contents of that initiative/program relevant to Cal/Val will be addressed in general without itemizing each focus. For some initiatives/programs, certain ECV will be used as an example to demonstrate the validation practice, while others not. More details and reference documents are referred to the supplement to this report[7], which will be kept up to date whenever the feedback regarding significant omissions was provided.

3.1. Copernicus

3.1.1. Global Land vegetation products (e.g. fAPAR)[8]

General Validation Procedures

The generic approach is based on the CEOS/WGCV LPV guidelines. This includes a protocol on the methods for validating the fAPAR data set, as well as guidelines for in-situ data processing accounting for up-scaling effects. Currently, a protocol is being established for the collection of in-situ data for fAPAR. All other vegetation ECVs of the Global Land Service are validated using the same validation practice as for fAPAR. The validation and quality monitoring is performed at different levels, i.e. the self-assessment and the independent assessment, as part of the external review every 6 months.

Generation of ECV

The fAPAR TCDR is entirely based on the normalized surface reflectance for a standard observational configuration.

Validation Datasets

Reference global biophysical products:

- CYCLOPES (CYCV31)
- MODIS (MODC5)
- GLOBCARBON (GLOV2) (only LAI)
- JRC FAPAR (http://fapar.jrc.ec.europa.eu/)
- GEOLAND V0 (GEOV0)
- Ground data

Validation Process

The specific approach consists of the evaluation of the listed criteria below, while taking into account the CEOS/WGCV LPV protocols. The validation is performed by a different partner in the consortium than those involved in the algorithm specification, implementation or operational processing to guarantee the independent nature of the validation.

[7] WP3 Report, 2014, Supplement to De3.32 “Generic Validation Strategy for CDRs/ECVs”
[8] http://land.copernicus.eu/global/products/fapar

The analysis is performed at regional, continental and global scale. For a number of evaluation criteria, the fAPAR estimates of the BELMANIP-2 sites were used only (Baret et al., 2006) according to the CEOS LPV validation protocol.

(1) Spatial Consistency Analysis

(2) Global Statistical Analysis

(3) Temporal Consistency Analysis

(4) Direct Validation

(5) Cross-Cutting Evaluation using a Land Data Assimilation System (LDAS)

Concerning the inter-comparison with other global fAPAR data sets, the OLIVE tool was developed, that provides the assessment against a number of thoroughly validated global fAPAR data sets (see previous section) for the BELMANIP-2 network of sites. (http://calvalportal.ceos.org/olive)

Error Characterization

The fAPAR product is delivered with two additional quality indicator layers, i.e. the quality flag and the uncertainty estimate. The quality flags are specified and related to different steps in the processing. The uncertainty estimate[9] is a quality indicator of the performance of the retrieval of fAPAR for a given pixel.

3.1.2. MACC II Atmospheric Composition[10]

The validation of atmospheric composition products consists of two main lines of inquiry:

(1) Direct comparison of retrieved gas column or aerosol property against the same quantity observed from ground-based instruments (typically spectrometers).
(2) Inter-comparison of retrieved atmospheric quantity with either the same quantity derived with other retrieval algorithms from the same or different satellite instruments. Some validation methods compare the current retrieval algorithm to a known reference algorithm with high accuracy, both having the exact same satellite data as input ("operational vs. scientific algorithm comparison")

While direct comparison is usually seen as the most desirable validation method, there are often substantial associated limitations. The satellite data may be temporally mismatched, have such coarse spatial resolution as to question the comparability to in-situ data, or there may be simply very few in-situ data available of the desired quantity. The lack of continuous, quality-controlled in-situ observations is lamented in several validation studies.

The satellite product inter-comparison has the benefit of providing information on the spatial structure of retrieval differences over a longer time period, potentially providing valuable information on the strength and weaknesses of the retrieval method(s). At the same time, great care needs to be taken to assess if the inter-comparison is first of all valid, as data from different satellite sensors may observe different parts of the atmosphere or be spectrally or temporally mismatched so as to inhibit comparability. The reference data source in the

[9] Details can be referred to [R1], [R2] and [R3].
[10] http://www.gmes-atmosphere.eu

inter-comparison also needs to be well documented in terms of its retrieval quality. Satellite instrument degradation (e.g. GOME-2A) is difficult to compensate for, and can seriously hamper inter-comparability[11].

In summary, both in-situ reference validation and inter-comparison of different satellite-based retrievals are common validation strategies for atmospheric composition products made by European providers. The two methods are generally seen as yielding complementary information on retrieval quality. Yet, circumstances often force the validation study to focus on one avenue of investigation. It is therefore noted that there is a need not only for more quality-controlled in-situ observations, but also a need to maintain multiple datasets on the same physical quantities to allow for an inter-comparison in the first place.

