EUMETSAT

CORE-CLIMAX European ECV CDR capacity assessment report

Doc.No. : CC/EUM/REP/14/004
Issue : v1A Draft
Date : 23 December 2014
WBSDBS:

EUMETSAT
Eumetsat-Allee 1, D-64295 Darmstadt, Germany
Tel: +49 6151 807-7

Document Signature Table

	Name	*Function*	*Signature*	*Date*
Prepared by:	Viju John	Core-Climax Scientific Support		
Prepared by:	Andrea Keiser-Weiss	DWD		
Prepared by:	Else Swinnen	VITO		
Approved by:	Joerg Schulz	Climate Services Product Manager		

Distribution List

Distribution list	
Name	*No. of Copies*

Document Change Record

Issue / Revision	*Date*	*DCN. No*	*Changed Pages / Paragraphs*

Table of Contents

EXECUTIVE SUMMARY

1 INTRODUCTION

1.1 Purpose and Scope

The purpose of this report is to document the capacity assessment conducted on behalf of the core-climax consortium on the European essential climate variable (ECV) climate data record (CDR) development.

1.2 Reference Documents

RD-1	Core-Climax System Maturity Matrix Instruction Manual	CC/EUM/MAN/13/002
RD-2	Tote, C., Coelst, F., Swinnen, E., Dong, Q (2014): CORE-CLIMAX Climate Data Record Assessment Web tool instruction manual	VITO
RD-a	5th Space Council - Council Resolution - Taking forward the European Space Policy, 26 September 2008	COUNCIL OF THE EUROPEAN UNION, 13569/08
RD-b	Wilson, J., M. Dowell and A. Belward (2010): European capacity for monitoring and assimilating space based climate observations – Status and prospects. JRC Scientific and Technical Report	*EUR 24273 EN*, 46 pp., DOI: 10.2788/70393
RD-c	Dowell, M., P. Lecomte, R. Husband, J. Schulz, T. Mohr, Y. Tahara, R. Eckman, E. Lindstrom, C. Wooldridge, S. Hilding, J. J. Bates, B. Ryan, J. Lafeuille, and S. Bojinski (2013): Strategy towards	

and architecture for climate monitoring from space. 39 pp., [available from: www.ceos.org, www.wmo.int/sat, www.cgms-info.org].

RD-d Bates, J. J. and J. L. Privette, (2012), A maturity model for assessing the completeness of climate data records, Eos Trans. AGU, 93(44), 441.

2 EUROPEAN AND INTERNATIONAL CONTEXT

Developing ECV climate data records poses many challenges because of the varied use of climate data, the complexities of data record generation, and the difficulties in sustaining the activities over extended periods of time. Therefore it is essential to assess the capability of the existing climate data record development activities to ensure the prolonged generation of high quality ECV climate data records so that they can help to produce the underpinning science that supports decisions on mitigation and adaptation for a changing Earth climate.

The European Union Council Resolution *"Taking forward the European Space Policy"*, as adopted by the Competitiveness Council meeting on 26 September 2008 was identifying Space and Climate Change as new priority area for the European Space Policy addressing the issue of climate change. The Council resolution [RD-a] invited the Commission *"... to conduct a study to assess the needs for full access to standardised data and for increased computing power, and the means to fulfil them, taking into account existing capacities and networking in Europe*."

Implementing this action the European Joint Research Centre conducted a workshop 2009 that did an ad hoc analysis of the European capacity on the means to provide these data and how Copernicus Services can effectively contribute to providing these data. The report by Wilson et al. [RD-b] is summarising the results of this workshop that identified 44 GCOS ECVs as the minimum set of standardised climate data that EC should be considering. This workshop did also a first attempt to analyse the capacity according to maturity, differentiating between sustained operational capacity and non-operational funded repetitive capacity and additional infrastructure needs in order to fill gaps identified.

Despite of these initial efforts a letter from GCOS and WCRP in May 2010 addressed the issue of ensuring transparency, traceability and good scientific judgement in the generation of climate data records that underpin climate research and climate change monitoring. This letter in particular stated:

"*However, there is currently no systematic international approach to ensure transparency, traceability and sound scientific judgement in the generation of climate data records across all fields of climate science and related Earth observations, and there are no dedicated sustained resources in place to support such an objective. For example, there are currently eight sea-ice concentration products produced by different organizations globally that differ significantly in providing an estimate of sea-ice extent and concentrations, mostly due to differences in methodology and not the variability or dynamics of underlying phenomenon. It is very confusing and frustrating for the non experts as to which one of these products they can use in their research and analysis, and the necessary documents to describe their attributes in a comparative manner akin to the global model inter-comparisons do not exist.*"

This letter clearly asked for a structured approach to generate, evaluate and maintain climate data records and also communicate information to the scientific and climate service communities in an understandable way providing guidance to the user.

A global group of space agencies has answered this request by developing a high level strategy for an architecture for climate monitoring from space [RD-c] that considers the whole value adding chain from making measurements to the development of policy and decision making. The report by Dowell et al. (2013) [RD-c] details two usage scenarios for such architecture:

- The promotion of a common understanding of the implementation implications of meeting the various climate monitoring requirements, and
- To support an assessment of the degree to which the current and planned systems that provide measurements from which climate data records are generated meet the requirements, and the generation of an action plan to address any identified shortfalls/gaps.

Essential for the second usage scenario is to assess what exists, what the degree of completeness and sustainability of the existing is, what quality the existing has and what is planned/committed for the future. The group of authors of the Dowell et al. [RD-c] report and the CEOS-CGMS Working Group Climate together with WMO established the so called GCOS ECV inventory (ecv-inventory.com) for climate data records derived from satellite measurements. Currently, the inventory consists of approximately 220 entries provided by space agencies around the world and provides a first basis for an analysis of the existing data records. Because the first call to populate the inventory was only directed to space agencies the current inventory holding is not complete and further work is needed to cover all relevant data records. In addition an analysis of the 'fit for purpose' of the data records needs to be done.

To support the international activities described above and the establishment of the Copernicus Climate Change Services one major objective of the CORE-CLIMAX project is to systematically assess the capacity of ongoing European activities in the area of generation and provision of climate data records. With respect to a Copernicus Climate Change Service also the role of in situ data and model-based reanalysis needs to be considered. As the in-situ data sets are generally under national responsibility the current assessment does not cover all relevant in situ data record generated by European providers. However, their inclusion in the assessment is demonstrating the value of the tools developed in CORE-CLIMAX and encourages more national managed data record to be included in a future assessment.

For an assessment of the European capacity in the most objective way possible we need tools that provide a basis for information preservation, expectations, and a metric for progress to completeness. The maturity matrix approach proposed by Bates and Privette [RD-d] offers a systematic mean to assess if the data record generation follows best practises in the areas science, information preservation and usage of the data. Some example uses of the matrix maturity are the assessments of data records developed in the NOAA Climate Data Record program and in the 2nd phase of SCOPE-CM to measure progress in the projects. For both these cases, maturity assessments were first done as self assessments. External assessments could be done in a form of audit.

The CORE-CLIMAX project's proposition is based on [RD-d], but extending the model to more general so that it can be applied not only for satellite data sets, but for all climate data records (*in situ*, combined satellite and *in situ*, reanalyses). The project discussed its adapted approach with many leading initiatives in Europe such as the EUMETSAT network of Satellite Application Facilities (SAF) and the ESA Climate Change Initiative but also internationally with WMO, the CEOS WG Climate, NOAA and USGS.
Basically, three different aspects of our capacity to generate data records need to be considered:

- Scientific, engineering and information preservation practises;
- Usage of products including feedback and update mechanisms ;
- Quality of products with respect to applications.

Assessing if data record generation follows best practises provides an internal view on strengths and weaknesses of the processes to generate, preserve and improve climate data records for agencies and each individual data record provider. It also provides a general information to the community concerning the status of individual data records as well as collective information on the state of all existing records, highlighting areas for development and improvement. The assessment of quality of products is facilitating an external view on data records trying to answer the most important user question: Is the quality good enough for my application?
The CORE-CLIMAX project defined three major elements for its capacity assessment:

- Data record descriptions that contain technical specifications and also information on quality, e.g., links to further documentation and/or inventories such as the CGMS-CEOS-WMO inventory (see http://ecv-inventory.com/ecv-inventory);
- A System Maturity Matrix (SMM) that evaluates if the production of a data record follows best practices for science, engineering, information preservation and facilitation of usage, and;
- An Application Performance Metric (APM) that attempts to evaluate the performance of an ECV CDR with respect to a specific application. To be able to apply the APM, user requirements for each application are needed to be compared to the actual technical specifications and validation results.

The three elements of the capacity assessment are designed to be independent of each other and represent means to support an assessment but do not provide the assessment results per se. The SMM is designed to principally be used without considering specific applications. With this the SMM does not depend on user requirements for specific applications and their change over time. In contrast the APM should facilitate a comparison of the real technical features of a data record and results of validation and other data quality assessment activities to user requirements for an application. It basically provides summary information on how close a specific data record is at fulfilling the requirements of a specific application. A prototype APM was developed during the capacity assessment workshop because the need of giving advice to data users what data record can be used for what application is needed. This need is manifested for instance in the huge amount of information provided on validation of data records that is unlikely to be processed by institutions that want to use the data records. The APM is intended to support institutions in making choices among different existing data records without the need to assess the full documentation of all potential data records. However, it shall be noted that the APM is a new concept that was used for the first time in the CORE-CLIMAX capacity assessment workshop. The development of an applicable tool

was beyond the scope of the CORE-CLIMAX project. It is expected that the APM will be further adapted in the future to become fully functional.

3 ASSESSMENT TOOLS AND PROCEDURE

3.1 Data Set Description Template

A data record description template which is given below has been filled for each individual data record that entered in the CORE-CLIMAX assessment. The template is structured very similar to the template used for data sets entering the Climate Model Inter-comparison Project (CMIP) exercise. Only the part on the applications has been extended as the usage of most climate data records goes beyond the climate model comparison. Keeping these templates very similar was done purposefully to support the usage of the assessed data records in the CMIP-6 exercise with preparations being started during 2014. The Data Set Description Template contains advice on how to fill the individual sections. The overall aim is that these descriptions do not extend to more than five pages.

CORE-CLIMAX DATA SET DESCRIPTION

(General Note:** This data set description shall not become longer than 5 pages per data set described. Please stay to the most important facts and use tables and bullet lists to provide information where appropriate.**)

(Type Data Set Name and if available digital identifier here):

- ***INTENT OF THE DOCUMENT***

(Provide information on what data set is described and for what application(s) it was created. Keep in mind that the information is targeted at users of any level who wish to use the dataset for climate applications. Users may not be expected to be experts for in situ, remote sensing or reanalysis techniques.)

- ***POINT OF CONTACT***

(Please provide a point of contact: Organisation and Contact details (at least a contact name, organisation and e-mail address)).

- ***DATA FIELD DESCRIPTION***

(Provide a link to an existing technical product specification or provide the information in a form of a table in this document. The specification shall at least include variable names and units (eventually including uncertainty estimates that come with the product), length of record, spatial coverage, spatial and temporal sampling.)

- ***DATA ORIGIN***

(Provide a basic description of the methodology used to derive the product including the input data used and the source (provenance) of the data. Also provide a description of data processing methods such as (inter-satellite) calibration, algorithms employed,

homogenization applied, mapping and averaging, etc. If the product makes heavily use of NWP and/or climate model data, e.g., as background fields this should be described as well.
In case of reanalysis data records please indicate what reanalysis system (coupled or single) has been used and name and version of the model(s).)

- ***VALIDATION AND UNCERTAINTY ESTIMATE***

(Provide a summary of validation activities performed for the product and provide a summary of systematic and random uncertainty of the product and how these vary with space, time and state (tabulated form appreciated). In particular information on temporal stability of the data which is an indication of whether the data can be used for longer term variability and trend analysis is appreciated.)

