World Meteorological Organization
World Climate Programme
Expert Meeting on Climate Data Modernization

KNMI, de Bilt, The Netherlands, 16-18 April 2018

EM-CDM/ Final report
28 May 2018

WMO EXPERT MEETING ON CLIMATE DATA MODERNISATION
16-18 APRIL 2018
FINAL REPORT
DRAFT

(Team member missing from picture: Cedric Bergeron (ECMWF))

DISCLAIMER

Regulation 43

Recommendations of working groups shall have no status within the Organization until they have been approved by the responsible constituent body. In the case of joint working groups, the recommendations must be concurred with by the presidents of the constituent bodies concerned before being submitted to the designated constituent body.

Regulation 44

In the case of a recommendation made by a working group between sessions of the responsible constituent body, either in a session of a working group or by correspondence, the president of the body may, as an exceptional measure, approve the recommendation on behalf of the constituent body when the matter is, in his opinion, urgent, and does not appear to imply new obligations for Members. He may then submit this recommendation for adoption by the Executive Council or to the President of the Organization for action in accordance with Regulation 9(5).

Chairperson, Publications Board
World Meteorological Organization (WMO)
7 bis, avenue de la Paix
P.O. Box No. 2300
CH-1211 Geneva 2, Switzerland

Tel.: +41 (0)22 730 84 03
Fax: +41 (0)22 730 80 40
E-mail: Publications@wmo.int

WMO EXPERT MEETING ON CLIMATE DATA MODERNIZATION

16-18 APRIL 2018

FINAL REPORT

EXECUTIVE SUMMARY

The WMO Expert Meeting on Climate Modernization was hosted at KNMI, De Bilt, Netherlands, from 16 to 18 October 2018.

The purpose of the meeting was a follow-up activity of the CCl led Inter-Programme Expert team on Climate Data Modernization (IPET-CDMP) for developing a High Quality Global Data Management Framework for Climate (HG-GDMFC). The meeting built on the outcomes of the international workshop on information management which was convened by CCl and CBS, 4-6 October 2017, and included a recommendation for a project plan for climate data access. A key intended outcome was the development of a catalogue of "trusted" high quality data sets for climate change monitoring, that will form the basis for calculating key climate indicators.

Key issues worked on during the meeting included:

- Identification of an initial limited number of climate related datasets by participating experts as “best trusted” datasets to become part of the WMO Climate Data Catalog of high quality data.

- Development of a WMO-Wide Stewardship Maturity Matrix for Climate Data (SMM-CD) based on US and European existing Maturity Matrix models. Climate Data Catalogue datasets will be assessed and scored by the SMM-CD with the goal to provide users with a level of trust in the management of the datasets.

- Development of a data discovery and access process for the Data Set Catalog through the WIS and the main Internet search engines with the goal to assist non-technical users to easily discover these high-quality data sets.

EXPERT MEETING ON CLIMATE DATA MODERNIZATION

1. Background

There are several challenges regarding climate data issues that hampers full implementation of high quality climate services. This includes among other things, much existing guidance on climate data management is out of date, due to rapid recent advances in technologies. Many countries' National Meteorological and Hydrological Services (NMHSs) report lack of capacity in (climate) data management and there is a significant gap in standardisation of terminology, processes, and practices due to lack of an appropriate regulatory framework. On the other hand, there is an opportunity to make better use of proliferation of new data sources and advances in technology that can support climate services. The WMO Commission for Climatology (CCl) led an inter-programme initiative on High Quality Global Data Management Framework for Climate (HQ-GDMFC) which was launched at the fifteenth session of CCl and aims at making use of an extended range of climate data types needed to support GFCS Climate Services Information System (CSIS) and harmonizing the definitions and processes that deal with data management in support of climate applications. The Framework helps also better harvesting the potential of other data sources to improve climate services, noting the potential to capture the benefits of modern information and remote sensing technologies, third party data and volunteer observations, especially in relation to standardized data management approaches associated with large data volumes. It provides the opportunity for addressing inconsistencies and gaps in the definitions and procedures relevant to climate data management. The HQ-GDMFC fosters collaboration with other programmes' data management initiatives, including CBS, JCOMM, GCOS and WCRP, and therefore it forms an essential element within wider WMO cross-commission data management initiatives, under the coordination of CBS. Furthermore, the Executive Council (EC) requested (Abridged final report EC-65) CCl to work closely with other Commissions and programmes, specifically with CBS on WIS, WIGOS and with the GFCS-relevant groups and teams, to move from a concept to the definition of a HQ-GDMFC.

- **Meeting Summary**

The WMO Expert Meeting on Climate Modernization was hosted at KNMI, De Bilt, Netherlands from 16 to 18 October 2018. **The objectives of the meeting were to:**

(1) Select an initial set of key climate datasets for the initial Global Data set catalog.

(2) Develop a WMO-Wide Stewardship Maturity Matrix for Climate Data (SMM-CD) that will allow the datasets from the Global Data Set Catalogue to be rated with a maturity score.

(3) Develop WIS and main search engine discovery Metadata requirements for users to easily discover high-quality data sets from the Catalog.

On Day 1 of the meeting two plenary sessions were held:

- **Session I - 'Setting the Stage'** was opened and chaired by Peter Siegmund of KNMI. Presentations were made summarizing the work leading up to the meeting including on the WMO Workshop on Data Information held 2-4 October 2017, the project on Climate Datasets, the need for a WMO-wide Maturity Matrix model and a way to improve the discovery of high quality data.

- **Session II - 'Available datasets and a Way-Forward'** was chaired by William Wright. Subject Matter Experts provided information on the status of on key climate indicator data such as the Essential Climate Variables (Temperature, Precipitation, Sea Level, Cryosphere), Climate Indices, Crowdsourced, Marine and Hydrological Data.