3.1.3. MyOcean II Baltic Sea Regional NRT SAR Based Sea Ice[12]

General Validation Procedures

Validation relies on the direct comparison with reference data for all three datasets: ice concentration, ice thickness and ice drift.

Generation of ECVs

(1) 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.
(2) 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.
(3) Ice concentration is based on the digitized FMI ice charts produced by ice analysts.

Validation Datasets

(1) Drill measurements from icebreakers in Baltic Sea are used for ice thickness validation.
(2) Ice concentration reference data is produced with ASI algorithm by University of Bremen with SSMI/S data.
(3) Ice drift reference data is collected with buoys in Gulf of Bothnia and Gulf of Finland.

Validation Process

(1) Required assignments are made to match observations of ice thickness for comparison
(2) Differences between product and reference data is calculated
(3) Ice concentration reference data (BR maps) are prepared for comparison using 3x3 homogenous windows to eliminate the possible errors due to the re-rectification
(4) Distributions of the difference between the maps is inspected
(5) Statistical analysis is done for the comparison
(6) Drift information from the buoys are gathered for relevant locations

[11] It is noted that the inter-comparison here is for validation practice, which is necessary to have the reference data established and characterized. Otherwise, it would be a pure inter-comparison practice, for which triple collocation approach is recommended.
[12] http://www.myocean.eu/

(7) Ice drift estimates are divided into two categories: short and long drift
(8) Motion magnitude is calculated for short drifts
(9) Motion magnitude and direction is calculated for long drifts
(10) Statistical analysis is done for the comparison of these values with reference data

Error Characterization

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.

3.1.4. GMES In-Situ Coordination

The GMES in-situ coordination (GISC) was an FP7 funded project running from January 2010 till October 2013.

The objectives of the project were i) to enable data providers or network of data providers to provide the required in-situ data for GMES/Copernicus, ii) to identify priorities and requirements of in-situ data for GMES/Copernicus core services and iii) to integrate in-situ assets and networks into long-term sustainable frameworks for GMES/Copernicus services.

Among the key deliverables of GISC was the identification and classification of the in-situ data required by the GMES/Copernicus services. The report on in-situ data requirements for a variety of ECVs can be accessed at http://gisc.ew.eea.europa.eu/deliverables/d2.1.pdf.

In April 2013 a GISC workshop 'Monitoring Matters' took place in Copenhagen. The workshop report is available at http://gisc.ew.eea.europa.eu/deliverables/monitoring-matters-workshop-report.pdf/download.

3.2. EUMETSAT SAFs

EUMETSAT SAFs include CM SAF, H SAF, LSA SAF, O3M SAF, OSI SAF, NWP SAF, ROM SAF and NWC SAF[13], among which CM SAF is aiming to generate and archive high-quality climate datasets on a continuous basis. Therefore, in the following, only the validation strategy of CM SAF is introduced. For other SAFs, they are designed for other specific climate application areas and readers are referred to the SAFs website[5] for more details.

The validation process is part of the CM-SAF CDR processing subsystems. The independent external review processes of the processing flow are part of the validation activities in CM-SAF. For example, the external review process includes requirements review (i.e. definition of product requirements, input data and validation methods), dataset readiness review and peer reviewed papers.

The validation is an integral part of the product characteristic. The validation report is reviewed and then archived, leading to committed service for CM-SAF. The annual review of products to ensure quality per product is implemented to evaluate product specific metrics, define service specifications and perform automatic validation. A synthesis report is published on the website.

[13] http://www.eumetsat.int/website/home/Satellites/GroundSegment/Safs/index.html

The validation procedures in CM-SAF include:

(1) The independent validation of the CDRs from satellite data, by using the part of in-situ data that are not merged with satellite data (i.e. independent reference source);
(2) The implementation of an operationalized validation of products in order to sustain established procedures and methods (i.e. 'validation facility'); and
(3) An initial analysis of consistency, stability and homogeneity between satellite instruments.

In general, there are two processes for validation of CDRs (i.e. validation-dependent product): automatic technical validation during processing (i.e. outliers, undefined, retrieval not successful), and scientific validation and evaluation (i.e. versus reference observations and available independent satellite data sets). In terms of scientific validation and evaluation, the importance of traceability to SI standards where possible/appropriate, is highlighted. In addition it is indicated that there is a need to regularly quantify the current state of the art in products being constructed for climate applications.

3.3. ESA CCI Projects

In the following, only the high level information relevant to validation strategy is presented. It is applicable to all CCI projects. For more details, readers are referred to ESA CCI website[14].

General Validation Procedures

It includes two levels of activities:
(1) Validation and error characterisation internal to each CCI Project;
(2) Externally, ECV data product assessment through the Climate Modelers User Group (CMUG);

Generation of ECVs

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.