- ***CONSIDERATIONS FOR CLIMATE APPLICATIONS***

(Provide information on the applicability of the product for the planned application (stated in section 1) including limitations. In particular observational products applicable for model evaluation should state the different character when compared to model data. For instance for satellite-derived products it is important to describe limitations such as validity in specific areas (e.g., ocean or land only), unresolved diurnal cycles or diurnal cycle aliasing due to orbit drifts for polar orbiting satellites, sampling issues such as in the presence of clouds, sensitivity of the instrument, etc and their respective impacts on the application. For in situ measurements or gridded data sets derived from station data limitations due to the representativeness of the data, etc. and their effect for an application shall be provided.)

- ***INSTRUMENT OVERVIEW***

(Provide information on the type of instruments (in situ/remote sensing) used to measure the variable provided including the measurement principle (e.g., infrared emission measured with a spectrometer) and give a description of the instrument science objective, capability, measurement principle, satellite and orbit characteristics or observation location and practice for in situ. Provide the strengths and weaknesses of the instrument measurement. If an instrument simulator is available, provide a short description and references later for details.
In the case of a re-analysis data set only indicate what instrument data relevant to the parameters considered have been assimilated. This can simply be a link to the information.)

- ***REFERENCES***

(Provide a complete list of references used in this document and may provide additional reading references on measurement principles, retrievals, modelling, validation, uncertainty characterisation, product, and applications.)

- ***REVISION HISTORY***

(Indicate the version number of this document, the date of writing and who has edited the document.)

3.2 System Maturity Matrix (SMM)

3.2.1 SMM description

The SMM is a tool to assess the system maturity of a CDR. SMM basically assesses whether CDR generation procedures have been compliant with best practices developed and accumulated by the scientific and engineering communities. The concept behind the CORE-CLIMAX system maturity matrix can be best illustrated as shown in Figure 1.

Creation of a climate data record is anchored on a number of assumptions and approximations, and thus is associated with significantly large uncertainties. This is mainly because the observing systems were designed to measure weather, but not for monitoring climate. Unless these assumptions and approximations are well understood and associated uncertainties are well characterized it is quite possible to misinterpret results of scientific analyses using these data records. Therefore uncertainty characterisation is a key area where CDRs need to achieve high levels of maturity.

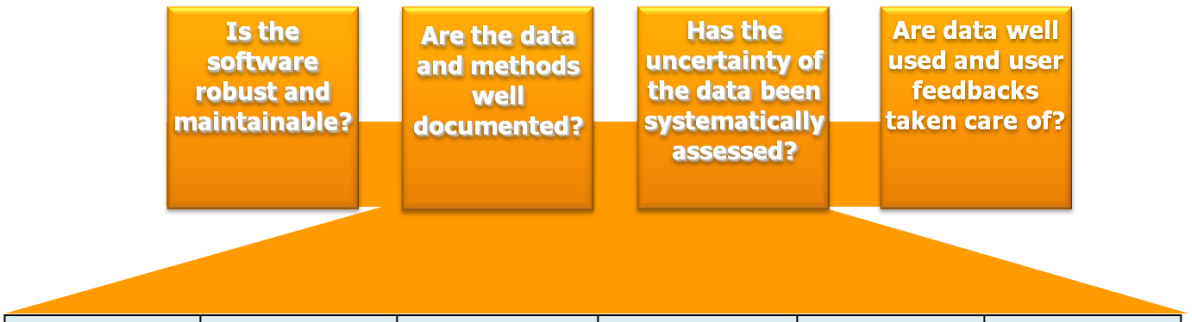
Stable and easily maintainable software is one of the essential components of successful CDRs. It should be easy to diagnose deficiencies, to make changes to the software, and to test the software after modification. Non-maintainable software can result in unexpected increase in the production cost of data sets. The metadata, especially describing the input raw data are essential because development of a CDR is often an evolutionary process and repeated reprocessing of the input dataset is necessary. This also demands the archival of the raw data for reprocessing. CDRs shall be archived in a way that allows easy access to the users with varying requirements and skills. Therefore it demands less complicated file structures and provisions for read and analyses (e.g., sub-setting, plotting) software. Availability of comprehensive descriptions of technical and scientific aspects of the production chain is another essential characteristic of a mature CDR.

Above all the most important maturity characteristic of a successful CDR is the acceptance and usage by the user community and whether there are mechanisms to receive and incorporate feedbacks from the user community.

There are 6 major categories where assessments are made:

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

For each of these categories the assessment will assign a score from 1 to 6 that reflects the maturity of the CDR with respect to a specific category. An overall score, e.g., an arithmetic mean for a CDR might be computed out of the six categories but it is not considered to be very useful.



Software readiness	Metadata	User documentation	Uncertainty Characterisation	Public Access, Feedback and Update	Usage
Are the codes compliant with standards, stable, portable and reproducible?	Do the metadata meet international standards, and allow provenance tracking?	Are the formal documents and peer-reviewed papers up-to-date and public?	Are the uncertainties assessed systematically in a standard manner?	Are the data, source code, and documents publicly available and regularly updated?	Are the data widely used in the scientific, and decision and policy making communities?

Figure 1: Illustration of CORE-CLIMAX system maturity matrix (SMM).

The maturity is also considered in three broad categories that give information on the grade of sustainment of the CDR generation process. The nomenclature for these broad categories has been imported from NOAA and follows [RD.3]:

- Maturity scores 1 and 2 establish Research Capability (RC): All aspects of the CDR are still under development and with the PI most likely in projects.
- Maturity scores 3 and 4 establish an Initial Operations Capability (IOC): At this stage the CDR and associated material are available to the user community. The CDR has reached a status where its usefulness is completely demonstrated and decisions need to be made to sustain its maintenance and further development. At this stage so called transitions of CDR generation capabilities from research units to more operational oriented units are happening. Good examples for this are the import of the HOAPS data record (www.hoaps.org) from the Max-Planck Institute in Hamburg into the EUMETSAT Climate Monitoring Satellite Application Facility or the transition of the well known International Satellite Cloud Climatology Project (ISCCP) from NASA into NOAA-NCDC.
- Maturity scores 5 and 6: Full Operations Capability (FOC): At this stage the production of the CDR has been transitioned into operational environments, e.g., the whole processing process is under configuration management, fully automated and performance is monitored. The production chain meets the goal of acquiring capabilities to provide uninterrupted and indefinite data provision for climate monitoring. The data provider, e.g., a space agency takes complete responsibility for the maintenance and also further development of the CDR. The specific development activities still are performed by scientists within or external to the responsible agency or both. A current example for a full operations capability is the EUMETSAT CM-SAF.

The major categories of the SMM are subdivided into several minor categories and assessment scores are assigned based on scores in these minor categories. After a long deliberation at the workshop we have decided to take the range of scores (that is the

minimum and the maximum) of the minor categories to represent a major category. The motivation for showing the minimum score is given by the fact that this score is informing about completeness of a major category. It directly points to an area for improvement. The maximum score will then indicate whether some minor categories have a higher score. It should be noted that the numbers need anyway an interpretation per assessed data record because the circumstances under which the data records were created hugely differ for satellite, in situ data records and reanalysis.

The minor categories sometimes include categories that cannot easily be assessed by an external assessor without asking the provider of the data which could be done in a formal audit type assessment but was not foreseen for the CORE-CLIMAX capacity assessment.

The SMM is provided as a multi-level Excel file where the scores shall be provided in the pages associated with the minor categories. These scores are then automatically be used to mark the range of scores for the major category highlighting the minimum score to directly point to an imminent issue. If a minor category is not filled a maturity of 1 will be set. There is one exception which is in the category Usage. In this category the usage of a data record is considered for applications in research and decision making. Which columns are taken into account depends on the intention of the data record. For instance, if the description is only pointing to use in research only that category shall be used to compute the overall maturity. It is planned to replace the Excel file with a web based tool and its availability will be communicated in due course of the project.
It is very important to use a unique CDR name and identification number (version) when the SMM is filled to assess a CDR. This shall match the name and identification information on the data set description form. Also a provision of the assessment date to follow the evolution in maturity of a particular CDR is very important.

3.2.2 SMM Web tool

A web tool has been designed so that CDR reviewers can perform the assessment of a CDR and thereby provide input to a database, from which information (statistics etc.) can be derived.

The purpose of this web tool instruction manual is to give a brief description specifically on how the web tool should be used. For a detailed description on the web tool please refer to RD.2.

Once all scores of the SMMs are entered one is able to generate overview of the maturity of the dataset. The overview shows the variable name, the assessment version, the earth-system which the dataset represents, the original project from where the variable is derived, type of assessment, and the minimum and the maximum scores as well as the time of the assessment.

The overview can be exported in a format of table where the red grids indicate the score between 1 and 2, the yellow grids represent the score between 3 and 4 while the green cells are those score between 5 and 6.

FAPAR , Self-assessment, Satellite

version= 1 ecv name= FAPAR earth system domain name= Land projectname= GIO modificationdate= 2014/01/17

Software Readiness	Metadata	User documentation	Uncertainty Characterisation	Public Access Feedback Update	Usage
Coding standards	Standards	Formal description of scientific methodology	Standards	Public Access/Archive	Research
Software Documentation	Collection level	Formal Validation Report	Validation	Version	Decision Support System
Numerical Reproducibility and Portability	File level	Formal Product User Guide	Uncertainty quantification	User Feedback Mechanism	
Security	Formal description of operations concept	Automated quality monitoring	Updates to Record		

Figure 2: An example of SMM web tool generated overview of a dataset

4 ASSESSMENT WORKSHOP

The CORE-CLIMAX Essential Climate Variable (ECV) Capacity Assessment Workshop was held at EUMETSAT Head Quarters in Darmstadt on 21 – 23 January 2014.

The workshop gathered about 40 participants representing all relevant European climate data record producers including EUMETSAT SAF network, ESA CCI projects, EC projects and EUMETSAT member state weather services as well as stakeholders (EC, CGMS, CEOS, WMO). The workshop covered data records for Land, Ocean, and Atmosphere constructed from satellite and in situ data as well as reanalysis. The workshop participants were asked to do a self assessment of their data records using the SMM prior to the workshop and there were 30 data sets (23 are based on satellite measurements, 6 are based on *in situ* measurements, and one is reanalysis based) with SMM results for analyses during the workshop.

First day of the workshop was dedicated to provide a big picture on the current status of international and European activities on the generation of ECV CDRs and the need of their sustained generation for effective climate services. The history, concepts, and initial feedbacks on the developed tools were also discussed to develop a common understanding on the assessment tools.

On the second day, the participants were divided into three subgroups (Land, Ocean, and Atmosphere themes) to discuss the SMM self assessments, to make more common understanding of the SMM, to reconcile the differences between different CDR communities, to come up with suggestions for improvements to the SMM, and to deliberate the idea of the APM.

The third day of the workshop presented and analysed the discussions from the three breakout groups and made recommendations and the way forward. The workshop participants endorsed the SMM as a useful tool for assessing completeness and identifying weaknesses in the development systems which produce CDRs from satellites and *in situ* measurements and reanalyses. The presentations and breakout group discussions provided a common

understanding of most of the elements in the SMM among different communities and programs and based on these understandings further improvements have been proposed for the SMM and its instruction manual. However, the workshop suggested that the process of introducing the concept of SMM in different communities needs to be open and inclusive. As the SSM is based on "best practises", reviews and changes to the SMM are expected on longer time scales (~10 years), but it is expected to have a review and eventual change prior to an implementation to a major service such as the Copernicus Climate Change Service. It was recommended to use the SMM in the review process of existing CDR programs such as EUMETSAT-SAF network and ESA-CCI. It was also recommended that SMM can be linked with the publication of CDRs in peer reviewed journals similar to the enforced use of unique digital object identifiers. The presentation of SMMs will follow the proposed colour coding (range of scores in each sub-category) and may include an interpretation of the scores.

