WMO Stewardship Maturity Matrix for Climate Data (SMM-CD)

The SMM-CD Working Group

Christina de Groot-Lief (WMO), Ge Peng (NCICS/NCEI),

Omar Baddour (WMO), William Wright (BOM),

Valentin Aich (WMO)

WEATHER CLIMATE WATER
TEMPS CLIMAT EAU

WMO OMM

World Meteorological Organization
Organisation météorologique mondiale

ESIP 2018
Tucson, AZ
July 17, 2018

Background

The **WMO Meteorological Organization (WMO)**

- is a specialized agency of the United Nations with 191 members states and territories.
- provides the framework for international cooperation for the development for meteorology, climatology and operational hydrology.
- is committed to, and continues to, facilitate free and unrestricted exchange of meteorology and related data and information, products, and services.

WMO OMM

Introduction

- The **WMO Commission for Climatology (CCl)** inter-programme initiative called the ***High Quality Global Data Management Framework for Climate*** **(HQ-GDMFC)** aims at making use of high quality climate data needed for developing climate services for policy and decision making in a variety of applications.
- A key priority for the HQ-GDMFC is harmonizing the definitions and processes, and developing a manual to guide collaborative entities on standards and best practices in the field of data management and stewardship.

SMM-CD: Development History

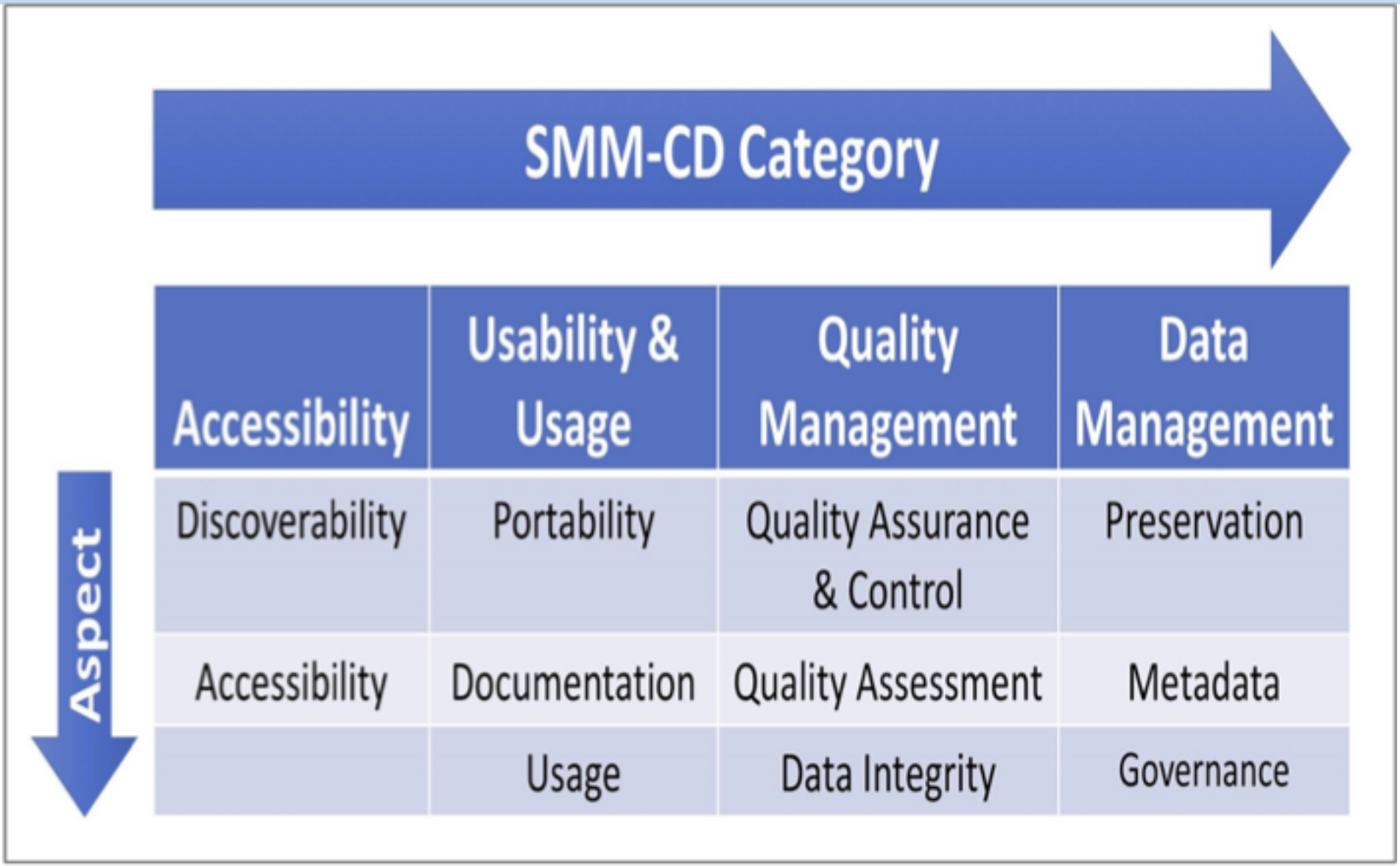
- The **WMO International Workshop on Information Management** (October 2017) concluded that a concept of trusted datasets needed to be defined.
- The **Expert Team on Climate Data Modernization (ET-CDM)** held a meeting (April 2018) to implement the concept of a climate data-specific maturity matrix model.
- The **WMO SMM-CD Working Group** is developing the matrix. Draft matrix and internal review have been completed. Presently the SMM-CD is being reviewed externally including the ESIP-wide review.