Validation Datasets

It includes requirements as listed below:

(1) For validation, CCI project teams shall use in-situ or suitable reference datasets that have not been used during the production of their CCI products;
(2) The independence of the geophysical process shall be considered. It shall ensure that if a particular auxiliary dataset is used in the production of their CCI products then the same dataset is not used in the validation and, if required, alternative auxiliary data are used.

When such unique independent reference data are not available, data for validation should be selected to ensure complete coverage of the various spatial and temporal scales in each CCI product. The selection of validation data sets should follow different levels of rigor depending on the level of independence of each data set, thus making sure that some level of

[14] http://www.esa-cci.org/

confidence can be given to every output product. Each CCI product should contain an indication of the level (/confidence) in the data quality resulting from the validation process.

Validation Process

(1) It shall ensure that the validation is carried out (or at least verified) by staff not involved in the final algorithm selection; ideally the validation of the CCI products should be carried out by external parties, i.e. by staff/institutions not involved in the production of the ECVs products.
(2) It shall use established, community accepted, traceable validation protocols where they exist. If such protocols do not exist then CCI projects may adapt existing protocols if appropriate and in any event shall offer their final protocol for future community acceptance;
(3) To assure the quality of an ECV data product, and that the product specifications are reached, a validation process shall be an ongoing process that takes into account requirements and responses from users. This should be fully documented in the Product Validation Plan

Error Characterization

The "Uncertainty Characterization" document from all CCI projects should follow a common table of contents as indicated by ESA CCI project guidelines.

3.4. CEOS WG CalVal

For CEOS-WGCV, the strategic objectives include[15]:

- Sensor-Specific Cal/Val: to document and establish forums for the assessment and recommendation of current techniques and standards for pre- and post-launch characterizations and calibration.

- Geophysical Validation: To document and establish forums for the assessment and recommendation of techniques for validation of geophysical parameters derived from Earth Observation satellite systems.

3.5. WMO Guidelines/Supplements/Requirements

The Commission for Instruments and Methods of Observations (CIMO) aims at facilitating international standardisation and compatibility of instruments and methods of observations to improve quality of products and services. The CIMO issues regular updates of the WMO guide to meteorological instruments and methods of observation. The guides cover several domains (atmosphere, ocean, land) and techniques (in-situ and satellite observations).

The WMO Integrated Global Observing System (WIGOS) issues GCOS guides providing requirements and guidelines to operate observations networks. For example, the GCOS Reference Upper-Air Network (GRUAN) guide contains recommendations on how to validate satellite retrievals[16].

[15] http://www.ceos.org/index.php?option=com_content&view=category&layout=blog&id=75&Itemid=113
[16] http://www.wmo.int/pages/prog/gcos/Publications/gcos/171.pdf

3.5.1 Satellite

The most relevant and comprehensive set of specific user requirements, supplements, guidelines and principles for the generation of satellite-based datasets and products can be found from the reference listed as below:

GCOS Documents
GCOS-82, 2003, The Second Report on the Adequacy of the Global Observing Systems for Climate in Support of the UNFCCC, 74 pp. GCOS-117, 2008: Future Climate Change Research and Observations: GCOS, WCRP and IGBP. GCOS-138, 2010, Implementation Plan for the Global Observing System for Climate in Support of the UNFCCC (2010 Update), GCOS-138 (GCOS-176, GTOS-84, WMO/TD-No. 1523). GCOS-143 (WMO/TD No.1530), 2010, Guideline for the Generation of Datasets and Products Meeting GCOS Requirements. An update of GCOS-128, WMO/TD-No. 1488 including in-situ datasets and amendments GCOS-153, 2011, Report of WOAP workshop on Evaluation of Satellite-Based Global Climate Datasets, Meeting Report, 43 pp. GCOS Climate Monitoring Principles also included in Resolution 9 WMO Cg-XIV (WMO No. 960)

From the above, calibration/validation was identified as an element embedded in the CDR processing chain.

3.5.2. In-Situ

Regulatory material provided by WMO and GCOS covers but is not limited to:

- Actual measurement setup recommendations
- Calibration procedures (and possible link to SI)
- Climate data management
- Current day metadata capture practice
- Quality control procedures
- Exchange format recommendations
- Data rescue activities
- Interpolation methods
- Validation approaches

In the following, relevant documents are listed:

WMO Regulatory Material

Technical regulations: WMO-No. 49 **(2011 edition updated in 2012), Volume I: General Meteorological Standards and Recommended Practices.**
Manuals: WMO-No. 306 **(2011 edition, updated in 2012), Manual on Codes, Volume I.** WMO-No. 386 **(2009), Manual on the Global Telecommunication System, Volume I** WMO-No. 485 **(2010 edition, updated in 2013), Manual on the Global Data-processing and Forecasting System** WMO-No. 544 **(2010), Manual on the Global Observing System, Volume I** WMO-No. 1060 **(2012), Manual on the WMO Information System**
Guides: *WMO-No. 8* (2008 edition, updated in 2010), Guide to Meteorological Instruments and Methods of Observation (CIMO Guide 2008 edition, Updated in 2010). → Guidance on measurement techniques with recommendations for quality management including metadata documentation *WMO-No. 100* (2011), Guide to Climatological Practices *WMO/TD-No. 1376* (2007), Guidelines on Climate Data Management → Guidance using standardized software tools *WMO/TD-No. 1186* (2003), Guideline on Climate Metadata and Homogenization → includes guidance on best practice for required metadata, site documentation and data analysis tools WMO/TD-No. 1210 (2004), Guidelines on Climate Data Rescue

GCOS Documents
GCOS-143 (WMO/TD No.1530), 2010, Guideline for the Generation of Datasets and Products Meeting GCOS Requirements. An update of GCOS-128, WMO/TD-No. 1488 including in-situ datasets and amendments → Guidance on documentation, quality assessment and publication of work related to CDR/ECV generation. Includes terminology of standards (stability, calibration, validation, traceability etc.) GCOS-138, 2010, Implementation Plan for the Global Observing System for Climate in

Support of the UNFCCC (2010 Update)

GCOS Climate Monitoring Principles also included in Resolution 9 WMO Cg-XIV (WMO No. 960)

3.6. In-situ Networks Climate Monitoring

Calibration and validation activities for in-situ-based datasets differ substantially from practices for satellite-based datasets. Usually, for in-situ ECVs the "quality control procedures and reference procedures" correspond to the terms of Cal/Val.

Linked to this, but also often treated separately, are following issues:

(1) Representativeness of stations
(2) Homogeneity
(3) Gridding
(4) Usual practice of combining of changing amounts of station density with data of various qualities
(5) Long-term stability
(6) Insufficient metadata

Information on this helps end-users to understand how in-situ data ECV were generated (e.g. starting from collection to data store) and to be used for which purpose by those descriptions.

The WMO Guide to Climatological Practices (refer to section 2.5.2) addresses the following issues regarding quality control, interpolation and validation:

Quality control

a. Procedures:
   - Error detection by objective, automated screening of data and manual review of the automated output
   - Automatic and semi-automated error-flagging procedures on data type (original data, corrected data, reconstructed data or calculated value), stage of validation (missing, eliminated, doubtful, validated etc.) and acquisition method (for data from multiple sources).
b. Documentation:
   - Documentation of quality control procedures and algorithms for each processing stage
   - User access for assessment of validity of observation and impact of changes in procedures on validity, continuity or homogeneity of data record
c. Types of error:
   - Metadata errors, such as incorrect station identifier or incorrect date stamp (often resulting in data errors)
   - Data errors, as results of instrumental, observer, data transmission, key entry and data validation process errors, changing data formats and data summarization problems
d. Format tests:
   - Checks for double entries or impossible format codes

- Procedures to eliminate or reduce format errors

e. Completeness tests:
   - Checks for missing data and flagging for future review
f. Consistency tests:
   - Internal consistency (checks on physical consistency with other elements and checks for consistency with definitions)
   - Temporal consistency (checks on the variation of an element in time)
   - Spatial consistency (comparison with observations at nearby stations)
   - Summarization tests (comparison of different summaries of data)
g. Tolerance tests:
   - Comparison against statistical thresholds

Estimation methods for the production of gridded datasets:

a. Mathematical estimation methods (e.g. inverse distance weighting, spline functions)
b. Estimation based on physical relationships (e.g. regression, discriminant and principal component analysis)
c. Spatial estimation methods (e.g. Kriging)

The combination of different methods (e.g. the use of a regression model and interpolation of residuals) is a common practice. Often, the production of gridded datasets follows a step-wise approach incorporating different estimation procedures.

Validation of gridded datasets:

a. Split validation (testing the methodology using a smaller subset excluded in the estimation procedure)
b. Cross-validation (repeated removals of observations from the sample and analysis of residuals between observed and estimated values)

Generally, validation strategies for in-situ data must be regarded separately for station measurements (point data) and gridded datasets (usually derived by interpolation procedures). Validation of station data using independent sources, e.g. validation of a temperature reading at one station, is generally not feasible (or at least not a common practice). Hence, other issues like homogeneity and data quality play a larger role. For gridded datasets, other issues become more important, such as the representativeness of the input data, number of stations per grid cell, interpolation methodology, inter-comparison with other gridded data or validation with independent ground stations etc.

However, there is no community-wide standard, best practice or agreement on protocols for quality control/validation procedures among the different in-situ-based activities.