5 RESULTS

The main workshop/assessment results can be summarised as follows:

- Operational fundamental climate data records generally gets high scores for Software Readiness, Metadata, and Usage, but low scores for User Documentation and Uncertainty Characterisation;

- Scientific data sets get high scores for User Documentation, Uncertainty Characterisation, and Usage. These datasets have mostly low scores for Software Readiness;

- In fact, getting consistently high scores in all categories is also a matter of time, i.e., maturing of the system to produce datasets takes time.

- *In situ* datasets are often updated continually and therefore the concept versioning is different to satellite datasets. It is encouraged to use versioning for better reproducibility, transparency, and for easier understanding and use of the data.

- An example to illustrate the problem by Nick Rayner, Met Office: "*As an example, let's take EN4, our sub-surface ocean data set. We're about to release EN.4.1.2.0, where the version number is incremented either when there is a software change and/or when new individual batches of observations are included (in the past, not for a monthly extension of the data set). So, extending that to land station data (as we do for CRUTEM4, currently on CRUTEM.4.2.0.0) this means bringing in either (i) a new set of stations or (ii) a number of years that have recently been digitised from stations that we already had; this changes the version number. This is important because people can then trace their results or their gridded data set back to a particular version of the in situ observations and so know why their results or their data set might differ from someone else's who used a different version.*"

- The maturity of *in situ* data largely depends on the selection of a set of stations or time period. If old and thus especially valuable data records would be excluded, this would immediately raise the maturity noticeably. In general, a more comprehensive

data set will have a lower maturity compared to a selected subset, which might lead to misinterpretation.

- There was a debate on the applicability of Software Readiness for in situ data. The cost of non-portable/non-documented software can be big. In "semi-operational" systems to generate data sets where people come and go, operating systems come and go, programming languages come and go, but the data set or the need to produce a CDR of a particular ECV remains for decades, and it is a lot more expensive when those systems are not portable and not well documented and so not easy to pass on to someone else to maintain/upgrade than when they are. It makes the difference between making incremental changes and having to start all over again.

- A point was made that there are no SMM assessment for source data. The SMM, in principle, is applicable to and applied to all kinds of climate data, including those from NMSs, so in the future there could be SMM assessments for those source data too.

- One should be wary of changing the SMM definition too often, it is encouraged to have a review and eventual change in a time frame of ~10 years.

- It is encouraged to use SMM as a part of peer-review when papers are published on data sets (in specialised journals).

6 ASSESSMENT RESULTS PER DATASETS

We conducted the assessment by categorising the datasets based on measurements target: Atmosphere, Ocean, Cryosphere, and Land. We treat fundamental climate records and reanalysis separately.

6.1 Fundamental Climate Data Records

We present SMM results of fundamental climate data records here. Fundamental Climate Data Records – FCDR – are sensor data that have been improved and quality controlled over time, together with ancillary data used to calibrate them. We treat them separately here because they can be used to create Thematic Climate Data Records – TCDR – which can pertain to different earth systems.

6.1.1 STRATOSPHERIC SOUNDING UNIT (SSU) FCDR

Name	SSU Level 1b radiances
Origin	NCDC/CLASS; Cheng-Zhi Zou cheng-zhi.zou@noaa.gov
Spatial Characteristics	Global
Temporal Characteristics	Dec 1978 – Jan 2006; Instantaneous

Category	Software Readiness	Metadata	User Documentation	Uncertainty Characterisation	Public Access, Feedback, Update	Usage
Range	1	2-3	2-3	2-3	3-4	3-4

The SSU level 1b radiances are stored at the NOAA-CLASS archive: http://www.class.ncdc.noaa.gov/saa/products/welcome/. They have been processed from the level 0 counts to calibrated radiances. The SSU observed radiances in the 15 micron carbon dioxide band, with three channels peaking at about 29 km, 35 km and 45 km in the stratosphere. The designation of these channels here is 1 to 3 respectively, although they are often referred to as channels 25 to 27 of the TOVS in the literature. The radiances originated from deep atmospheric layers. For nadir views the half-width varies from 16 km deep (channel 1) to about 22 km (channel 3). The SSU only flew on TIROS-N and NOAA's 6, 7, 8, 9, 11 and 14. In 1998 the AMSU-A instrument was launched which replaced the SSU in operations but measurements continued to be made by the SSUs until 2006. The data are stored as an orbit by orbit format with one record per scan line.

This is an important dataset for monitoring changes in stratospheric temperatures. However the current state of SSU level 1b radiance datasets still require work for climate monitoring (e.g., Thompson et al., 2012[1]).

The software for this data set is in a preliminary stage which means a good effort is necessary to make this data set operational. All sub categories have score 1.

Metadata are sufficient to understand and use the data at both collection and file levels (score 2 each), however no metadata standards are considered.

Formal documentations on scientific methods, validation are available in the peer reviewed literature and 2 recent papers describing the data processing have been accepted for peer reviewed publication. There is no description of operations concept, but a limited product user guide is available. Thus this category is primarily under research grade.

Significant efforts have been made recently to quantify uncertainties for this dataset and peer-viewed papers have been submitted or published.

Data and documentation are archived and available from NCDC, however there is no regular or standard mechanism established for collecting user feedbacks and updating the dataset.

The product has been moderately used for research and decision making.

[1] Thompson et al., 2012: The mystery of recent stratospheric temperature trends, *Nature* 491, 692–697, doi:10.1038/nature11579.

6.1.2 SSM/I FCDR Edition 1.0

Name	SSM/I FCDR
Origin	CMSAF; contact.cmsaf@dwd.de
Spatial Characteristics	Pixel resolutions varying with channels.
Temporal Characteristics	Jul 1987 – Dec 2008; Instantaneous

Category	Software Readiness	Metadata	User Documentation	Uncertainty Characterisation	Public Access, Feedback, Update	Usage
Range	1 – 4	5 – 6	2 – 5	3 - 5	4 - 5	1 - 2

The CM SAF Fundamental Climate Data Record of SSM/I Brightness Temperatures provides homogenised and inter-calibrated brightness temperatures from the six SSM/I radiometers aboard F08, F10, F11, F13, F14, and F15.

Software used to generate the data set is in the initial operations capability except that security aspects of the software are not tested yet.

Metadata has the highest score for all categories. This implies that the dataset meets international standards for metadata and the compliance has been systematically tested and complete collection and file level metadata are available for the entire data period.

Formal scientific descriptions of the methods used to generate the data are available from the data provider. Data provider keeps updated validation report and user manual and a peer-reviewed paper on validation is published and comprehensive description of operations concept is available. More work on publishing the scientific methods will move the User Documentation category to full operational capability.

The dataset is in initial operations category for Uncertainty Characterisation category. The data set has not been participated in an international data assessment, but such assessments are generally not available for FCDRs.

Data and documentation are archived, under version control, and available from the data provider. However, the source codes which are used to generate the data are neither archived nor available to users. Regular mechanisms for collecting user feedbacks are established and the data are being regularly updated whenever new input data or new feedbacks are available.

These data are not very well used yet.

6.1.3 Baseline Surface Radiation Network (BSRN)

Name	Archive of the Baseline Surface Radiation Network (BSRN)
Origin	World Radiation Monitoring Center (WRMC); Gert.Koenig-Langlo@awi.de
Spatial Characteristics	Global stations, the number is varying over time, e.g., 9 stations 1992 and 58 stations 2014.
Temporal Characteristics	selected research stations, which provide typically 1-minute averaged short- and long-wave surface radiation fluxes of the best possible quality currently available, and aggregated products (daily, monthly).

Software Readiness	Metadata	User Documentation	Uncertainty Characterisation	Public Access, Feedback, Update	Usage
Not rated	6	Not rated	2-4 some not rated	4-6	5-6

Main objectives of the Baseline Surface Radiation Network (BSRN; http://bsrn.awi.de) dataset are: to monitor the shortwave and longwave surface radiative fluxes, to provide data for the calibration and validation of satellite-based estimates of the surface radiative fluxes and to produce high quality observational data for comparison with models and to derive long-term regionally representative climatologies.

As this dataset is generally a collection of raw measurements, the system for generating it does not require a robust software system, for example, in case of a satellite data production system. However, a check of the website reveals that at the Data Warehouse, there must be some software to handle the data, to check, compress and format change, to produce station lists and maps, and some software for quality checks are provided to the users. Therefore, Software Readiness maturity of these components could have been rated.

The highest score for metadata represents the completeness of this category.

User documentations is not rated, however, users are guided in a convenient manner to their desired product via the web-pages. There are papers and other material explaining the methods of observations. Thus, this category could also have been rated.

Papers and reports on uncertainty information are linked, Validation using external reference data done for limited locations and times, a quality toolbox is provided, and limited information on uncertainty arising from systematic and random effects in the measurement.

Data and documents are archived and version controlled by the data provider. Regular mechanisms for user feedbacks are established and the product is regularly updated and interim data records are available.

The product has become a reference dataset for multiple applications and has been widely used for policy making.

6.2 Atmosphere

In this session we present assessment results of the atmospheric data sets.

6.2.1 ESA GHG-CCI datasets

Dataset attributes

Name	GHG-CCI; XCO_2 and XCH_4
Origin	ESA GHG-CCI; buchwitz@uni-bremen.de
Spatial Characteristics	Global, 30 x 60 km^2, global (land only for CO_2)
Temporal Characteristics	2002 – 2012 (SCIAMACHY); 2009 – present (GOSAT); Instantaneous

System Maturity Matrix Scores

Category	Software Readiness	Metadata	User Documentation	Uncertainty Characterisation	Public Access, Feedback, Update	Usage
Range	1-2	3-4	2-3	2-4	3-5	3

Climate prediction requires a good knowledge on the Greenhouse Gas (GHG) sources and sinks, in particular CO_2 and CH_4 and other GHG. The GHG-CCI (http://www.esa-ghg-cci.org/) data products contain information on regional CO_2 and CH_4 surface fluxes (emissions and uptake) and therefore can be used to improve our knowledge on GHG surface fluxes.

These products are generated with GHG-CCI "ECV Core Algorithm" (ECAs). ECAs are algorithms to retrieve dry-air column-averaged mole fractions of carbon dioxide (CO_2) and methane (CH_4), denoted XCO_2 (in ppm) and XCH_4 (in ppb) from currently two satellite instruments: SCIAMACHY onboard ENVISAT (2002-2012) and TANSO-FTS onboard GOSAT (2009-ongoing). The GHG-CCI products are Level 2 products, i.e., detailed information such as time and location is provided for each single satellite observation (ground pixel). Requirements are formulated in the corresponding User Requirements Document (URD, http://www.esa-ghg-cci.org/?q=webfm_send/173).

The software used to generate the dataset in under research grade, i.e., coding standards are not fully systematically applied, limited documentation of the code, and the software is not fully checked for portability, numerical reproducibility, and security.

Metadata are in initial operational capability, i.e., standards are defined but not systematically applied, collection level metadata are sufficient to use and understand the data without external assistance, and limited location level metadata are available. Recently the data format and documentation has been significantly improved (see, http://www.esa-ghg-cci.org/ -> Documents and -> CRDP (Data), in particular the PSD (http://www.esa-ghg-cci.org/index.php?q=webfm_send/160)).

User Documentation is in initial operational capability category. Comprehensive descriptions of scientific methods and validation are available on the GHG-CCI website (http://www.esa-ghg-cci.org/ -> Documents and -> CRDP (Data)) and journal papers are published on the methods and validation (http://www.esa-ghg-cci.org/ -> Publications). Product user guides are available (http://www.esa-ghg-cci.org/ -> CRDP (Data)) and the operations concept (http://www.esa-ghg-cci.org/?q=webfm_send/193).