On Day 2, three breakout groups worked on the following:

- **Breakout group A: Data Catalog of High Quality Global Climate Data**

Participants: Peter SIEGMUND (KNMI) Lead, Dominique BEROD (WMO), Robert DUNN (Met Office), David GALLAHER (NSIDC), Lydia GATES(JCOMM), Christina LIEF (WMO) and Markus ZIESE (GPCC)

Goal: identify a list of high quality global data sets to be considered in the WMO catalog on global trusted climate data sets

Terms of Reference (TOR)

(a) Develop WMO Data Catalog of initial Global level authoritative, trusted data sets for informing on key climate indicators (Temperature, Precipitation, Sea Ice, Sea Level, Ice Sheets, Glaciers, Climate Indices, Crowdsourced, Marine and Hydrology) based on GCOS data set list.
(b) Review the format of the Data Catalog.
(c) Define trusted data sets from what came out of the October workshop.
(d) Develop a process for evaluating/qualifying the data sets to be part of the Catalog.
(e) Review the list of proposed data sets and see how they fit into the trusted definition and may be look if they have been scored in a Maturity Matrix already.
(f) Finalize the list of initial datasets for the Data Catalog.
(g) Decide how to handle future additions to the Data Catalog.

Outcomes

The group assembled inventory/catalog of relevant and (potentially) trusted data sets and metadata based on WMO/GCOS provided list of key ECV data sets and recommendation from Subject Matter Experts. List includes reanalysis, satellite, model data and merged datasets. (See Annex 4 for list of data sets). It proposed a mechanism for assigning data set/system maturity scores with the WMO Maturity Matrix as data sets will need to be evaluated for their maturity scores in the newly developed WMO Convergence Maturity Model. If they score at a satisfactory level, they can be included in the Data Catalog as trusted key climate indicators data sets. Scores rank from Level 1 (Not Managed) to Level 5 (Optimal) for 8 criteria (Documentation; Governance; Storage and Preservation; Quality Management; Data Access and Sharing; Discoverability; Change Management; Competencies) that were decided on in cooperation with the CBS TT-IM team. The total score of a data set will be based on the lowest scoring of the 8 categories. The group will need to establish the threshold score that will define a data sets as trusted. It is expected that this work will be further expanded and include additional global data set as well as regional and national datasets using the same mechanism.

- **Breakout group B: Develop WMO-Wide Stewardship Maturity Matrix for Climate Data (SSM-CD)**

Goal: Develop WMO Maturity Matrix based on existing models such as from NOAA/NCEI Data Stewardship Maturity Model (DSMM) and the ESA CoreClimax Product Maturity Matrix

Participants: Ge PENG (NCEI) Lead, Omar BADDOUR (WMO), Valentin AICH (WMO/GCOS) and William WRIGHT (BOM)

Terms of Reference (TOR)

(a) Review & discuss the maturity models (NOAA-NCEI, Europe, others) for assessing maturity of Earth Science data sets
(b) Review and potentially finalize the categories of the WMO-MMCM
(c) Discuss the way to represent the maturity ratings
(d) Start to define aspects of each categories
(e) Define an action plan for a final draft and for draft review

Outcomes:

The group worked on developing the WMO Stewardship Maturity Matrix for Climate Data 8SMM-CD) with the goal to provide a method of assigning a level of trust to datasets that was based on objective measures and would best help users identify data sets that can be trusted.It drafted an initial version of WMO SMM-CD based on scoring of level 1 to 5 of four categories (Accessibility, Usability, Quality Management and Data Management) with each category consisting of 2 to 4 sub-categories which are referred to as Aspects. For the Accessibility category there are 2 aspects: Discoverability and Accessibility. For the Usability category also two: Media/Format and Documentation. For the Quality Management category there are four aspects: Quality Assurance, Quality Control, Quality Assessment and Data Integrity. And for the Data Management category there are three aspects: Accountability, responsibility and competency. The team drafted descriptions and rationale of each of the categories and aspects for every level.

- **Breakout group C: Developing promotion mechanism of the Data Catalog in the WMO Information System (WIS) and the main search engines.**

Participants: Anna MILAN (NCEI) Lead, Kate ROBERTS (BOM), Thorsten BUSSELBERG (WIS), Cedric BERGERON (ECMWF)

Terms of Reference (TOR)

(a) Define the requirements to promote “trusted” WMO Datasets and their discovery in a prominent way through the WMO information System and Internet search engines
(b) Define the mechanism of promoting the “trusted” WMO Datasets
(c) Identify implementation steps

The group developed an approach to promote these trusted data sets in the WIS but also in the general-purpose search engines such as Google. The metadata describing the datasets will include reference to the level of trust in a way that the search algorithms could process. With the most trusted datasets appearing high in search rankings.

On Day 3, the participants met together again in plenary. The breakout groups presented their reports and recommendations on next steps and implementations were discussed.

Main outcomes of the meeting:

(1) Completion of the initial list of 16 potential high-quality data sets to be included in the Data Catalog. These data sets were taken from the GCOS ECV data set list and confirmed as being good potentials for trusted data set list of the Catalog by the team Subject Matter Experts. These data sets will now need to be run through the newly developed WMO Stewardship Maturity Matrix for Climate Data to evaluate their maturity with a score from 1-5. Only high scoring data sets above 3 will be included in the catalog.
(2) Development of first draft of the WMO SMM-CD was completed by breakout group B. The draft is presently going through internal review till June 22, 2018 and will then go through a period of external review which involves a presentation at ESIP 2018 in July.
(3) Identification of an approach based on a key attribute to be described in the Discovery Metadata that will help in promoting Data Catalog data sets in the WIS and on the main search engines such as Google. Next step is to compile list of associated WIS compliant metadata and other authoritative metadata or data home page for each data asset once they have been chosen for the Data Catalog.

Recommendations

1) Experts to test the first version of the WMO/SMM-CD and report back on their findings. This will help in further refining the SMM-CD
2) Present SMM-CD for external review at The Earth Science Information Partners (ESIP) 2018, "Data for our changing Earth: Realizing the Socioeconomic Value of Earth Science Data", July 17-20, 2018, Tucson, Arizona. (Lead Ge Peng, William Wright, Christina Lief)
3) Present the work being done at the AGU 2018 Fall Meeting, December 10-14, 2018, Washington, DC. (lead: David Gallaher)

The way forward for finalisation of SMM-CD:

(1) Collect feedback from experts based on testing SMM-CD on their climate datasets developed by their institutions
(2) Collect feedback from ESIP and AGU meetings
(3) Collect Feedback from CCl Management Group
(4) Develop guidelines on SMM-CD to be included as an appendix in the WMO Manual on High Quality Global Data Management Framework for Climate (HQ-GDMFC)

5 Meeting close

- The meeting closed at 1630 on 18 April 2018.