Generally, gaps and research needs are identified as follows:

1. upscaling of in-situ measurement and downscaling of national statistics needs international efforts in developing a common terminology and definitions of accounting schemes
2. representativeness of stations
3. gridding procedures and connected uncertainties
4. the usual practice of combining of changing amounts of station density with data of various qualities (i.e., change of the observing system)
5. long term stability and reference procedures

6. homogenization procedures
7. insufficient metadata

The GCOS Adequacy Reports (GCOS-48, 1998; GCOS-82, 2002) as well as the 2010 Update of the GCOS Implementation Plan (GCOS-138, 2010) provide valuable information on gaps in observing systems for climate and actions needed to fully implement the GCOS. Both the adequacy report and the implementation plan are being updated and shall be available by end 2015 and 2016 respectively (FCCC/SBSTA/2012/5, §37). More details are referred to the supplement to this report[17].

3.7 Earth System Reanalyses

Reanalyses of the Earth system typically cover several elements. For example, the recent reanalysis from NCEP (CFSR[18]) covers atmosphere, land, ocean (surface and sub-surface), and sea-ice. The current ECMWF reanalysis (ERA-Interim[19]) covers atmosphere, ocean surface (waves), and land.

For the sake of the present discussion, an Earth system element is said to be covered if:

- it is represented by a model which attempts to reproduce variations from known physical laws or parameterizations (e.g., for sea-ice: a sea-ice model), and
- if the states produced by such model are updated for one or several geophysical variable(s) by corresponding observations of this Earth system element (e.g., for waves: satellite altimeter wave-height observations).

The output products of a reanalysis system include ECV datasets that can be considered as a blend of model data and observational data. When all ECV datasets produced by an assimilation-based reanalysis are the products of a single reanalysis-assimilation system, the system validation is necessary, but not sufficient.

Indeed, models of the various elements can be as far as coupled or fully integrated, but in general there are still significant disconnections between the various Earth system elements in the assimilation: analysis updates from observations (so-called increments) are typically computed separately for each Earth system element's data assimilation, and it is only during the model integration that all states are made physically consistent between one another. This point has consequences for the validation, which entails considering not only validation of the overall system but also the validation of the individual ECV datasets.

The processes involved in reanalysis production are validated by monitoring several metrics. For an assimilation-based reanalysis system, these involve the following components:

- the observations input
- the forcing or boundary datasets input
- the model configuration (for the various Earth system elements)
- the data assimilation system (in the various Earth system elements)

[17] WP3 Report, 2014, Supplement to De3.32 "Generic Validation Strategy for CDRs/ECVs"
[18] Saha, S., and Coauthors, 2010: The NCEP Climate Forecast System Reanalysis. Bull. Amer. Meteor. Soc., 91, 1015–1057. http://dx.doi.org/10.1175/2010BAMS3001.1

[19] Dee, D.P., and Coauthors, 2011: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q.J.R. Meteorol. Soc., 137: 553–597. http://dx.doi.org/10.1002/qj.828

For all of the above, the reanalysis producers generally rely on reports and published papers to make their choices regarding the maturity and reliability and to decide whether the component selected is suitable for their own application (bearing in mind that a production covers several years or decades and may take months or years to complete). Once this 'input checking' has been done, all reanalysis components are validated during production as follows.

The set of forcing and boundary dataset inputs is validated by generating time-series of the mean state (averaged across some spatio-temporal domain) and its fluctuations (time increments and standard deviations within the spatio-temporal grid). In the case of ensemble datasets, the variations in ensemble spread are compared against expectations (e.g. expecting lower spread in areas of lower natural variability and/or greater observation concentration).

The observation input is validated by sub-setting the various observation types and variables and plotting their general characteristics (again, using some form of averaging to reduce the amount of information to a few selected granules). An important validation is to verify that the input is as expected, i.e. by inspecting timelines of observation sources. In particular the reanalysis producers validate that the observations get used as expected (number of observations in total, accepted, rejected, and separating by reason to identify potential problems)[20].

As a reanalysis production can last months or years and be ported across several computing environments, the model configuration is checked for unwanted changes. This validation usually happens at regular intervals during the lifetime of a reanalysis production, by comparing the production logs with a reference log (such as one produced in the early stages of production). The behaviour of the underlying forecast model is itself usually validated by running a long model integration without assimilation (similar to so-called AMIP integrations[21]), and inspecting the climate trends produced, as well as global energy and water budgets.