The dataset falls into initial operational capability for Uncertainty Characterisation category (latest assessment: PVIR: http://www.esa-ghg-cci.org/index.php?q=webfm_send/160, AIECAR: http://www.iup.uni-bremen.de/~buch/ghgcci_public/AIECARv1_GHG-CCI_Final.pdf). The only sub-category under research grade is Automated Quality Monitoring which is only partially in place yet (see http://www.esa-ghg-cci.org/ -> CRDP (Data)).

The data, documentation, and source code are archived by the data provider, which are version controlled by the PI. User feedback is collected and regular updates to dataset are made by the PI.

The data are extensively used for research (http://www.esa-ghg-cci.org/ -> Publications) but not yet for decision support system.

6.2.2 GPCC Full Data Reanalysis Version 6

Name	GPCC Full Data reanalysis Version 6
Origin	Deutscher Wetterdienst; gpcc@dwd.de
Spatial Characteristics	0.5°, 1.0° and 2.5° regular lat-lon grid, global, over land, without Antarctica
Temporal Characteristics	1901-2010; monthly

Software Readiness	Metadata	User Documentation	Uncertainty Characterisation	Public Access, Feedback, Update	Usage
2 – 4	6	2 – 6	2 – 6	3 – 5	6

The GPCC monthly gridded precipitation fields are calculated from internationally collected and quality controlled rain-gauge data (GTS-based and historical data). The objectives of the GPCC Full Data Reanalysis (GPCC_FD) are for verification of models, for analysis of historic global precipitation, for research concerning the global water cycle, e.g., trend and tele-connection.
Full details are at ftp://ftp.dwd.de/pub/data/gpcc/html/fulldata_v6_doi_download.html.

Software readiness maturity ranges between 2 and 4, which are appropriate for the production of *in situ* datasets. Coding standards are identified, moderate documentation of the software is available, and PI affirms portability, numerical reproducibility, and no security issues.

Metadata maturity is self-assessed to highest scores. However, it is very ambitious for such a large collection of data including historical data.

Maturity of User Documentation is very high, except for the formal validation report, where maturity 2 is reached with limited validation available from the PI. Potential documentation on comprehensive validation and inter-comparisons would increase the maturity of category.

The maturity scores with respect to Uncertainty Characterization range from 2 to 6. SI traceability is established, comprehensive validation and inter-comparisons have been done, and automated quality monitoring is fully implemented. However, only limited information is available on uncertainties arising from systematic and random effects in the measurements.

Public Access, Feedback, Update is of adequate maturity (range 3-5), with data and documentation are version controlled and archived with public access. Feedbacks are collected from scientific community and they are used for regular update of the product.

The dataset is extensively used for research and policy making (for instance, it is used in the IPCC AR5 WG1 Summary for Policy Makers).

6.2.3 NKDZ

Name	NKDZ station data, historical version v002
Origin	Deutscher Wetterdienst; datenservice@dwd.de
Spatial Characteristics	Stations covering Germany
Temporal Characteristics	1/1/1781 to 31/12/2013; hourly, daily, monthly, annual

Software Readiness	Metadata	User Documentation	Uncertainty Characterisation	Public Access, Feedback, Update	Usage
1-4	3-6	1-3	4-6	3-6	6

This is a national data set, produced by the Deutscher Wetterdienst (DWD). In July 2014, DWD had released its historical measurements of climatological parameters from the DWD climatological and meteorological stations together with station-specific metadata for free public access at ftp://ftp-cdc.dwd.de/pub/CDC/observations_germany/

The special value of this data set is in its comprehensive collection (back in time to 1781, and in the level of detail the station specific metadata (such as, e.g., known station history, instrument change or change in averaging formula).

Software Readiness ranges from research to initial operational capability. Coding standards are not used, moderate software documentation is available, and PI affirms no security issues. Third party affirms portability and numerical reproducibility.

Metadata standards are not systematically applied, but the data set is self assessed to have complete collection and file level metadata. However, this is ambitious in the light of lost metadata of historical measurements.

User Documentation is generally under research grade, that is, only limited documentation is available grade, which points to need for improvements in documentation.

In the Uncertainty Characterisation, automated quality monitoring has the highest score, but there is much room for improvement for standards, validation, and uncertainty quantification.

Public Access, Feedback, Update category is under initial to full operational capability with ICDRs being produced.

The dataset is widely used for research and policy making.

6.2.4 ESA-CCI Aerosol datasets

Name	ESA Aerosol_cci datasets
Origin	ESA Aerosol_cci; thomas.holzer-popp@dlr.de
Spatial Characteristics	Global, different resolution (0.1 to 10 degrees)
Temporal Characteristics	~weekly-monthly sampling (between 1995 and 2012)

Category	Software Readiness	Metadata	User Documentation	Uncertainty Characterisation	Public Access, Feedback, Update	Usage
Range	1-3	4-5	3-4	2-4	3-5	1-2

Aerosol CCI provides following datasets after three years of intensive algorithm development, sensitivity analysis, validation and inter-comparison activities (Holzer-Popp, et al., 2013) and a round robin exercise of eight different algorithms (de Leeuw et al, 2013):

- AOD and Ångström exponent (AE)
- Stratospheric extinction profiles and aerosol optical depth
- Absorbing aerosol index (AAI)

AOD/AE datasets are provided for two main purposes: climate aerosol model evaluation, data assimilation into global aerosol re-analysis / forecasting model systems. The stratospheric dataset is provided as correction to the total column AOD retrievals (mostly relevant in case of major volcanic eruptions) and for stratospheric climate model evaluation. The AAI dataset (so far as only information for aerosol absorption, though qualitative) has been prepared for comparison to model datasets by developing a model AAI simulator and analyzing major sensitivities.

The software used to generate the datasets is partially compliant with coding standards and contains moderate documentation. However, the code is not evaluated for portability and numerical reproducibility, but the PI confirms of no security problems.

The metadata associated with the datasets are compliant with international standards and there is complete file level metadata and the collection level metadata are sufficient to understand and use the data.

The dataset falls into category of initial operational capability for User Documentation by maintaining comprehensive description and peer reviewed paper on scientific methods. Validation report is available from the PI and a paper on validation is submitted, user guide and limited operations concept are available from the data provider.

Uncertainty characterisation category is also in the initial operational capability. The data are inter-compared with other corresponding CDRs and quantitative estimates of uncertainty provided at location level. However, automatic quality monitoring is not in place.

Data, source code, and documentation are archived by the data provider with PI's versioning. The datasets are regularly being updated by using feedbacks from scientific community.

The data usage is on initial level (e.g. data assimilation test by ECMWF/MACC, trend and model comparison by AEROCOM).

6.2.5 GNSS Radio Occultation

Name	GNSS-RO DATA
Origin	ROM SAF
Spatial Characteristics	Global, 0-40 km (5x5 degrees in horizontal, 200 m in vertical)
Temporal Characteristics	09/2006 to present

Category	Software Readiness	Metadata	User Documentation	Uncertainty Characterisation	Public Access, Feedback, Update	Usage
Range	3-6	4-6	5-6	4-6	5-6	2-3

ROM SAF monthly climate data products publically are available in the ROM SAF archives (http://www.romsaf.org). The data cover the troposphere and the lower and middle stratosphere, and are primarily intended for global climate monitoring and climate research applications, including model validation.

The software used to generate the dataset under initial or full operational capability. The code is partially compliant with coding standards, moderately documented, portability and numerical reproducibility checked, and passed data provider's security review.

The metadata associated with the datasets are compliant with international standards and there is complete file level metadata and the collection level metadata are sufficient to understand and use the data.

The User Documentation category is in full operational capability with comprehensive documentation on methods, validation, user guide, and operations concept.

Uncertainty characterization category is in initial to full operational capability. Procedures to achieve SI traceability are defined, datasets participated in multiple international data assessment, uncertainties are quantified at location level, and fully automated quality monitoring feeding back to metadata or documentation.

Full operational capability is also achieved for Public Access, Feedback, and Update category with fully version controlled data, source code, and documentation are archived by the data provider. Regular feedback mechanisms are established and interim data records are being produced.

The data are yet to be fully utilized as a CDR.

6.2.6 Free Tropospheric Humidity (FTH)

Name	FTH edition 1.0
Origin	CMSAF; contact.cmsaf@dwd.de
Spatial Characteristics	(45W – 45E; 45S – 45N), gridded res. - 0.625°×0.625°
Temporal Characteristics	July 1983 to December 2009, 3 hourly sampling

Category	Software Readiness	Metadata	User Documentation	Uncertainty Characterisation	Public Access, Feedback, Update	Usage
Range	1-3	5-6	3-5	3-4	4-5	2-3

FTH is an important climate variable due to its large sensitivity to outgoing longwave radiation (OLR) and thus a strong feedback factor in the Earth's climate system. Monitoring of FTH globally is therefore important to understand our changing climate. However, measuring humidity in the free troposphere is a challenge and there are only a few datasets available for FTH. One of these datasets is the FTH dataset produced by the EUMETSAT's CMSAF (http://wui.cmsaf.eu/safira/action/viewDoiDetails?acronym=FTH_METEOSAT_V001). FTH are derived from METOSAT 1st and 2nd generation geostationary satellites.

The dataset has highest score in Metadata sub-category, which meets Full Operational capability.

The scores are very high for User Documentation sub-category which is a reflection of rigorous documentation process in the SAF network. The scores are meeting Initial or Full Operational capability.

The Uncertainty Characterisation section is in the Initial Operational capability.

The Access, User Feedback and Update are also in Initial Operational capability.

The usage of this product is mainly for Research and it is under Initial Operational capability. The expectation is to have an impact on decision support system in the long run.

The dataset can potentially make some improvements in the Software Readiness category because the scores in this category are either in Research Grade or Initial Operational capability.

6.2.7 HOAPS release 3.2

Name	HOAPS release 3.2
Origin	CMSAF; contact.cmsaf@dwd.de
Spatial Characteristics	Only over the ice-free ocean surfaces; 0.5° x 0.5°
Temporal Characteristics	Monthly averages and 6-hourly composites

Category	Software Readiness	Metadata	User Documentation	Uncertainty Characterisation	Public Access, Feedback, Update	Usage
Range	1-5	5	4-5	2-4	4-5	2-4

The Hamburg Ocean Atmosphere Parameters and Fluxes from Satellite (HOAPS) data set is a satellite-based climatology of precipitation, evaporation, freshwater budget (evaporation minus precipitation), related turbulent heat fluxes and atmospheric state variables as well as liquid water path and total column water vapour over the global ice free oceans.

The scores in the Software Readiness vary from research grade to full operations capability. Coding standards are applied, but not compliant to the standards yet. The software is portable and results are numerically reproducible, but security aspects of the code are not tested yet.

Metadata associated with this dataset is compliant with international standards and complete at file and collection levels. This implies full operational capability.

User Documentation falls under initial to full operational capability. Journal paper is published on methods and comprehensive description of scientific methods is maintained by dataset developer. Report on inter-comparison with other CDRs available and the product user guide is regularly updated. Comprehensive description of the operations concept is also available.

The dataset generally falls under initial operational category for Uncertainty Characterisation. Standard uncertainty nomenclature is applied, comprehensive validation including comparisons against other CDRs are done, but there is only limited information available on uncertainties arising from systematic and random effects in the measurements. Automated quality monitoring is partially implemented.

The dataset generally falls under initial operational category for Public Access, feedback, and update. Data and documentation are under institution's version control, archived and, publicly available. The dataset is regularly updated taking into account established user feedback mechanism.

The data have been used in research which is evidenced by citations.