A drafting team comprised of William Wright, Christina Lief, Peter Siegmund and Omar Baddour continued work on the first version of the draft Manual on HQ-GDMFC. The team agreed to further work on the draft to be submitted for consideration by the WMO Executive Council which will meet at its seventeenth session (EC-70) in Geneva 20-29 June 2018. At the writing of this report, the second draft was completed on 15 June 2018 and made available for consideration by EC-70.

_____

ANNEX 1: AGENDA

Day 1 – Monday April 16, 2018

1	**Plenary - Session I: Setting the Stage**	P Siegmund (Netherlands)
1.1	Outcomes of International Workshop	O Baddour (WMO)
1.2	Project on Climate Datasets	W Wright (Australia)
1.3	Ad hoc Working Group, TOR and Thematic Focus	C Lief (France)
1.4	Maturity Models and need for a WMO Maturity Matrix Convergence Model	P Siegmund (Netherlands)
1.5	Need for Prominent Discovery and the use of WIS Infrastructure	K Roberts (Australia)
1.6	What Metadata are Needed, and Citation Issues	K Roberts (Australia)/T Busselberg (Germany)
2	**Plenary - Session II: Available Datasets and a Way-Forward**	W Wright (Australia)
2.1	Data Issues	R Dunn (UK)
2.2	Status of ECV and Process of Evaluation of ECV Datasets (GCOS Perspective)	V Aich (WMO)
2.3	Climate Indices	P Siegmund (Netherlands)
2.4	Temperature Data	R Dunn (UK)
2.5	Precipitation Data	M Ziese (Germany)
2.6	Sea Level Data	A Cazenave (France) (C Lief (France) presented summary)
2.7	Cryosphere Data	D Gallaher (US)
2.8	Hydrology Data, Status and Quality Process	D Berod (WMO)
2.9	Marine Data	L Gates (Germany)
2.10	Copernicus Data	C Bergeron (UK)
2.11	Presentation on Maturity Models	G Peng (US)
2.12	Current Status of the Maturity Matric Convergence Model	W Wright (Australia)
2.13	A way forward for prominent discovery for the WMO Data Catalog of High-quality Climate Data	A Milan (US)

DAY 2 - **Tuesday April 17, 2018**

3 Session III – Breakout Groups A, B, C Meetings

DAY 3 - **Wednesday April 18, 2018**

4 Session-IV- Session chair: Christina Lief/Rapporteur: Omar Baddour

4.1 Reports and Finalization

5 Session-V - Session chair: Christina Lief / Rapporteur: Omar Baddour

5.1 Recommendations on next steps and implementations

_____

ANNEX 2: PARTICIPANTS

Affiliation	Participant
Australia	Kate ROBERTS
Australia	William WRIGHT
Germany	Lydia GATES
Germany	Thorsten BUSSELBERG
Germany	Markus ZIESE
Netherlands	Peter SIEGMUND
UK	Cedric BERGERON
UK	Robert DUNN
US	Ge PENG
US	Anna MILAN
US	David GALLAHER
WMO	Omar BADDOUR
WMO	Valentin AICH
WMO	Dominique BEROD
WMO	Christina LIEF

_____

ANNEX 3: SUMMARIES OF PRESENTATIONS

Working documents and presentations supplied for the meeting are available from http://wis.wmo.int/page=WWIM.

1 Session I: Setting the Stage (Session chair: Peter Siegmund (KNMI))

1.1 Outcomes of International Workshop (Omar Baddour (WMO))

- The WMO Workshop on Information Management met at WMO headquarters in Geneva, October 2-4 2017.

- 35 participants attended from world data centers, IPCC, WRCP, GCOS, UN-BIG Data, Hydrology, Marine, Agriculture, Climatology and Weather.

- The key objectives of the workshop were to:
  - Identify key climate data sources that should be promoted as reference sources and outline a project to enhance their visibility in the WMO Information System (WIS)
  - Simplify user access to these trusted data sets.
  - Provide guidance on the development of a WMO reference on information management
- Outcome was a project plan for climate data access in context of the WIS which proposed to develop:
  - a data catalog of authoritative, trusted data sets for informing on key climate indicators for global policy users of climate change information, as well as for regional and national level data sets.
  - A mechanism to assign a maturity score to these data sets.
  - Provide a discovery mechanism in the WIS and the main search engines for non-technical users to discover these trusted data sets.

1.2 Project on climate data-sets (William Wright (BOM, Australia)

- The requirement for high quality data set must include that the coverage of the data is as long as possible, and as far as possible complete, is available digitally, is properly quality controlled with confidence assessments, free of spurious discontinuities and trends, has adequate metadata, has complete lineage documentation, and has proper management of data and metadata. The quantitative measure of data management can be assessed by a maturity matrix approach.

- The HQ-GDMFC will identify high quality datasets and provide guidelines and best practices to improve climate data, such as harmonizing the Climatological Day definition, and management through WMO Technical Regulations. An internationally agreed Catalog providing a living list of high quality datasets will be developed and maintained under the auspices of WMO. This Data Catalog will be made available to users using the WIS and datasets will be easily discoverable and accessible though optimized search in main search engines.

- GCOS identifies high quality datasets using methods that include peer-review processes. These sets must provide evidence of conforming with the GCOS Monitoring Principles on the observing side. The purpose of the HQ-GDMFC Climate Datasets project is to ensure that the datasets are also well and sustainably managed.

1.3 Ad hoc Working Group, Thematic Focus and TOR (Christina Lief (WMO))

- The Ad hoc Expert Working Group was formed as a follow up of the WMO International Workshop on Information Management, held at WMO in October 2017, and is focusing on the group A recommendations. The Working group is composed of some of the members who attended the workshop and some invited experts from the major data centers, WMO Commissions, Programmes and Initiatives. The Workshop Participants also recommended to jointly develop with the WMO Commission for Basic Systems (CBS) Tasks-Team on Information Management (TT-IM), a WMO Maturity Matrix Convergence Model based on the NOAA/NCEI Data Stewardship Maturity Matrix (DSMM) and the ESA CORE-CLIMAX System Maturity Matrix (SMM).