The data assimilation system is validated by inspecting the basic metrics such as, for a variational system, the observational cost function and its reduction during the analysis minimization, and validating them against expectations. All quantities produced by the assimilation system are monitored, such as analysis increments, reduction in the distance to assimilated observations, observation bias corrections, background errors when these are dynamically updated, and reduction in ensemble spread caused by the analysis in the case of ensemble productions[8]. In that respect an important validation is the comparison between assumed (background and observation) errors and effectively observed background

---

[20] Poli, P., and Coauthors, 2013: The data assimilation system and initial performance evaluation of the ECMWF pilot reanalysis of the 20th-century assimilating surface observations only (ERA-20C). ERA Report Series 14, available from ECMWF, Shinfield Park, Reading, UK, 59 pp. http://www.ecmwf.int/publications/library/do/references/show?id=90833

[21] Hersbach, H., C. Peubey, A. Simmons, P. Poli, D. Dee and P. Berrisford, 2013: ERA-20CM: a twentieth century atmospheric model ensemble. ERA Report Series 16, available from ECMWF, Shinfield Park, Reading, UK, 44 pp. http://www.ecmwf.int/publications/library/do/references/show?id=90989

departures in the system, treating the observations as an imperfect reference (i.e. taking into account its errors).

The ECV datasets produced by a reanalysis system are then compared against independent validation datasets. Such datasets include typically CRUTEM (for surface temperatures), GPCC and GPCP (for precipitation), KNMI's Multi Sensor Reanalysis (for ozone). Validation is also conducted by cross-referencing departures before and after assimilation for the assimilated observations. Important conclusions about the relative biases in the various observations can still be retrieved from such detective/investigative work[22].

4. Analysis of Current Validation Practices

4.1. Essentials of Validation Procedure for CDRs/ECVs

4.1.1. Satellite Products

Based on the current validation practices of different initiatives/programmes, indicated in section 2, the validation procedure for CDRs/ECVs may be identified as: direct validation (direct comparison against reference data that independent and superior in accuracy relative to the monitored data record (CDR/ECV), e.g. in-situ data or high quality satellite observations), indirect validation (inter-comparison against reference data of similar products that already exists, are not independent, and have similar quality as the monitored data record) and scientific quality check (e.g. spatial-temporal consistency, product continuity, precision and smoothness). This implies the establishment of reference data, which should be kept as much independent as possible. The first essential of validation procedure is therefore:

A. The specification of the source of reference data

It is important to document how reference data are produced, especially for the satellite-based reference data. For example, for indirect validation, satellite products will be compared with each other. The first of all is to check whether different satellite sensors are measuring the same geophysical parameters. Without the detailed documentation on how reference data being produced, it is difficult to understand that 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 reference. The typical disadvantage, in terms of its usage for validation, is the mismatch of scales of satellite products and in-situ observations. Another point is that there are simply very few in-situ data available for the desired ECVs/CDRs (e.g. lack of continuous, quality-controlled in-situ observations). It is therefore needed to identify:

A-1. Requirements of in-situ sites (e.g. representativeness, homogeneity, long-term stability etc.);

A-2. Guidelines for in-situ data processing (e.g. especially accounting for up-scaling effect, may also include detection of outliers, undefined, record not successful);

[22] Simmons, A.J., P. Poli, D.P. Dee, P. Berrisford, H. Hersbach and C. Peubey, 2013: Estimating low-frequency variability and trends in atmospheric temperature using ERA-Interim. ERA Report Series 15, available from ECMWF, Shinfield Park, Reading, UK, 48 pp. http://www.ecmwf.int/publications/library/do/references/show?id=90834

A-3. Protocol for in-situ data collection (e.g. metadata format etc.)

Note that direct validation is less hampered by above described disadvantages when, instead of in-situ observations, satellite or flight campaign based observations serve as reference and these observations are independent and superior in quality as compared to the monitored ECV/CDR product.

With the establishment of reference data and technical data processing, the specific validation procedures should then be identified, which needs:

B. Definition of validation methods (e.g. direct/indirect/cross-cutting comparisons, spatial-temporal consistency analysis, large statistic, case studies, etc.)

It is well accepted that different ECVs/CDRs have different validation procedures, as well as the establishment of reference data. Nevertheless, the above points are essentials when implementing the validation process (e.g. for different ECVs/CDRs), which should be practiced with essential strategies/mechanisms including:

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

As a final step in the validation procedure, an effort should be made to assess the consistency of the monitored products (ECV/CDR) with products of other physical variables. (For example, the consistency of the surface soil moisture derived from ASCAT with the precipitation products). It is to avoid the physical mismatch in the first place:

F. Consistency check (e.g. an initial analysis of physical consistency among different products that are independent from each other)

The self-assessment serves as the internal validation process to ensure quality per product (e.g. technical quality, scientific quality and consistency across variables), while the independent assessment is to ensure the unbiased validation process to ensure quality per product. The entities to implement independent assessment should be a separate/external entity other than the one generating the product. Such entities are entitled, at the same time, to implement the external review process (e.g. evaluation) of the internal validation process. The external evaluation process shall include the review of all validation documents, for example, in CM-SAF, the requirements review (i.e. definition of product requirements, input data and validation methods), dataset readiness review and peer reviewed papers.