6.2.8 ECA&D

Name	ECA&D (European Climate Assessment & Dataset)
Origin	ECA&D Project Team, KNMI, eca@knmi.nl
Spatial Characteristics	10233 meteorological stations are collected throughout Europe and the Mediterranean, providing observations of up to 12 meteorological parameters each.
Temporal Characteristics	Time series of daily data

Software Readiness	Metadata	User Documentation	Uncertainty Characterisation	Public Access, Feedback, Update	Usage
1-2	2-5	1-6	1-6	1-6	6

The European Climate Assessment & Dataset (http://eca.knmi.nl/) provides information on changes in weather and climate extremes, as well as the daily dataset needed to monitor and analyse these extremes. It is based on the contributions of long-term high quality daily observational data from 61 National Meteorological and Hydrological Services, observatories and universities from Europe and the Mediterranean. The special value of this data set is on one hand in the comprehensive collection, which makes it most useful for many applications, and on the other hand in derived indices.

The software used to produce this dataset is under research grade. Improvements are suggested in this category.

The Metadata is under research or initial operational capability with limited collection level metadata and sufficient file level metadata with location level metadata.

Maturity related to User Documentation spans the whole range 1 – 6. Some formal validation report would raise the minimum score to 3. Comprehensive tools for cross-comparisons and plenty of scientific papers addressing specific parameters are available from the website.

In the Uncertainty Characterization, the low scores hinge on the input data. However, the automated quality monitoring is established and results are fed back to metadata.

Public Access, Feedback, Update is generally under full operational capability except that there is no version control established.

The dataset are extensively used for research and policy making.

6.2.9 EOBS

Name	E-OBS dataset (gridded data based on station observations in Europe) Version 9
Origin	ECA&D Project Team, KNMI, eca@knmi.nl
Spatial Characteristics	area: 25N-75N x 40W-75E; grids: 0.25 and 0.5 degree regular lat-lon grid, as well as 0.22 and 0.44 degree rotated pole grid
Temporal Characteristics	Daily, 1950-01-01 to 2013-06-30

Software Readiness	Metadata	User Documentation	Uncertainty Characterisation	Public Access, Feedback, Update	Usage
1-5	5-6	4-6	1-5	5-6	6

E-OBS provides gridded fields based on the ECA&D station collection. More details at http://eca.knmi.nl/.

Software readiness scores vary from research to full operational capability. No coding standards are used, moderate software documentation is available, the software can be installed operationally by a third party and produces numerically reproducible results, and the PI confirms no security issues.

The Metadata for this dataset is in full operational grade.

Maturity for User Documentation is generally under full operational capability. Comprehensive descriptions and peer-reviewed papers on scientific methods, validation and inter-comparison are available. Product user guide and operations concept are regularly updated.

Uncertainty Characterization is under research grade except for automated quality monitoring which is fully established. Low scores here are hinges on the input data which points to the need of uncertainty characterization by the original data providers or a community effort.

Public Access, Feedback, Update is under full operational capability with data, documentation, and source code are archived and under fully established version control. ICDRs are being produced with fully established user feedback mechanisms.

The dataset is extensively used for research and policy making.

6.2.10 Heleosat Surface Radiation

Name	CM SAF Surface Radiation MVIRI Data Set
Origin	CM SAF; contact.cmsaf@dwd.de
Spatial Characteristics	METEOSAT 0 deg disk; 0.03x0.03 degree grid
Temporal Characteristics	Hourly, daily, and monthly means; 1983-2005

Category	Software Readiness	Metadata	User Documentation	Uncertainty Characterisation	Public Access, Feedback, Update	Usage
Range	2-4	4-6	5-6	3-4	5-6	4-5

The CM SAF Surface Radiation MVIRI Data Set is a satellite-based climatology of the surface irradiance, the surface direct irradiance and the effective cloud albedo derived from satellite-observations from the visible channel of the MVIRI instruments onboard the geostationary Meteosat satellites. The dataset was generated for climate applications.

Overall, this dataset is in the initial to full operational capability range, except for the "Security" category in the Software Readiness sub-matrix. This can be overcome by establishing a process to check security aspects in the SAF framework.

Scores in the Software Readiness could be improved by increasing compliance to coding standards and by completing the software documentation.

The dataset has high scores in Metadata and User Documentation, which are in full operational capability.

The scores for Uncertainty Characterisation are in the initial operational capability. Low scores in the Standards are only due to non-existence of procedures to achieve SI traceability. Participation in international data assessment(s), quantifying spatial and temporal error co-variances, and full implementation of automated quality monitoring will take this dataset towards full operational capability.

The scores for Public Access, Feedback, and update are already in full operational capability. The dataset is regularly updated taking into account user feedbacks.

The dataset has high scores for usage which are corroborated by separate lists of research and decision support applications in the Data Record Description.

6.2.11 CLARA-A1 Surface Radiation

Name	CM SAF CLARA A1 surface radiation
Origin	CM SAF; contact.cmsaf@dwd.de
Spatial Characteristics	Global, 0.25x0.25 grid

Temporal Characteristics	daily and monthly mean, 1982-2009

Category	Software Readiness	Metadata	User Documentation	Uncertainty Characterisation	Public Access, Feedback, Update	Usage
Range	3 – 5	4 - 5	4 - 5	3 - 4	5	2 - 4

CM SAF CLARA A1 surface radiation properties dataset is a global dataset of surface radiation derived from measurements of the series of Advanced Very High Resolution Radiometer (AVHRR) on the NOAA satellite series including METOP-A satellite.

Software readiness is generally in the initial operational category. Coding standards are partially applied, complete software installation/user manual available, reproducibility and portability confirmed by 3rd party, and the software passes data provider's security review.

Metadata are compliant with international standards, file and collection level metadata are sufficient to understand and use the data with enhanced discovery metadata and limited location level metadata. Thus this category is under initial operations category.

User documentation is under full operational capability. Comprehensive description on scientific methods and journal papers are published on this. Reports and peer-reviewed publications on comprehensive validation and inter-comparison are available from the data provider. Product user guide is regularly updated and maintained by data provider. Operations concept and description of practical implementation are available from the data provider.

Uncertainty characterisation falls into initial operations category. Standard uncertainty nomenclature is applied. Comprehensive validation is done and inter-comparisons against other datasets have been performed. Comprehensive and quantitative estimates of uncertainty are provided with the product and automated quality monitoring is partially applied.

Public access, feedback, and update are fully operational with dataset, documentation, and source code are version controlled and archived by the data provider and are publicly available (except the software). The dataset is regularly updated taking into account user feedbacks.

The data are used for research and citations are occurring.

6.2.12 CLARA-A1 Cloud Properties

Name	CM SAF CLARA A1 cloud properties
Origin	CM SAF; contact.cmsaf@dwd.de
Spatial Characteristics	Global, 0.25 x 0.25 grid
Temporal Characteristics	daily and monthly mean, 1982 – 2009

Category	Software Readiness	Metadata	User Documentation	Uncertainty Characterisation	Public Access, Feedback, Update	Usage
Range	3 – 5	4 - 5	4 - 5	3 - 4	5	2 - 4

CMSAF CLARA A1 cloud properties dataset is a global dataset of surface radiation derived from measurements of the series of Advanced Very High Resolution Radiometer (AVHRR) on the NOAA satellite series including METOP-A satellite.

Software readiness is generally in the initial operational category. Coding standards are partially applied, complete software installation/user manual available, reproducibility and portability confirmed by 3rd party, and the software passes data provider's security review.

Metadata are compliant with international standards, file and collection level metadata are sufficient to understand and use the data with enhanced discovery metadata and limited location level metadata. Thus this category is under initial operations category.

User documentation is under full operational capability. Comprehensive description on scientific methods and journal papers are published on this. Reports and peer-reviewed publications on comprehensive validation and inter-comparison are available from the data provider. Product user guide is regularly updated and maintained by data provider. Operations concept and description of practical implementation are available from the data provider.

Uncertainty characterisation falls into initial operations category. Standard uncertainty nomenclature is applied. Comprehensive validation is done and inter-comparisons against other datasets have been performed. Comprehensive and quantitative estimates of uncertainty are provided with the product and automated quality monitoring is partially applied.

Public access, feedback, and update are fully operational with dataset, documentation, and source code are version controlled and archived by the data provider and are publicly available (except the software). The dataset is regularly updated taking into account user feedbacks.

The data are used for research and citations are occurring.

6.3 Oceanic datasets

In this session we present assessment results of the oceanic data sets.

6.3.1 Baltic Sea Automated Sea Ice

Name	SEAICE_BAL_SEAICE_L4_NRT_ OBSERVATIONS_011_011/004
Origin	MyOcean, FMI, Juha.Karvonen@fmi.fi
Spatial Characteristics	1 km grid
Temporal Characteristics	Baltic Sea ice season

Category	Software Readiness	Metadata	User Documentation	Uncertainty Characterisation	Public Access, Feedback, Update	Usage
Range	1 – 5	5 – 6	3 – 6	3 – 5	2 – 5	3 – 6

This dataset is a part of operational sea ice services in FMI and covers automated ice thickness and concentration charts which are produced by SAR data. Operational production is in FMI while data is provided by MyOcean project. The parameters are based on ice chart produced on daily basis during the Baltic Sea ice season and show the ice concentration in a 1 km grid. Ice thickness chart (ITC) is a product based on the most recent available ice chart (IC) and a SAR image.

Software Readiness is generally under initial operational capability: software is portable, and third party can install with numerically reproducible results. The software has moderate documentation. However, neither coding standards are applied nor security is evaluated.

The Metadata is under full operational capability.

User Documentation is from initial to full operational capability with reports and papers on scientific methods, comprehensive validation and data assessment, regularly updated product user guide, and comprehensive description of operations concept including practical implementation.

Uncertainty Characterisation is generally under initial operational capability with application of standard uncertainty nomenclature, partial establishment of SI traceability, comprehensive validation and inter-comparison, comprehensive information on uncertainties, and partially implemented automated quality monitoring.

Public Access, Feedback, and Archive category is also generally under initial operational capability with version controlled data and documentation are archived by the data provider. The data are regularly and operationally updated with feedbacks collected from scientific community.

The data are moderately used for research and extensively used for policy making.

6.3.2 ESA-SST-CCI-Analysis

Name	ESA SST CCI Analysis long-term product V 1.0
Origin	ESA-CCI; c.j.merchant@reading.ac.uk
Spatial Characteristics	Global; 0.05° lat-lon grid resolution
Temporal Characteristics	~20 years; Daily

Category	Software Readiness	Metadata	User Documentation	Uncertainty Characterisation	Public Access, Feedback, Update	Usage
Range	1 – 4	3 – 5	2 – 5	4 – 5	3 – 5	1 – 3

ESA SST CCI Analysis long-term product version 1 are derived from infra-red imagery obtained from several Earth-observing satellite missions, combined to give daily, spatially complete information over the global oceans over 20 years. The principal recommended applications are for climate research applications requiring ~20 years of stable, low bias

records of SST. The dataset is particularly valuable if a representation of global SST is required that is independent of in situ SST measurements.

The software used to produce this data set is under research to initial operational capability. Coding standards are not applied and security is not checked. Software is sufficiently documented and the PI affirms numerical reproducibility and portability.

The metadata for this dataset is generally under initial operational capability. International metadata standards are applied, collection level metadata are sufficient understand and use the data, and there are complete location level metadata at the file level.

User documentation is also generally under initial operational capability. Comprehensive documentation on scientific methods, validation, and operations concept is available from the PI. Product user guide is maintained by the data provider and is being regularly updated.

Uncertainty Characterisation generally is under initial operational capability. Standard uncertainty nomenclature is applied and procedures to establish SI traceability are defined. Comprehensive validation is done and the product is inter-compared with other CDRs. Qualitative information of uncertainty is provided with temporal and spatial error covariance. Automated quality monitoring is partially established.