- Working group experts provided the status on specific data themes/variables presenting the problems/challenges with data completeness, gaps, and other issues. This helped the group identify the data that are available for each of the themes/variables and best determine what dataset meet the high-quality label to be included in the WMO HQ Data Catalog. The data management maturity of these data sets will be established using a maturity matrix score.

- Terms of Reference (TOR):

    - Development of a WMO Data Catalog of High Quality data (ECV: Temperature; Precipitation; Sea Ice; Sea Level; Ice sheets; Glaciers) Climate Indices, Crowdsourced, and Hydrology datasets; (Breakout Group A lead by Peter Siegmund)

    - Joint development with the CBS TT-IM of a WMO Maturity Matrix Convergence Model; (Breakout Group B lead by Ge Peng)

    - Promotion mechanism of the Data Catalog in the WMO Information System (WIS) and the main search engines. (Breakout Group C lead by Anna Milan)

1.4 Maturity Models and need for WMO Maturity Matrix Convergence Model (Peter Siegmund (KNMI, The Netherlands)

- In the final report of the First Meeting of the Task Team on Information Management (TT-IM), 5-6 October 2017, participants agreed that a maturity model approach would be appropriate for developing, managing and monitoring information management practices within WMO.
- A proposed WMO Maturity Matrix would use 8 categories (documentation, governance, storage and preservation, quality management, access and sharing, discovery, change management, competencies) that would each be sub-divided into aspects. The matrix would be developed considering behaviors for each of the aspects that would be expected in practice at each of 5 maturity levels (1=ad hoc; 2=minimal; 3=intermediate; 4=advanced; 5=optimal).
- Development of the WMO Maturity Matrix would involve defining the aspects for each of the categories, as well as the behaviors for the maturity levels. A maturity score formula would need to be defined.
- Only data with a high enough maturity could be labeled as trusted and included in the high-quality data catalog of climate data.

1.5 Need for Prominent Discovery and the use of WIS infrastructure (Kate Roberts (BOM, Australia)/Thorsten Busselberg (DWD, Germany)

- WIS Centers are global information center that provide data, metadata, distribute data to GISCs, and among others, ensures area of responsibility has effective telecommunications.

- The WIS Discovery metadata provides what a data set contains, who is responsible for the data set and how to access that data.

1.6 What Metadata are Needed, and Citation Issues (Kate Roberts (BOM, The Netherlands)

- Metadata are required for high quality datasets to be integrated into current WIS (and WIS2.0)
- All metadata for inclusion in the current WIS must be WMCP v1.2 or 1.3 compliant: http://wis.wmo.int/2013/metadata/version_1-3-0/
- WMCP records must be ISO19115/19139:2007 compliant
- GISCS are intended to handle ISO19115-compliant metadata
- Ideal metadata provides an understanding of the data content
- In the ISO19115:2006 metadata standard, Data Lineage and Data Quality assessments both sit under “Data Quality”.
- Any known quality issue should be state or referenced.

2 Session II: Available datasets and a way-forward (Session chair: William Wright (BOM, Australia)

- High quality climate datasets are to provide authoritative climate assessments, ensure the integrity of climate data, to develop climate change knowledge to inform adaptation strategies at global, regional and national levels, and to support climate modelling and climate change projections.

- Specific requirements for quality climate datasets are: that they should be as long a time series as possible, accessible in digital forms, as far as possible complete, properly quality controlled, have confidence assessments, as far as possible free of spurious discontinuities and trends (requiring attention to observing practices), and accompanied by adequate metadata and complete documentation of lineage.

- The Commission for Climatology is planning a High Quality Global Data Management Framework for Climate (HQ-GDMFC) that includes a manual on climate data management. The need for such a framework was driven by the following factors. Much existing guidance on climate data management was out of date, due, for example, to rapid recent advances in technologies. The NMHS of many countries reported lack of capacity in (climate) data management. There was a lack of a regulatory framework, standardisation of terminology, processes and policies. The framework was also needed to enable better use of new data sources and advances in technology.

- Initial findings of the team responsible for creating the HQ-GDMFC (IPET-CDMP – the Inter-Programme Expert Team on the Climate Data Modernization Programme) were that there was scope for collaborative arrangements for data management between different scales and domains (international datacentres, regional climate centres, National Meteorological and Hydrological Services (NMHS) and the WIS. The team also concluded that there was a need for new standards on management of changes to observing systems and processes, process documentation, and authentication and publication of climate data sets.

2.1 Data Issues (Robert Dunn (Met Office Hadley Center, UK)

- Importance of making transparent assessment of data qualities for the users so they can make informed decision on the use of the data.

- There are processing differences depending on the various data types.

- Uncertainties need to be address for measurements, processing (parametric and structural), ensembles of opportunity and implementation.

- There is a need for procedural aspects such as version control, quality assurance, open code, timely access and updates, maintained documentation, communications via peer review and user engagement and collaboration and knowledge sharing.

- There are issues with accessibility such as using standardized data formats, synchronicity between mirrored versions, up to date documentation, standards, licenses and controlled vocabulary.

2.2 Status of ECV and process of evaluation of ECV datasets (GCOS perspective) (Valentin Aich (WMO))

- GCOS specifies 54 Essential Climate Variables (ECV) that are key for sustainable climate observations.

- Data sources need to be reviewed against certain criteria before a data set can be approved as part of the GCOS ECVs such as have global coverage, free and open access, quality content documentation, include metadata and be considered by the GCOS Science Panels. Also, each ECV has one or more ECV stewards as point of contact who consult with their communities, keep data lists up to date and review new suggestions.

- GCOS data sets are available on the GCOS web site and are trusted, accessible and sustainable.

- GCOS suggest using their list of ECVs for the base of the WMO Data Catalog.