On top of the six points (A-F) listed above, validation facility should be developed to implement an operationalized validation of products (e.g. to sustain established procedures and methods), in order to perform automatic validation (e.g. the validation process become less and less time-consuming and more and more mature). At operational/mature validation level, the internal validation should be reviewed regularly (e.g. semi-annually or annually).

It is noted that by the nature of validation independency is the most essential element of validation strategy. For example, the reference data and the assessment of products should be kept as much independent as possible. It therefore comes to a point to understand the scale of independency of the whole validation process. In regards to reference data, the potential independency levels can be scaled based on:[23, 24]

1. Independent ('true') in-situ data;
2. Other in-situ data;
3. Airborne campaign datasets for medium-scale comparisons;
4. Other satellite datasets for large-scale comparisons;
5. Historic datasets, trends, climatology for large-scale comparisons;
6. Impact studies using other products (e.g. consistency among different variables);
7. Cross-cutting comparison using data assimilation approach, standard metrics.

For independent assessment and external review process, the independency is represented by the entities implementing assessment and reviewing, which may be potentially scaled among:

1. The 'volunteer' external parties (e.g. pursuit of scientific excellency) that have no any connection/involvement in the production of products
2. The contracted external parties (e.g. under a separate different contract than the one producing products)
3. Other external parties (e.g. end-users, stake holders, commercial companies)

4.1.2. In-situ Products

The essential of validation procedures for in-situ based products is 'quality control', which has been well addressed by the WMO guide to climatological practices. The general validation strategies should be regarded separately for station measurement (point data) and gridded datasets (e.g. derived by interpolation procedure). Validation of station data using independent sources is generally not feasible and is not a common practice, in case of which homogeneity and data quality are essential to ensure quality of products. On the other hand, for gridded datasets, it is more important to check representativeness of the input data, number of stations per grid cell, interpolation approaches, inter-comparison with other gridded data and validation with independent ground stations.

It is noted that there is no community-wide standard, best practice or agreement on protocols for quality control/validation procedures among the different in-situ-based activities.

4.1.3. Reanalysis Products

The essential components involved in reanalysis production are validated by monitoring several metrics, which include the observation inputs, the set of forcing and boundary dataset inputs, the model configuration and the data assimilation system. The ECV datasets produced by such a validated reanalysis system can then be compared with independent validation datasets, which are not involved in the production of reanalysis outputs. It actually can then follow the essentials of validation strategy for satellite products.

4.2. Maturity Matrix

In WP2, system maturity matrix is developed to offer a systematic mean to assess if the data record generation procedure follows best practises. It includes six major categories where assessments are made:

---

[23] ESA Climate Change Initiative: CCI Project Guidelines, 2010
[24] http://lpvs.gsfc.nasa.gov/index.html

1. Software readiness
2. Metadata
3. User documentation
4. Uncertainty characterisation
5. Public access, feedback, and update
6. Usage

The maturity of the validation procedures adopted covered in "user documentation" and "uncertainty characterisation". To evaluate the extent to which the product has been validated, the 6 maturity scores are showed as below:

Score	Description
1	None
2	Validation using external reference data done for limited locations and times
3	Validation using external reference data done for global and temporal representative locations and times
4	Score 3 + (Inter)comparison against corresponding CDRs (other methods, models, etc)
5	Score 4 + data provider participated in one inter-national data assessment
6	Score 4 + data provider participated in multiple inter-national data assessment and incorporating feedbacks into the product development cycle

It is expected that the essentials of validation strategy identified in section 3.1 can be assessed by the above table.

5. Proposed Generic Validation Strategy for CDRs/ECVs

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

1. The generation of independent reference datasets
2. Assessing independency levels of reference datasets (see section 4.1.1)
3. Self-assessment
4. Independent assessment
5. External review/evaluation of self-assessment validation practice
6. Assessing independency levels of point 4 and 5 (see section 4.1.1)
7. Consistency check for inter-related CDRs/ECVs
8. Sustaining established procedures and methods

The above generic validation strategies/procedures differ from protocols (e.g. see A-1 to A-3 in section 4.1.1), and are abstracted on the basis of analysing essentials of current validation practices (e.g. see section 3 and section 4). It serves as a checking-list for identifying how far the validation procedure has been approaching the current identified completeness (e.g. 'best practice').