Public Access, Feedback, and Update is also in the initial operational category. Data, documentation, and source code are under PI's versioning and archived by the data provider. The product is irregularly updated by using the feedbacks from scientific community collected by PI and data provider.

The benefits of using this dataset for research have been demonstrated.

6.3.3 ESA-SST-CCI-AVHRR

Name	ESA SST CCI AVHRR L2P long-term product V 1.0
Origin	ESA-CCI; c.j.merchant@reading.ac.uk
Spatial Characteristics	Global; ~4 km at nadir view
Temporal Characteristics	~20 years; Instantaneous

Category	Software Readiness	Metadata	User Documentation	Uncertainty Characterisation	Public Access, Feedback, Update	Usage
Range	1 – 4	3 – 5	2 – 5	3 – 5	3 – 5	1 – 3

ESA SST CCI AVHRR L2P long-term product version 1.0 provides global Sea Surface Temperature (SST) as estimated from Advanced Very High Resolution Radiometer (AVHRR) imagery.

The principal recommended applications are climate research applications requiring ~20 years of global SSTs observed by satellite, with no gap-filling/interpolation. Since the SSTs are harmonized independently of in situ observations, use is recommended for applications

where it is beneficial to have SST datasets that are independent. Skin and depth SSTs are distinguished and both are provided.

The software used to produce this data set is under research to initial operational capability. Coding standards are not applied and security is not checked. Software is sufficiently documented and the PI affirms numerical reproducibility and portability.

The metadata for this dataset is generally under initial operational capability. International metadata standards are applied, collection level metadata are sufficient understand and use the data, and there are complete location level metadata at the file level.

User documentation is also generally under initial operational capability. Comprehensive documentation on scientific methods, validation, and operations concept is available from the PI. Product user guide is maintained by the data provider and is being regularly updated.

Uncertainty Characterisation generally is under initial operational capability. Standard uncertainty nomenclature is applied and procedures to establish SI traceability are defined. Comprehensive validation is done and the product is inter-compared with other CDRs. Qualitative information of uncertainty is provided with temporal and spatial error covariance.

Public Access, Feedback, and Update is also in the initial operational category. Data, documentation, and source code are under PI's versioning and archived by the data provider. The product is irregularly updated by using the feedbacks from scientific community collected by PI and data provider.

The benefits of using this dataset for research have been demonstrated.

6.3.4 HadISST1

Name	HadISST
Origin	Met Office – HadObs; nick.rayner@metoffice.gov.uk
Spatial Characteristics	Global oceans; 1 degree lat-lon grid
Temporal Characteristics	1870 to date; monthly

Category	Software Readiness	Metadata	User Documentation	Uncertainty Characterisation	Public Access, Feedback, Update	Usage
Range	1 – 3	3 – 4	1 – 6	1 – 5	3 – 5	4 – 6

The HadISST1 data set is a blend of historical SST and modern SST observations from ships, buoys, drifters, etc. and sea ice observations, partly from historical ship- and air-borne and partly from satellite data.

The code used produce the dataset is generally under research category. Coding standards are identified, but not applied, minimal documentation is available from the PI, portability and numerical reproducibility are confirmed by the PI, and the security is not evaluated.

Metadata is under initial operational capability. International standards are systematically applied, sufficient collection and file level metadata to understand and use the data without external assistance. Also, limited location-level metadata are available.

The scores for User Documentation vary from research to full operational capability. Highest score is for the formal description of scientific methods due to comprehensive description of methods maintained by the data provider and journal papers published when the product is updated. Comprehensive validation report, publications, and data assessment results are available. Limited description of operations concept is available, but there is no formal product user guide.

The uncertainty characterisation scores also range from research capability to full operational capability. Standard uncertainty nomenclature is applied, comprehensive validation is done and the data product has been part of international assessments, and automated quality monitoring is partially implemented.

Public Access, Feedback, and Update is under initial to full operational capability. The data, documentation, and source code are archived by the data provider and under institution's version control. The data set is irregularly updated following the updates collected from scientific community.

The dataset has been used for research and citations have been occurring. The dataset is extensively used for decision making and influence on policy making is demonstrated.

6.3.5 ESA Ocean Colour CCI

Name	Ocean Colour CCI
Origin	ESA-CCI; Shubha Sathyendranath, PML; help@esa-oceancolour-cci.org
Spatial Characteristics	Global; 4 km
Temporal Characteristics	1997 – 2014; daily composite

Category	Software Readiness	Metadata	User Documentation	Uncertainty Characterisation	Public Access, Feedback, Update	Usage
Range	1 – 4	6	3 – 6	3 – 5	3 – 6	2 – 3

The ESA-CCI Ocean Colour ECV dataset provides ocean colour data, with a focus on Case 1 (Open Ocean, not Coastal) waters, which can be used by climate change prediction and assessment models. The dataset is created by band-shifting and bias-correcting atmosphere corrected MERIS and Aqua MODIS data to match SeaWiFS data, merging the datasets and computing per-pixel uncertainty estimates.

Software readiness is generally under initial operational capability: coding standards are partially applied and compliance results available, moderate description of the software available and PI affirms portability and numerical reproducibility. However, security aspects of the software are not evaluated.

Metadata is under full operational capability.

User Documentation is under initial or full operational capability. Comprehensive descriptions and papers on methods, validation, and operations concepts are maintained by data provider and product user guide is regularly updated.

Uncertainty Characterisation is also under initial or full operational capability. Standard uncertainty nomenclature is applied and SI traceability is partially established, comprehensive validation is done, quantitative estimates of uncertainty provided within the product characterising more or less uncertain data points, temporal and spatial error covariance are quantified, and automated quality monitoring is partially established.

Public Access, Feedback, Update is also from initial to full operational capability with data and documentation are under version control and archived by the data provider. ICDRs are being produced by taking into account user feedbacks.

The dataset is being moderately used.

6.4 Cryosphere

We present the assessments of Cryosphere datasets here.

6.4.1 Sea Ice Volume Flux

Name	Sea Ice volume flux through Fram Strait 79N
Origin	NERSC; kjetil.lygre@nersc.no
Spatial Characteristics	At ~79 deg N latitude with 1 deg step from 5 deg E to 15 deg W
Temporal Characteristics	6 Campaigns from 2005 to 2007

Category	Software Readiness	Metadata	User Documentation	Uncertainty Characterisation	Public Access, Feedback, Update	Usage
Range	2	3	2	2	3	2

Sea ice volume flux along 79 degrees north has been computed with a 1 degree longitudinal resolution. The flux was computed as a deliverable during the EU Monarch-A project to be compared to various climate parameters such as sea ice cover in the Arctic, ice thickness distribution, wind field etc. Grid cells for all data sets are centered on 79 deg N and from 5 deg E to 15 deg W. Ice area flux extracted from the time series of K. Kloster based on SAR and passive microwave observations. 4 columns contain: (1) ice concentration (%) (2) ice displacement per time interval 3-4 days (km), (3) azimuth angle (deg) (4) area ice flux (km2/day/deg). All values are given for ~79 deg N latitude with 1 deg step from 5 deg E to 15 deg W. Volume flux is computed by multiplying area flux and IceSAT sea ice thickness.

Software used to create this dataset is research grade.

Metadata is under initial operational capability with defined Standards and sufficient to use and understand the data and to extract discovery metadata.

User Documentation is under research grade with comprehensive scientific description of the methodology, report on limited validation, and limited product user guide available from PI.

Uncertainty Characterisation is also under research grade with Standard uncertainty nomenclature defined, limited validation done, and limited information.

Data are available from PI and irregular updates using feedback through scientific exchange which makes this dataset falls under initial operational capability for Public Access, Feedback, and Update.

The data are not very well used.

6.4.2 Cryoland Glacier Products

Name	Cryoland Glacier products
Origin	Cryoland Project; Dr. Thomas Nagler, ENVEO IT GmbH, thomas.nagler@enveo.at
Spatial Characteristics	Regional/local
Temporal Characteristics	Annual

Category	Software Readiness	Metadata	User Documentation	Uncertainty Characterisation	Public Access, Feedback, Update	Usage
Range	1 – 5	6	2 – 5	2 – 3	2 – 5	2 – 3

Cryoland Glacier products datasets is a part of Cryoland Project, operational as a Copernicus service on snow and land ice. The product glacier outlines / area is provided on user request according to the internationally accepted GLIMS standards. The products are: Glacier Outlines/Area Map, Snow/Ice Maps, Glacier Lake, and Glacier Ice Velocity Map.

The software used to generate the dataset is under full operational capability except that the portability and numerical reproducibility are evaluated.

The metadata for the dataset is under full operational capability.

The scores for user documentation vary from research to full operational capability. Comprehensive descriptions and papers on scientific methods are available, but only limited documentation on validation exists. User guide and operations concepts have highest scores.

Uncertainty characterisation is generally under research capability.

Public Access, Feedback, Update scores are varying from research to full operational capability with fully version controlled dataset is available from the PI. The dataset is regularly updated by the PI using feedbacks obtained from established mechanisms.

Benefits for research are demonstrated by publications and potential benefits for decision making are identified.

6.5 Land surface

We present the assessments of land surface datasets here.

6.5.1 METEOSAT Surface Albedo

Name	Meteosat Surface Albedo
Origin	EUMETSAT; ops@eumetsat.int
Spatial Characteristics	METEOSAT first generation coverage; 2.5 km at sub-satellite point
Temporal Characteristics	1982 – 2011; 10 day composite

Category	Software Readiness	Metadata	User Documentation	Uncertainty Characterisation	Public Access, Feedback, Update	Usage
Range	2 – 5	3 – 5	3 – 5	3 – 4	3 – 5	1 – 4

METEOSAT Surface Albedo (MSA) data set is derived from measurements of the METEOSAT Visible and Infra Red Imager (MVIRI) instrument on METEOSAT first generation satellites. Surface albedo is generally defined as the instantaneous ratio of surface-reflected radiation flux to incident radiation flux (dimensionless).

Scores for Software Readiness varies from research to full operational capability: coding standards are defined, but not applied, software documentation is complete, software can be installed by 3rd party with numerically reproducible results, and passes data providers security review.

Metadata is generally under initial operational capability. International standards are systematically applied, collection level metadata are sufficient to understand and use the data, and there is complete file level metadata.

User Documentation is also generally under initial operational capability. Comprehensive scientific description is maintained by data provider and journal papers are published. Report on comprehensive validation is available. Comprehensive user guide is available from the PI and description of operations concept is available.

Uncertainty characterisation is under initial operational capability. Standard uncertainty nomenclature is applied, comprehensive validation is done and quantitative uncertainty estimates available characterising more or less uncertain data point. Automated quality monitoring is partially applied.
Public Access, archive and feedback is under initial or full operational capability. Data, documentation, and source code are under fully established version control and archived by the data provider. The dataset is irregularly updated using user feedbacks.

The product has been used for research and citations are occurring.

6.5.2 GEOV1 Leaf Area index (LAI)

Name	GEOV1 LAI
Origin	Copernicus Global Land Service; rl@hygeos.com
Spatial Characteristics	Global, 1 km grid
Temporal Characteristics	1999 to present; 10 daily means

Category	Software Readiness	Metadata	User Documentation	Uncertainty Characterisation	Public Access, Feedback, Update	Usage
Range	6	6	3-6	6	4-6	3

The GEOV1 Leaf Area index (LAI) time series is provided by the European Copernicus Global Land Service and is based on SPOT-VEGETATION satellite observations, from 1999 until May 2014 (end of lifetime of SPOT-VGT). Since January 2014, the same data set is derived from the PROBA-V sensor, which is the successor of SPOT-VGT.

The software used to produce the dataset has highest maturity in all categories which implies the codes are completely compliant with standards and which has been verified, the software is fully documented with complete installation/user manual and the software is a turnkey system. Also the software is checked for security issues and passed all the tests. This implies that the system is in a fully mature state and no improvements needed.