2.3 Climate Indices (Climate Indices (Peter Siegmund (KNMI, The Netherlands))

- As defined by the WMO CCl, Clivar and JCOMM Expert Team on Climate Change Detection and Indices (ETCCDI) there are 27 standardized indices for the GCOS ECV. They are near the surface over land, internationally coordinated and used for both observations and models.
- The global data sets for the climate extreme indices are: HadEX (doi:10.1029/2005JD006290), HadEX2 (doi:10.1002/jgrd.50150), HadGHCND and GHCNDEX
- The WCRP grand challenge on weather and climate extremes is evaluating whether the existing observations are sufficient to underpin the assessment of extremes. So far it has found that there are critical gaps in the amount, quality, consistency and availability of observation especially with respect to extremes. It also found that the current climate extreme data sets are inadequate to properly assess climate variability and change, needed for detection, attribution and model evaluation. This is due to spatial and temporal data limitations, difference in how extremes are defined, spatial representation of point-based measurements, scaling issues between observation (in-situ and remote sensed) and models and uncertainties in variables estimates from satellite retrievals.

2.4 Temperature Data (Robert Dunn (Met Office Hadley Center, UK))

- Temperature data can be based on:
  - Surface: daily or monthly station data (in situ), reanalyses data or blended land and surface marine data, sea surface temperature, marine air temperature, ice shelf

    - Non-surface: ‘Surface T’ temperature, radiosonde data for upper air and sub-surface land such as permafrost from boreholes, and sub-surface marine.
    - Satellites measure radiance and brightness, and upper atmosphere temperature.
- Scientist need to define what is meant by temperature and users need to let know what they want from temperature data. Each data set has features to be aware of depending on use-case. Information on comparing data sets, describing uncertainties are helpful and education and informed decision making is vital.

2.5 Precipitation Data (Markus Ziese (DWD, Germany))

- There are many challenges to produce a precipitation data set such as that some countries have very restrictive data policies, some historical data sets are not in digital format, not accessible and not in a user-friendly format. Some observation networks have maintenance failures which causes gaps in time series. Other problems are with the quality of the data caused by orographic effects, real extremes over thresholds and time stamps of aggregation interval as well as with the interpolation and grid creation of the data set.
- Some examples of precipitation data sets are the GHCN-Daily (station data), CRU (in situ, station based, gridded) and the GPCC (in situ, gridded).
- There are many precipitation data sets but they differ in data source, quality control, gridding scheme, temporal and spatial resolution, update cycle and references (DOI, URL).

2.6 Seal Level Data (Anny Cazenave (LEGO, France)) (Anny was unable to attend due to a schedule conflict. A couple of slides were presented on the status of seal level data)

- Sea Level is an important indicator of global climate change. It is an integrated response of ocean heat content, land ice, land water storage to external forcing factors (including anthropogenic forcing) and internal climate variability.
- From 1/1993 to 3/2018, the mean rate of sea level rise has been 3.0 +/- 0.3 mm/yr with an acceleration of 0.09 mm/yr2
- The global mean sea level rise is mostly the result of anthropogenic global warming.
- Regional seal level changes still dominated by the internal climate variability
- Since 2017, there have been operational generation of sea level products by the Copernicus Climate Change Service (C3S) (Europe) but continuing upstream research is strongly needed to provide improved altimetry-based sea level data sets.
- The Global Mean Seal Level Budget is the Observed Global Mean Sea Level = Ocean Thermal Expansion + Ocean Mass.
- Motivations for regularly assessing the seal level budget is to foster a better understanding process at work and to follow the temporal changes of individual components as well as the place bounds on missing or poorly know contributions.
- The challenges are a need for more research to fully reach the GCOS requirements; a need to ensure continuity of the high-quality sea level record for climate research and include new satellites in orbit; a need for continuing the R&D activities for feed operational production of sea level by C3S and develop multi-mission coastal sea level products globally.

- Conventional satellite altimetry does not work well within 20km of the coast due to perturbation from land (shelf currents, waves, river runoff in deltas and estuaries, etc.)

2.7 Cryosphere Data (David Gallaher (NSIDC, US))

- Cryosphere data include ice sheets, sea ice, permafrost, snow and glacier data.
- There is a risk of continuing satellite coverage for cryosphere data. High gap potentials are for sea ice extent (due to loss of F-20 and AMSR-2 age), sea ice concentration (due to loss of F-20 and AMSR-2 age), permafrost (not known replacement of satellite), ice sheet elevation (IceSat-1 failed in 2009), ice sheet mass (Grace-1 failed in 2017).
- Data gaps threaten overall quality of the data record when satellite failures prevent cross correlation of satellite sensors.
- Specific data sets impacted by gaps are IceSat and Grace.
- Data sets with a high risk of gaps are DMSP/SSM/I passive microwave and Cryosat-2
- There is an overall risk with decreased funding for Earth observations in the US and possible cancellation of satellite missions and deactivation of operational satellites.
- Commercial remote sensing may be able to fill gaps but how does science access these data.

2.8 Hydrology Data, Status and Quality (Dominique Berod (WMO))

- The structure of Hydrology at WNO fall under two branches: The Climate Prediction and Adaptation Branch and the Hydrology and Water Resources Branch which overseas the Hydrological Forecast and Water Division, The Basic Systems in Hydrology Division and the Capacity Building in Hydrology and Water Resources Management Office.
- WMO has a regulatory framework for hydrological data that provide in the Technical regulations, Manuals and Guides.
- WHYCOS is the World Hydrological Cycle Observing System with world-wide regional systems such as the Pacific-HYCOS. Some of these are operational while some are not.
- WMO also hosts a hydrological observing system portal (WHOS) that represents the hydrological component of WIGOS and provides access to hydrological data.
- The WMO HydroHub is being developed as the WMO Global Hydrological support facility and will integrate WMO hydrological components and data access including the WHYCOS and WHOS.
- WMO HydroSOS is also being developed and will provide the first global hydrological monitoring and reporting system for assessing surface and groundwater status and warning on impending floods and droughts.