For each procedure identified above, the specified protocol shall be defined and followed. For example, for the point 1, in the case of using in-situ data as references, the protocols listed in A-1 to A-3 (in section 4.1.1) should be considered. On the other hand, in the case of using satellite or flight campaign based observation as references, the protocol with similar concept as for in-situ data should be defined. It is imperative to acknowledge the importance of understanding how the reference data are established to ensure its transparency to enable traceability. In fact, each procedure from the above should be made as transparent as possible to ensure traceability.

Point 1-3 are internal validation procedures. For the point 2, the independency scale of reference data shall be defined (see section 4.1.1). Similar to point B as indicated in section 4.1.1, the point 3, self-assessment, requires the need to define and document the validation methods, including and not limited to validation plans (e.g. case studies, approaches) and error characterization (e.g. error propagation, error budget).

The external validation procedures are composed of point 4-6. The independent assessment (i.e. the point 4) holds the same demand as the self-assessment does, which is to document the validation methods. One step further is that it necessitates the following point 6 to assess the independency of the external assessments. Additionally, one of the external validation procedures (i.e. the point 5) facilitates the independent review/investigation on the self-assessment validation practices, to check the transparency and traceability of the validation activities. The independency of the point 5 should be checked as well by the point 6, in which the independency level of external assessments can be represented by the entities implementing assessment and reviewing (see section 4.1.1). For the point 5, the dataset release readiness review mechanism from CM-SAF (www.cmsaf.eu) can be followed as the starting point.

Except for the internal and external validation procedures, point 7 and 8 serve as the synergic components of the validation procedures. The consistency check (e.g. the point 7) can help identify gaps in gridded information in the interactions and exchanges between the domains atmosphere, ocean and land. It is to check the physical consistency of the monitored products (ECV/CDR) with products of 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 and precipitation data. For the point 8, it aims to achieve an operational validation level, at which validation activities and data release were regularly implemented. Both point 7 and 8 shall be documented in details to enable traceability.

6. Conclusion and Recommendations

6.1. Satellite based

A review of the current validation practices of European data providers was performed. The main satellite-derived ECV products are provided by EUMETSAT (SAFs and CAF), by ESA (ESA-CCI projects), and by Copernicus projects. Based on the analysis of current practices, the essentials of general validation strategy are identified. By the nature of validation, independency is the most essential element. The proposed generic validation strategy is based on this nature to emphasize independency at each step. It is recommended to scale different independency levels for the reference datasets (i.e. point A) and the external reviewers (i.e. point D& E). This may contribute to the assessment of the completeness of validation process. Such assessment can be subsequently implemented by using the developed tool (System Maturity Matrix) in WP2.

Assessing the consistency between products highlights the possible synergies between various products, which in the end may help improve the overall quality of the ECV products. In the field of atmospheric variables, a high level of integration and international validation standards was achieved thanks to the leadership of WMO. The validation guideline for GRUAN is a good example of this achievement. The use of data assimilation, i.e. the integration of various ECV products into models for reanalyses 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. There is a strong heritage related to the data assimilation techniques for the atmospheric and ocean variables. The importance of such cross-cutting validation, i.e. check the consistency between distinct ECV products (e.g. LAI and surface albedo over land), even across domains (e.g. soil moisture and precipitation) has to be emphasized. The data assimilation systems used to produce reanalyses 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 as the impact of using the products in applications on the production is the key information for further improvement of the products, and as the users can provide in-situ observations from channels that cannot be accessed by the data producers. At the same time, generic global models able to ingest the satellite products (e.g. in reanalyses) are needed to assess the consistency of the products at a continental or global scale.

6.2. In-situ based

Among the different in-situ-based activities, there is no community-wide standard, best practice or agreement on protocols for quality control/validation procedures. The general validation strategies should be regarded separately for station measurement (point data) and gridded datasets (e.g. derived by interpolation procedure), where the essentials of quality control procedures may vary from homogeneity and data quality to number of stations per grid cell. In addition, it is anticipated that atmospheric, terrestrial and oceanic ECVs have special problems to be taken into account in the validation, so that completely identical methodology may not be advisable for all these ECV groups. The gaps and research needs in this area in terms of validation of in-situ based ECVs/CDRs include issues relevant to upscaling/downscaling, lexicons, long term stability, metadata, and homogenization and reference procedures. How to incorporate these into the SMM validation categories needs further investigations.

6.3. Reanalysis based

When all ECV datasets produced by an assimilation-based reanalysis are the products of a single reanalysis system, the system validation is necessary, but not sufficient; it is also necessary to conduct specific validation of the individual ECV datasets. The validation of data assimilation system was discussed based on its components, including: the observation inputs, the forcing inputs, the model configuration and the data assimilation system itself. It represents the current practice the reanalysis producers are following. It would be beneficial to cover this aspect in the SMM validation categories during its further development, considering the increasingly important role of reanalysis in producing ECV/CDR datasets.