The dataset has highest scores in the metadata category as well. It has compliance with international standards and the compliance is systematically checked by the data provider. The dataset also has complete metadata collection and file levels. This implies that the system is in a fully mature state and no improvements needed.

In the User Documentation category, the data set has highest scores for all but formal description of operations concept. This implies the dataset has all scientific documentation as complete as possible.

The data has highest scores for all sub-categories in Uncertainty Characterisation. This implies no more additional efforts required for this.

Data, documentation, and source code are archived and fully under version control. Making the source code available to the public will make it achieve the highest scores. The data are produced in a continuous manner with availability of Interim Climate Data Records and there are regular mechanisms to obtain user feedbacks and incorporate them in updates.

The data set is used moderately for research and decision making.

6.5.3 GEOV1 fAPAR

Name	GEOV1 fAPAR
Origin	Copernicus Global Land Service, rl@hygeos.com
Spatial Characteristics	Global, 1km resolution
Temporal Characteristics	1999 to present; 10 daily means

Category	Software	Metadata	User	Uncertainty	Public Access,	Usage

	Readiness		Documentation	Characterisation	Feedback, Update	
Range	6	6	3-6	6	4-6	3

The GEOV1 fraction of Absorbed Photosynthetically Active Radiation (fAPAR) time series is provided by the European Copernicus Global Land Service and is based on SPOT-VEGETATION satellite observations, from 1999 until May 2014 (end of lifetime of SPOT-VGT). Since January 2014, the same data set is derived from the PROBA-V sensor, which is the successor of SPOT-VGT.

The Copernicus Global Land Service products fAPAR, LAI and surface albedo are processed using the same processing chain (except for the algorithm part) and distributed through the same portal. Also, the metadata and user documentation are treated in the same way. Therefore, the maturity of these products is very similar or even identical.

The software used to produce the dataset has highest maturity in all categories which implies the codes are completely compliant with standards and which has been verified, the software is fully documented with complete installation/user manual and the software is a turnkey system. Also the software is checked for security issues and passed all the tests. This implies that the system is in a fully mature state and no improvements needed.

The dataset has highest scores in the metadata category as well. It has compliance with international standards and the compliance is systematically checked by the data provider. The dataset also has complete metadata collection and file levels. This implies that the system is in a fully mature state and no improvements needed.

In the User Documentation category, the data set has highest scores for all but formal description of operations concept. This implies the dataset has all scientific documentation as complete as possible.

The data has highest scores for all sub-categories in Uncertainty Characterisation. This implies no more additional efforts required for this.

Data, documentation, and source code are archived and fully under version control. Like for the other Copernicus Global Land Service products, making the source code available to the public will make it achieve the highest scores. The data are produced in a continuous manner with availability of Interim Climate Data Records and there are regular mechanisms to obtain user feedbacks and incorporate them in updates.

The data set is used moderately for research and decision making.

6.5.4 GEOV1 Surface Albedo

Name	GEOV1 Surface Albedo
Origin	Copernicus Global Land Service, rl@hygeos.com
Spatial Characteristics	Global, 1km resolution
Temporal Characteristics	1999 to present; 10 daily means

Category	Software Readiness	Metadata	User Documentation	Uncertainty Characterisation	Public Access, Feedback, Update	Usage
Range	6	6	3-6	4-6	4-5	2

The Copernicus Global Land Service products fAPAR, LAI and surface albedo are processed using the same processing chain (except for the algorithm part) and distributed through the same portal. Also, the metadata and user documentation are treated in the same way. Therefore, the maturity of these products is very similar or even identical.

The software used to produce the dataset has highest maturity in all categories which implies the codes are completely compliant with standards and which has been verified, the software is fully documented with complete installation/user manual and the software is a turnkey system. Also the software is checked for security issues and passed all the tests. This implies that the system is in a fully mature state and no improvements needed.

The dataset has highest scores in the metadata category as well. It has compliance with international standards and the compliance is systematically checked by the data provider. The dataset also has complete metadata collection and file levels. This implies that the system is in a fully mature state and no improvements needed.

In the User Documentation category, the data set has highest scores for all but formal description of operations concept. This implies the dataset has all scientific documentation as complete as possible.

The data has highest scores for all sub-categories in Uncertainty Characterisation, except for validation. Participating in international data set assessment(s) and incorporating feedbacks into the product development cycle would increase the maturity of the uncertainty characterisation.

Data, documentation, and source code are archived and fully under version control. Like for the other Copernicus Global Land Service products, making the source code available to the public will make it achieve the highest scores. The data are produced in a continuous manner, but without availability of Interim Climate Data Records and there are regular mechanisms to obtain user feedbacks and incorporate them in updates.

The data set is started being used for research and decision making, for which potential benefits have been identified.

6.5.5 ESA-CCI Soil Moisture

Name	ESA-CCI Soil Moisture
Origin	ESA CCI; ww@ipf.tuwien.ac.at
Spatial Characteristics	Global; 0.25°~1.5°
Temporal Characteristics	1978 – 2010; between 1 and 3days

Category	Software Readiness	Metadata	User Documentation	Uncertainty Characterisation	Public Access, Feedback, Update	Usage
Range	3	4-5	4-6	5	3-4	4

This Soil Moisture product is being developed and provided in the frame of the ESA Climate Change Initiative. The global soil moisture data set has been generated using active and passive microwave space borne instruments and covers the 32 year period from 1978 to 2010.

The software used to produce the Soil Moisture product has an overall maturity that corresponds with the level of Initial Operation Capability, which is in agreement with the data set development status.

The metadata associated with the products has a somewhat higher maturity, which is fully mature for the subcategory File level. Using international standards for the Collection level and regular update and checks of the metadata would increase the maturity of this category.

The User Documentation reaches the highest maturity for the formal description of the scientific methodology. The other user documentation subcategories have moderate maturity. Again, this is in line with the product development status. A formal description of the operations concept and its implementation, and regularly updates of the product user manual and validation results to the user would increase the maturity.

The Uncertainty characterization reaches nearly the highest maturity. Increasing the effort in international data assessments and organizing automated quality monitoring would result in the highest maturity.

The data and documentation are archived and available from the data provider and the version control is institutionalised. A gain in maturity could be reached by implementing a full version control and archiving the source code. User feedback is requested, but no feedback mechanisms are put in place and only ad hoc updates to the data set is performed.

The data is used extensively in research and moderately in decision making.

6.5.6 ESA-CCI Land Cover

Name	ESA LandCover_CCI datasets
Origin	ESA CCI; pierre.defourny@uclouvain.be
Spatial Characteristics	See text below
Temporal Characteristics	See text below

Category	Software Readiness	Metadata	User Documentation	Uncertainty Characterisation	Public Access, Feedback, Update	Usage
LC maps	2 – 3	2 – 5	2 – 4	1 – 4	1 – 3	1 – 2
LC cond	1 – 2	2 – 3	1 – 4	1	1 – 3	1 – 2
SR	2 – 5	3 – 5	2 – 5	3 – 4	1 – 4	1 – 2

ESA –CCI Global land cover databases consist of:

- 3 global LC maps at 300m spatial for 3 epochs cantered on the years 2010 (2008-2012), 2005 (2003-2007) and 2000 (1998-2002);
- 3 LC condition products which describe the natural variability of the land surface. They correspond to three aggregated land surface seasonality products: the NDVI, the Burned Areas (BA) and the snow.
- Global surface reflectance (SR) time series made of 7-day composites from 2003 to 2012.

The assessment is done taking into account the scores for the 3 datasets together.

Software Readiness is under research grade for LC datasets and generally under full operational capability for SR dataset.

Metadata scores: Generally under initial operational capability for all datasets.

User Documentation: Generally under research grade for LC datasets and under initial operational capability for SR dataset.

Uncertainty Characterisation: No uncertainty characterisation is done for LC conditions, research to initial operational capability for LC maps, and initial operational capability for SR products.

Public Access, Feedback, Update: Research to initial operational capability for LC datasets and initial operational capability for SR products.

Usage: Benefits for research applications identified for all datasets.

6.5.7 CLARA-A1-SAL (Surface Albedo)

Name	CLARA-A1-SAL (Surface Albedo) Edition 1
Origin	CMSAF, FMI; aku.riihela@fmi.fi
Spatial Characteristics	Global; 0.25x0.25 degrees lat-lon grid
Temporal Characteristics	1982-2009; five-day and monthly mean

Category	Software Readiness	Metadata	User Documentation	Uncertainty Characterisation	Public Access, Feedback, Update	Usage
Range	1 – 3	3 – 4	4 – 5	3 – 4	4 – 5	2 – 3

The CLARA-A1-SAL data set is derived from measurements of AVHRR visible/infrared imagers on NOAA and Metop-A satellites. The quantity described by the CLARA-A1-SAL dataset is the broadband Directional-Hemispherical Reflectance (DHR) of Earth's surface, also called Black-sky albedo.

The software used produce the dataset is generally under initial operational capability, but security issues are not evaluated yet.

The Metadata is under initial operational capability: International standards are systematically applied and there are sufficient file and collection level metadata to understand and use the data.

User Documentation is under initial or full operational capability. Comprehensive descriptions and journal papers available on scientific methods, reports on comprehensive validation and inter-comparisons are available, product user guide is regularly updated, and comprehensive operations concept is available.

Uncertainty Characterisation is under initial operational capability: Standard uncertainty nomenclature is applied, comprehensive validation is done and uncertainty arising from systematic and random effects in the measurement is available, and automated quality monitoring is partially applied.

Public Access, Feedback, Update is under initial or full operational capability: Data, documentation, and source code are under fully established version control and archived by the data provider. The dataset is regularly being updated by using feedback through established mechanisms.

Benefits for research applications are demonstrated in publications.

6.5.8 GlobSnow Snow Extent

Name	GlobSnow Snow Extent (SE) version 1.2
Origin	FMI; kari.luojus@fmi.fi
Spatial Characteristics	Global; 1 km
Temporal Characteristics	1995 – present;

Category	Software Readiness	Metadata	User Documentation	Uncertainty Characterisation	Public Access, Feedback, Update	Usage
Range						

The European Space Agency (ESA) Data User Element (DUE) GlobSnow Snow Extent (SE) product set version 1.2 for the Northern Hemisphere represents information on snow coverage retrieved from ERS-2 ATSR-2 and Envisat AATSR from 1995 until present.

The software used to generate this dataset is fully compliant with the coding standards, moderately documented, portable and results are numerically reproducible, no security issues are found. These make the Software Readiness suitable for initial operational capability.

Metadata and User Documentation get highest scores for all aspects.

Uncertainty Characterisation is under initial to full operational capability. Standard uncertainty nomenclature is used and SI traceability is partially established, comprehensive

validation and inter-comparison were done. Quantitative estimates of uncertainty provided within the product characterising more or less uncertain data points. Automated monitoring is fully established and results fed back to metadata.

Public Access, feedback and update is also under initial to full operational capability. Data and documentation are under fully established version control and archived and available to the public from Data Provider. The dataset is regularly updated using fully established feedback mechanisms.

The data are extensively used for research and for decision making.

6.5.9 H-SAF Daily Snow Cover (H10)

Name	H-SAF Daily Snow Cover (H10/SN-OBS-1)
Origin	H-SAF, FMI; matias.takala@fmi.fi
Spatial Characteristics	The H-SAF area [25-75°N lat, 25°W - 45° E lon]
Temporal Characteristics	2004 – present; Daily

Category	Software Readiness	Metadata	User Documentation	Uncertainty Characterisation	Public Access, Feedback, Update	Usage
Range	2 – 5	5 – 6	6	3 – 6	4 – 6	2

H-SAF Daily Snow Cover (H10/SN-OBS-1) data set is a product under H-SAF Project of EUMETSAT SAF Network. As the hydrological project of the SAF Network, snow parameters are the key parameters to the project. H10 dataset gives the snow information on a pixel basis for SAF Europe area as a snow mask, having binary information for the snow coverage on cloud-free and non-dark locations.