2.9 Marine Data (Lydia Gates (JCOMM, Germany))

- JCOMM provides international cooperation and has strategic goals of gathering observations, data management, services and forecasting systems, developing standards and best practices and coordination and collaboration and strategy.
- A well-formed Marine Climate Data System (MCDS) enables JCOMM to establish a much-needed operational international path for acquisition and for sharing relevant public access to delayed-mode global marine meteorological and oceanographic data leveraging existing tested resources.

- The most extensive freely available global surface marine data archive is ICOADS with over 455 million individual marine reports, observations and metadata reported from a given ship, buoy, coastal platform or oceanographic instrument. It provides a common access point and is version controlled.
- The CMOC World Ocean Database is the world's largest publicly available oceanographic profile database and pulls all relevant data stream together. It also provides climate products such as the WOD Ocean Heat Content and contributes to assessment and ocean knowledge. New limited number of CMOC data management areas were being proposed such as CMOC-China, -World Ocean Database, and -ICOADS.
- There are other systems that provide data access to marine data such as WIGOS and JCOMM's Data Discovery Interface.

2.10 Copernicus Data (Cedric Bergeron (Copernicus/ECMWF, UK))

- Copernicus, previously known as the Global Monitoring for Environment and Security (GMES) is the European Programme for the establishment of a European capacity for Earth Observation.
- ECMWF has been entrusted with the implementation of the Copernicus Climate Change Service (C3S) which provides information to increase the knowledge base to support adaptation and mitigation policies.
- Copernicus hosts the Climate Data Store (CDS) which contains observations, global and regional climate reanalyzes, global and regional climate projections and seasonal forecasts. It is designed as a distributed system, providing improved access to existing datasets through a unified web interface. The CDS also provides a toolbox that allows users to develop applications and make use of the CDS content.
- The C3S provides seasonal forecasts using data from ECMWF, the Met Office, Meteo France, DWD and the CMCC and makes them publicly available.
- Global and regional climate projections are also produced using state of the art climate models such as CMIP and CORDEX.
- Url: climate.copernicus.eu

2.11 Presentation of Maturity Models (Ge Peng (NCEI, US))

- A Maturity Matrix is defined as a maturity assessment model with desired evolution in discrete, progressive stages from a more ad hoc approach to a more managed process
- WMO is dedicated to ensuring the highest possible quality of all its endorsed data, products and service and would benefit from a data maturity assessment model to provide users with a way to assess data sets as trusted using a maturity score.
- Trusted data mean it is drawn from carefully selected sources, transformed in accordance with data's intended use and delivered in formation and time frames that are appropriate to specific users.
- Over the past few years, there have been several efforts to measure the maturity of data such as the Gap Analysis for Integrated Atmospheric ECV CLimate Monitoring (GAIA-CLIM) measurement system maturity matrix (Thorne et al 2015), NOAA's Center for Satellite Applications and Research (STAR) data product algorithm maturity matrix (Reed 2013; Zhou et al. 2016), and the European COordinating Earth observation data validation for RE-analysis for CLIMAte ServiceS (CORE-CLIMAX) production system maturity matrix

(EUMETSAT 2013), and the NCEI/CICS-NC scientific data stewardship maturity matrix (Peng et al. 2015).

2.12 Current Status of the Maturity Matrix Convergence Model (William Wright (BOM, Australia))

- The climate program is modernising its data and information management practices, in the context of a broader WMO initiative and will be working closely with the information management component of WIS, i.e. the CBS Task Team-Information Management (TT-IM), whose goal is to provide standards, best practice and guidance for Members so that information management within WMO Programmes is consistent and reliable.
- TT-IM is using a maturity model approach to specify WIS information management.
- A maturity model is a tool for monitoring and improving aspects of organizational performance and is often presented as a two-dimensional matrix. The rows identify the various requirements and aspects while the columns describe typical behaviours representing increasing capability against each aspect. Typically, a maturity model uses five capability levels from Level 1 ‘Ad hoc nor managed’ to level 5 ‘Optimal level 4+ measured, controlled, audit’.
- Benefits of a maturity matrix approach are that it divides information management activities into a manageable set of aspects that can be addressed by Members, allows Members to assess their information management practices to identify those aspects that would benefit most from process improvement activities, provides a way of measuring progress towards improving information management capability in support of WMO Programmes, allows Members to identify an appropriate level of process maturity that should be used for the information they are managing.
- The eight identified practices in high quality data and information management are good documentation, good governance, effective storage and preservation, effective quality assurance processes, accessibility and sharing mechanisms, discoverability, accountable change management processes and provenance, adequate competencies and training (with new WIS competencies).
- Process to date has been the following: WMO Workshop on IM defined the requirements; the following TT-IM meeting defined the Aspects and sub-aspects; Members of the TT-IM were asked to describe in more detail the aspects; One row of the Maturity Matrix was filled out in terms of Aspects and behaviours (for Governance). Presented at CBS TECO.
- Proposed next steps are that the behaviours corresponding to each maturity level for each Aspect will be identified and described by the TT-IM with the aim to populate this by October 2018, Guidance will then be developed to support members in developing their competency in each aspect, Regulations, describing the high-level requirements, will be drafted post-EC and put through a formal consultation process late in 2018, Goal is to have these ready for approval by Cg-18, initial draft guidance will be placed on the Web during 2018 and early 2019, and feedback and road-testing is encouraged; Further development of the standards, recommended practices and guidance will continue and be approved through standard procedures.

2.13 A way forward for prominent discovery for the WMO Data Catalog of High-quality Climate Data (Anna Milan (NCEI, US))

- The NOAA/NCEI OneStop discovery and access project focuses on users being able to identify key data sets through an intuitive interface using the cloud and applying maturity scales. Search is being driven using standardized metadata content. OneStop also provide guidance and training materials.

- Discovery should promote the most trusted datasets and would include sufficient information in the metadata records to allow search algorithms to promote the datasets in their results. With such information it would then be possible to use searches of standard catalogues (such as the WIS catalogue) to prepare listings of the most trusted datasets for application areas. It was recommended that those listings should be linked from the home pages of GISCs (Global Information System Centres). Although based on the ISO 19115 geographic metadata standard, the information in the listings should be presented in a user-friendly way. In addition to listing the trusted datasets in searches, the datasets should also appear in the results of the main search engines such as Google.