Software Readiness varies from research grade to full operational capability: partially compliant to coding standards, minimal documentation is available, PI affirms portability and numerical reproducibility, and continues to pass data provider's security review.

Metadata and User Documentation are under full operational capability.

Uncertainty Characterisation is under initial to full operational capability: Standard uncertainty nomenclature is applied, comprehensive validation and inter-comparison are done, comprehensive quantification of more or less uncertain points are available, and temporal and spatial error covariance are quantified. Automated monitoring is fully established and results fed back to metadata.

Public Access, feedback and update is also under initial to full operational capability. Data and documentation are under institution's version control and archived and available to the public from Data Provider. The dataset is regularly updated using fully established feedback mechanisms.

The data are not well utilised for research and policy making.

6.5.10 H-SAF Daily Effective Snow Cover (H12)

Name	H-SAF Daily Effective Snow Cover (H12/SN-OBS-3)
Origin	H-SAF, FMI; matias.takala@fmi.fi
Spatial Characteristics	The H-SAF area [25-75°N lat, 25°W - 45° E lon]
Temporal Characteristics	Daily

Category	Software Readiness	Metadata	User Documentation	Uncertainty Characterisation	Public Access, Feedback, Update	Usage
Range	2 – 6	5 – 6	6	3 – 6	4 – 6	2

H-SAF Daily Effective (Fractional) Snow Cover (H12/SN-OBS-3) data set gives the snow information on a pixel basis for SAF Europe area as a snow mask, having fractional information for the snow coverage on cloud-free and non-dark locations.

Software Readiness varies from research grade to full operational capability: partially compliant to coding standards, minimal documentation is available, PI affirms portability and numerical reproducibility, and continues to pass data provider's security review.

Metadata and User Documentation are under full operational capability.

Uncertainty Characterisation is under initial to full operational capability: Standard uncertainty nomenclature is applied, comprehensive validation is done, comprehensive quantification of more or less uncertain points are available, and temporal and spatial error covariance are quantified. Automated monitoring is fully established and results fed back to metadata.

Public Access, feedback and update is also under initial to full operational capability. Data and documentation are under institution's version control and archived and available to the public from Data Provider. The dataset is regularly updated using fully established feedback mechanisms.

The data are not well utilised for research and policy making.

6.5.11 LSA-SAF Daily Snow Cover

Name	LSA-SAF Daily Snow Cover
Origin	LSA-SAF, FMI; Niilo.Siljamo@fmi.fi
Spatial Characteristics	3 km at sub-satellite point
Temporal Characteristics	Daily

Category	Software Readiness	Metadata	User Documentation	Uncertainty Characterisation	Public Access, Feedback, Update	Usage
Range	1 – 2	6	6	4 – 6	4 – 5	4 – 5

LSA-SAF Daily Snow Cover data set are snow cover map produced from MSG data contains a classification of each surface pixel or resolution cell into one (and only one) of the following classes:

- totally snow covered
- partially snow covered
- no snow
- unclassified
- non-processed
- water (sea, lake, river etc.)

An additional set of quality/processing flags for each pixel indicates the certainty of the classification and integration and also gives information on the processing and conditions.

The software used to generate this dataset is under research category: coding standards are not applied, minimal software documentation is available, PI affirms portability and numerical reproducibility, and security issues are not evaluated.

Metadata and User Documentation are under full operational capability.

Uncertainty Characterisation is under initial to full operational capability: Standard uncertainty nomenclature is applied and SI traceability is partially established, comprehensive validation and inter-comparison are done, comprehensive quantification of more or less uncertain points are available, and temporal and spatial error covariance are quantified. Automated monitoring is fully established and results fed back to metadata.
Public Access, feedback and update is also under initial to full operational capability. Data and documentation are under institution's version control and archived and available to the public from Data Provider. The dataset is regularly updated using fully established feedback mechanisms.

The dataset is widely used for research and decision making.

6.5.12 Global Fire Assimilation System

Name	Global Fire Assimilation System
Origin	Copernicus Atmosphere Monitoring Service (CAMS); j.kaiser@mpic.de
Spatial Characteristics	Global, .1 and .5 deg
Temporal Characteristics	2000 – present; Daily

Category	Software Readiness	Metadata	User Documentation	Uncertainty Characterisation	Public Access, Feedback, Update	Usage
Range	2 – 3	2 – 6	1 – 6	1 – 4	4 – 6	3 – 4

Global Fire Assimilation System (GFAS) provides global daily estimates of open biomass burning, a.k.a. vegetation fires with 0.5deg and 0.1deg resolutions. It is based on satellite

observations of Fire Radiative Power (FRP), currently by NASA's two polar orbiting MODIS instruments. The dataset is primarily intended as lower boundary condition input for atmospheric composition and air quality modeling in real time and retrospectively.

The software used is under research to initial operational capability. Coding standards are partially applied, moderate documentation is available, and PI affirms portability, numerical reproducibility, and of no security issues.

Metadata are under initial to full operational capability with full compliance to international standards. Sufficient collection level metadata and complete file level metadata are available.

User Documentation scores are varying from research grade to full operational capability. Comprehensive descriptions and journal publications are available on scientific methods. Limited descriptions of validation and operations concepts are available. However, there is no product user guide available.

Uncertainty Characterisation is either under research grade or initial operational capability. Standards are not applied, but comprehensive validation and inter-comparison are done. Limited information is available on uncertainty arising from systematic and random effects in the measurement. Automated quality monitoring is partially implemented

Public Access, Feedback, Update is under initial or full operational capability. Data and documentation are under institution's version control and archived and available to the public from Data Provider. Interim Climate Data Records are being produces using feedbacks through established mechanisms.

The dataset is widely used for research and citations are occurring. Use in decision making is occurring and benefits are emerging.

6.6 Reanalysis

This subsection discusses the ERA-Interim data set which is the only reanalysis data which is assessed in this report.

6.6.1 ERA-Interim

Name	ERA-Interim
Origin	ECMWF; Data.Services@ecmwf.int
Spatial Characteristics	Global, gridded
Temporal Characteristics	1979 – now; 6-hourly, with daily and monthly averages

Category	Software Readiness	Metadata	User Documentation	Uncertainty Characterisation	Public Access, Feedback, Update	Usage
Range	1 - 3	2 - 3	2 - 4	2 - 6	3 - 5	6

The ERA-Interim dataset comes from a comprehensive global reanalysis based on data assimilation, i.e. a process that blends model forecasts with a range of observational data,

taking into account their respective uncertainty characteristics. Originally created to support both numerical weather prediction and climate studies, ERA-Interim has subsequently been adopted in other applications including climate monitoring and earth-system science.

Coding standards are partially applied to the software used to generate the ERA-Interim dataset, the software documentation is substantially more than minimal with header and process description (comments) in the code but it lacks a README needed for the next maturity level, the PI affirms numerical reproducibility and portability under identical conditions, and the security problems are not evaluated. The software category is thus rated as research grade.

Metadata category is rated as initial operational grade. Standards are identified and applied at the file level, but there are only limited collection-level metadata available. The file-level metadata are often sufficient to use and understand the ERA-Interim data.

User Documentation is initial operational grade except for the lack of a description of the operations concept.

The ERA-Interim dataset is extensively validated for the principal parameters (e.g., temperature, humidity, precipitation). The data provider (ECMWF) has participated in multiple international data assessments and the results have been fed back to the product development cycle.

The ERA-Interim dataset, documentation, and source code are under institutional version control and archived. Data and documentation are available to the public. The data provider collects feedback from the scientific community and this is used for irregular updates to the dataset. This category falls into initial operational capability.

Improvements in all these categories are achievable given sufficient resourcing.

ERA-Interim data are used extensively in research and decision support systems.

7 SUPPORTING DATA SET USERS TO MAKE THE RIGHT CHOICE

Another tool developed in the project, the Application Performance Matrix (APM), attempts to evaluate the performance of an ECV CDR with respect to a specific application. The APM was added to the capacity assessment during the discussions about the SMM because it became clear that the SMM cannot answer the question on how good a data record is for a specific application. To be able to assess suitability of a data record with the APM, user requirements for each considered application are needed to compare the actual technical specifications and validation results to them.

APM is basically posing a set of typical questions which a user may ask when a data record is being searched for. Whereas questions towards the spatiotemporal coverage may be easy to answer from the technical specifications of a data record, questions towards results of uncertainty analysis are more difficult and a suggestion on the suitability of a data record for an application may need interaction between the application and data record experts. Key for

any suggestion for usage based on this is an understanding of the user requirements for an application. For instance GCOS provides useful requirements for its ECVs which can be used as guidelines for suggestions of data records for applications in climate system analysis. However, a detailed analysis of user requirements per application would be useful to enhance the usability of the APM in the future.

The basic principle of the APM is easy as it evaluates how well the data record's technical specifications and accessible validation results, which should be listed in a Product Specification Table (PST), match the user requirements for the application considered, which should be listed in a User Requirement Table (URT). The (PST) is a database that consists of all relevant details on the climate data record, such as the technical specifications (e.g. period covered, temporal and spatial resolution/sampling, temporal and spatial coverage, etc) and a summary of validation results (e.g. uncertainty arising from systematic and random effects, temporal stability, etc). The PST ideally is part of a climate data record inventory where all PSTs are coming together.

The User Requirement Table (URT) consists of the user query of requirements expressed in parameters that are provided in the PST. Essentially, the APM evaluation process refers to performing a query on the PST. When such a query is made on several data records simultaneously, the search query result that is returned comprises the APM and a suggestion on which data records are matching requirements for the application.

These concepts were discussed in the workshop and the participants endorsed further developments to the APM concept which will be pursued in the near future.

8 CONCLUSIONS

8.1 Completeness of the assessment

All major European satellite data producers participated in the assessment: ESA-CCI, CM-SAF, H-SAF, LSA-SAF, and MAAC. Some national or university developed dataset could be missing.

There is only one global reanalysis that is being produced in Europe (ERA Interim) and that has participated in the assessment. It could be possible to extent the assessment to regional reanalyses such as UERRA project in future. It is suggested to perform SMM assessments for such datasets in CCCS.

It has to be emphasised that such an assessment had been done for the first time for *in situ* datasets and the great response from the dataset developers are highly appreciated. Major European players such as Met Office Hadley Centre (HadObs), DWD (GPCC), KNMI (ECA&D), AWI (BSRN) participated in the assessment. Their fruitful discussions have contributed towards adapting the initially satellite oriented SSM to a more general one.

Although the assessment was done in a self-assessment mode, an audit-type assessment for some of the randomly picked datasets revealed that the data providers have been very honest in assigning SMM scores. This provides an overall large credibility of the self assessment

presented here. This is actually the reason why we did not perform 'audit' assessments for all datasets.

8.2 Value and potential usage

The following are some applications of the SMM:

- Implementation in the Quality Assurance and Enhancement (QA&E) pillar of C3S;

- The Maturity Matrix is already used in CEOS-CGMS WG Climate to assess status of data records in GCOS ECV inventory. This can be periodically repeated;

- SMM is used in H2020 EUSTACE project to assess the maturity of data set development;

- SMM and APM were included in QA concept of QA4ECV project;

- It is also planned to link the SMM results (e.g., graphical output shown in Figure 2) with CHARMe, i.e., assessment results become available through CHARMe nodes as annotations to data records.

APPENDIX A LIST OF ACRONYMS
APPENDIX B DATA SET DESCRIPTIONS
APPENDIX C DETAILED ASSESSMENT RESULTS PER DATA SET