ANNEX 4: INITIAL PROPOSED DATA CATALOG OF HIQH QUALITY GLOBAL CLIMATE DATA

Climate Indicator: TEMPERATURE (ECV)			
Data set name	**Scale: Daily, Monthly, Annual, Other**	**Citation**	**Location: Web link to the dataset**
NOAAGlobalTemp (gridded)	Monthly	Smith, T.M., R.W. Reynolds, T.C. Peterson, and J. Lawrimore, 2008: Improvements to NOAA's historical merged land–ocean surface temperatures analysis (1880–2006); Journal of Climate, 21, 2283–2296, doi:10.1175/2007JCLI2100.1	**Global Gridded 5° x 5° Data** Metadata
HadCRUT.4.6.0.0	Monthly	Morice, C. P., J. J. Kennedy, N. A. Rayner, and P. D. Jones (2012), Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The HadCRUT4 dataset, J. Geophys. Res., 117, D08101, doi:10.1029/2011JD017187	**Data download** Metadata
NASA Surface Temperature Analysis (GISTEMP)	Monthly	Hansen, J., R., Ruedy, M. Sato, and K. Lo 2010: Global temperature change, rev. Geophys., 48 RG4004, doi:10.1029/2010RG000345	**Data download** Metadata

Climate Indicator: PRECIPITATION (ECV)			
Data set name	**Scale: Daily, Monthly, Annual, Other**	**Citation**	**Location: Web link to the dataset**
GPCC Full Data Reanalysis	Monthly	Schneider, U., Finger, P., Meyer-Christoffer, A., Rustemeier, E., Ziese,	**Data**

version 7.0 at 0.5		M. and Becker, A.: "Evaluating the Hydrological Cycle over Land Using the Newly-Corrected Precipitation Climatology from the Global Precipitation Climatology Centre (GPCC)", Atmosphere 2017, 8(3), 52, doi:10.3390/atmos8030052	**Download** Metadata
Climate Indicator: SEA LEVEL (ECV)			
Data set name	**Scale: Daily, Monthly, Annual, Other**	**Citation**	**Location: Web link to the dataset**
GLOSS Global Mean Sea Level	Monthly & Annual	Simon J. Holgate, Andrew Matthews, Philip L. Woodworth, Lesley J. Rickards, Mark E. Tamisiea, Elizabeth Bradshaw, Peter R. Foden, Kathleen M. Gordon, Svetlana Jevrejeva, and Jeff Pugh (2013) New Data Systems and Products at the Permanent Service for Mean Sea Level. Journal of Coastal Research: Volume 29, Issue 3: pp. 493 – 504. Doi:10.2112/JCOASTRES-D-12-00175.1	**Data download** Metadata

Climate Indicator: SEA ICE (ECV)			
Data set name	**Scale: Daily, Monthly, Annual, Other**	**Citation**	**Location: Web link to the dataset**
Sea Ice Index, Version 3	Daily	Fetterer, F., K. Knowles, W. Meier, M. Savoie, and A. K. Windnagel. 2017, updated daily. Sea Ice Index, Version 3. Boulder, Colorado USA. NSIDC: National Snow and Ice Data Center. doi: http://dx.doi.org/10.7265/N5K072F8	**Data download** Metadata
Climate Indicator: ICE SHEETS (ECV)			
Data set name	**Scale: Daily, Monthly, Annual, Other**	**Citation**	**Location: Web link to the dataset**
GLAS/ICESat 500 m Laser Altimetry Digital Elevation Model of Antarctica, Version 1	Operational periods	John P. DiMarzio 2007. GLAS/ICESat 500 m Laser Altimetry Digital Elevation Model of Antarctica, Version 1. [Indicate subset used]. Boulder, Colorado USA. NSIDC: National Snow and Ice Data Center.	**Data download** Metadata

(satellite)		doi:10.5067/K2IMI0L24BRJ	
GLAS/ICESat 1 km Laser Altimetry Digital Elevation Model of Greenland, Version 1 (satellite)	Operational periods	John P. DiMarzio 2007. GLAS/ICESat 1 km Laser Altimetry Digital Elevation Model of Greenland, Version 1. [Indicate subset used]. Boulder, Colorado USA. NSIDC: National Snow and Ice Data Center. doi:10.5067/FYMKT3GJE0TM	**Data download** Metadata
Antarctica Mass Variability Time Series Version 1 from JPL GRACE Mascon CRI Filtered (Satellite)	Operational periods	D. N. Wiese, D.-N. Yuan, C. Boening, F. W. Landerer, M. M. Watkins. 2017. Antarctica Mass Variability Time Series Version 1 from JPL GRACE Mascon CRI Filtered. Ver. 1. PO.DAAC, CA, USA. Dataset accessed [YYYY-MM-DD] at http://dx.doi.org/10.5067/TEMSC-ANTS1.	Data download Documentation
Greenland Mass Variability Time Series Version 1 from JPL GRACE Mascon CRI Filtered	Operational periods	D. N. Wiese, D.-N. Yuan, C. Boening, F. W. Landerer, M. M. Watkins. 2017. Greenland Mass Variability Time Series Version 1 from JPL GRACE Mascon CRI Filtered. Ver. 1. PO.DAAC, CA, USA. Dataset accessed [YYYY-MM-DD] at http://dx.doi.org/10.5067/TEMSC-GRTS1.	Data download Documentation
Climate Indicator: GLACIERS (ECV)			
Data set name	**Scale: Daily, Monthly, Annual, Other**	**Citation**	**Location: Web link to the dataset**
World Glacier Database, version 1	1 January 1900 to 31 December 2003 (updated 2012)	WGMS, and National Snow and Ice Data Center (comps.). 1999, updated 2012. World Glacier Inventory, Version 1. [Indicate subset used]. Boulder, Colorado USA. NSIDC: National Snow and Ice Data Center. doi:10.7265/N5/NSIDC-WGI-2012-02	**Data download** Metadata
Climate Indicator: CLIMATE INDICES			
Data set name	**Scale: Daily, Monthly, Annual, Other**	**Citation**	**Location: Web link to the dataset**
HADEX2: GRIDDED TEMPERATURE AND	Annual	Donat, M. G., et al. (2013), Updated analyses of temperature and precipitation extreme indices since the beginning of the twentieth century: The HadEX2 dataset, J.	**Data download** Metadata

PRECIPITATION CLIMATE EXTREMES INDICES (CLIMDEX DATA)		Geophys. Res. Atmos., 118, 2098–2118 DOI:10.1002/jgrd.50150	
Climate Indicator: CROWDSOURCED DATA			
Data set name	**Scale: Daily, Monthly, Annual, Other**	**Citation**	**Location: Web link to the dataset**
Community Collaborative Rain, Hail and Snow Network (CoCoRaHS) (observations)	Daily	Reges, H.W., N. Doesken, J. Turner, N. Newman, A. Bergantino, and Z. Schwalbe, 2016: COCORAHS: The evolution and accomplishments of a volunteer rain gauge network. Bull. Amer. Meteor. Soc., 97, 1831-1846 doi:10.1175/BAMS-D-14-00213.1	**Data download** Metadata
Climate Indicator: HYDROLOGY			
Data set name	**Scale: Daily, Monthly, Annual, Other**	**Citation**	**Location: Web link to the dataset**
Global Runoff Data Center (GRDC) Database	Monthly	Fekete, B. M., C. J. Vörösmarty, and W. Grabs, High-resolution fields of global runoff combining observed river discharge and simulated water balances, Global Biogeochem. Cycles, 16(3), 2002, doi:10.1029/1999GB001254	**Data download** **Metadata**
Climate Indicator: MARINE			
Data set name	**Scale: Daily, Monthly, Annual, Other**	**Citation**	**Location: Web link to the dataset download and metadata**
ICOADS 3.0 (gridded data)	Monthly	Freeman, E., S.D. Woodruff, S.J. Worley, S.J. Lubker, E.C. Kent, W.E. Angel, D.I . Berry, P. Brohan, R. Eastman, L. Gates, W. Gloeden, Z. Ji, J. Lawrimore, N.A. Rayner, G. Rosenhagen, and S.R. Smith, 2017: ICOADS Release 3.0: A major update to the historical marine climate record. *Int. J. Climatol.* (CLIMAR-IV Special Issue), 37, 2211-2237 (doi:10.1002/joc.4775).	Data download Metadata

World Ocean Database 2013 (WOD13) (observations)	Multiple	Boyer, T.P., J. I. Antonov, O. K. Baranova, C. Coleman, H. E. Garcia, A. Grodsky, D. R. Johnson, R. A. Locarnini, A. V. Mishonov, T.D. O'Brien, C.R. Paver, J.R. Reagan, D. Seidov, I. V. Smolyar, and M. M. Zweng, 2013: World Ocean Database 2013, NOAA Atlas NESDIS 72, S. Levitus, Ed., A. Mishonov, Technical Ed.; Silver Spring, MD, 209 pp.,(doi.org/10.7289/V5NZ85MT)	Data download Metadata

ANNEX 5: BIBLIOGRAPHY

BIBLIOGRAPHY

Intergovernmental Oceanographic Commission, 2017. *IOC Strategic Plan for Oceanographic Data and Information Management(2017-2021).* Paris: Intergovernmental Oceanographic Commission.

International Civil Aviation Organization, 2014. *Manual on the Digital Exchange of Aeronautical Meteorological Information.* Montreal: International Civil Aviation Organization.

International Civil Aviation Organization, 2015. *Meteorological Service for International Air Navigation. Annex 3 to the Convention on International Air Navigation..* Nineteenth ed. Montreal: International Civil Aviation Organization.

International Civil Aviation Organization, 2016. *Global Air Navigation Plan.* 5 ed. Montreal: International Civil Aviation Organization.

Mosely, P., 2001. *Exchange of hydrological data and products, technical reports in hydrology and water resources.* Geneva: World Meteorological Organization.

NCEI, 2015. *The scientific data stewardship maturity assessment model template,* s.l.: NCEI.

Open Geospatial Consortium and the World Wide Web Consortium, 2017. *Spatial Data on the Web Best Practices.* s.l.:World Wide Web Consortium (W3C).

Peng, G. et al., 2015. A unified framework for measuring stewardship practices applied to digital environmental datasets.. *Data Sciene Journal,* Volume 13, p. 23.

World Meteorological Organization and Intergovernmental Oceanographic Commission, 2017. *Draft Joint WMO and IOC Strategy for Marine Meteorological and Oceanographic Data Management (2018-2021).* s.l.:World Meteorological Organization.

World Meteorological Organization and Intergovernmental Oceanographic Commission, 2017. *The Ocean Data Standards and Best Practices Project (ODSBP).* Paris: World Meteorological Organization and Intergovernmental Oceanographic Commission.

World Meteorological Organization, Japan Meteorological Agency, Deutscher Wetterdienst, 2012. *Assessment of the observed extreme conditions during late boreal winter 2011/2012.* Geneva: s.n.

World Meteorological Organization, 1994. *Data acquisiiton and processing, analysis, forecasting and other applications. Guide to hydrological practices..* 5 ed. Geneva: World Meteorological Organization.

World Meteorological Organization, 2010. *Assessment of the observed extreme conditions during the 2009/2010 boreal winter.* Geneva: World Meteorological Organization.

World Meteorological Organization, 2013. *The Global Climate 2001-2010: a decade of climate extremes.* Geneva: World Meteorological Organization.

World Meteorological Organization, 2014. *Climate Data Management System Specifications....* s.l.:World Meteorological Organization.

World Meteorological Organization, 2016. *The Global Climate in 2011-2015.* Geneva: World Meteorological Organization.

World Meteorological Organization, 2016. *WMO Statement on the state of the global climate in 2016.* Geneva: World Meteorological Organization.

World Meteorological Organization, 2016. *WMO Techical Regulations, Basic Documents No. 2, Volume II - Meteorological Service for International Air Navigation..* 2016 ed. s.l.:World Meteorological Organization.
World Meteorological Organization, 2017. *WMO Information System (WIS) 2.0 strategy.* s.l.:World Meteorological Organization.