SMM-CD: Development Principles

The WMO SMM-CD:

- provides datasets with a score from 1 to 5 to help users assess their 'trustworthiness.'
- is based on existing models such as the NOAA/NCEI Data Stewardship Maturity Matrix (DSMM) and the European Space Agency (ESA) COordinating Earth observation data validation for RE-analysis for CLIMAte ServiceS (CORE-CLIMAX) System Maturity Matrix (SMM).
- has 4 categories: *Accessibility (Data Access), Usability & Usage, Quality Management and Data Management*. Each of these categories have several aspects.

SMM-CD: Categories and Aspects



Accessibility	Usability & Usage	Quality Management	Data Management
Discoverability	Portability	Quality Assurance & Control	Preservation
Accessibility	Documentation	Quality Assessment	Metadata
	Usage	Data Integrity	Governance

SMM-CD: Maturity Levels

Level 1	Level 2	Level 3	Level 4	Level 5
AD HOC	MINIMAL	INTERMEDIATE	ADVANCED	OPTIMAL
Not Managed	Limit-Managed	Managed	Well-Managed	Level 4 +
	Not Defined	Defined	Well-Defined	
		Partially Implemented	Fully Implemented	
				Measured, Controlled, Audited

WMO OMM

SMM-CD: Categories and Aspects

Accessibility (Data Access) Category

Refers to the ability to locate (**Discoverability**) and get to the datasets in question (**Accessibility**), with higher levels of maturity corresponding to the ease for a potential user to find and gain access to the dataset.

Aspects:

- Discoverability
- Accessibility

WMO OMM

ACCESSIBILITY (DATA ACCESS) Category and Aspects Table

Aspect	Level 1	Level 2	Level 3	Level 4	Level 5
Discoverability	By personal contact only; Dataset information not discoverable	Limited dataset information, such as scientific description of the methodology, in the literature	Minimal catalog-level metadata; Dataset searchable online	Complete set of collection-level discovery metadata + minimal granular metadata	Level 4 + available on an international catalogue, prominently displayed online and routinely updated
Accessibility	Data not available publicly; Person-to-person contact needed	Basic online services available for data access (e.g. FTP/HTTP direct download).	Non-standard data service	Standard-based interoperability data service	Level 4 + full capability of sub-setting, aggregation and visualization

SMM-CD: Categories and Aspects

Usability & Usage Category

Describes how easily the data products may be understood and used and be incorporated into the user's own working environment.

Aspects:

- Data Portability
- Documentation
- Usage

WMO OMM

USABILITY & USAGE Category and Aspects Table

Aspect	Level 1	Level 2	Level 3	Level 4	Level 5
Data Portability	Non-machine readable	Machine readable	Standards-based machine readable	Machine independent, self-describing, interoperable format	Level 4 + capability of providing user required format
Documentation	Product information not publicly available online	Limited online documentation (e.g., User Guide)	Document on how the data product was created and how to use it, is available online	Full documentation based on a standard template and available online	Level 4 + online tutorial on using and analysing the dataset Complete production system information available online
Usage	Weak citations in scientific publication in peer-review journal or as institutional reports	Intermediate citations + referenced in institutional climate assessment reports (e.g., by NOAA)	Strong citations + referenced in national climate assessment reports (e.g., by USGCRP)	Level 3 + referenced in international climate assessment reports (e.g., by IPCC)	Level 4 + referenced in international decision/policy making published reports (e.g., by UNFCCC, UN-ISDR, World Bank, etc.)

SMM-CD Categories and Aspects

Quality Management Category

Encompasses quality assurance procedures including quality monitoring, quality control and quality assessment and communication of reliability.

Aspects:

- Quality assurance & control
- Quality assessment
- Data integrity

WMO OMM

QUALITY MANAGEMENT Category and Aspects Table

Aspect	Level 1	Level 2	Level 3	Level 4	Level 5
Quality Assurance & Control	Data quality assurance (QA) & control (QC) procedure unknown or none	QA/QC procedure are defined, documented, and partially implemented.	QA/QC procedure are well-defined according to community best practices, documented and fully applied.	Level 3 + provision of error statistics published or tracked with results made available online and communicated to data providers; Procedure for user feedback, improvement prioritization in place	Level.4 + detailed analysis of errors and gaps at space-time unit level: (Station, grid-points, daily, monthly and or annual time-scale, etc.) QA/QC procedure monitored
Quality Assessment	Product quality assessment not done or done internally and information not available	Assessed by PI or data producer; Assessment results available online including error source estimates	Level 2 + collection-level uncertainty estimates and quality flags are available and methodology documented online.	Level 3 + cell-level uncertainty estimates and quality flags Product validation and evaluation published in peer-reviewed journal	Level 4 + The complete product provenance is captured and publicly available.
Data Integrity	Unknown or no data ingest integrity check	Data ingest integrity verifiable (e.g., checksum technology)	Level 2 + Data archive integrity verifiable	Level 3 + Data access integrity verifiable Conforming to community data integrity technology standard	Level 4 + Data authenticity verifiable (e.g., data signature technology) Performance of data integrity check monitored and reported

SMM-CD: Categories and Aspects

Data Management Category

Refers to the processes to ensure the data and the 'contextual' metadata are securely archived.

Aspects:

- Preservation
- Metadata
- Governance

WMO OMM

DATA MANAGEMENT Category and Aspects Table

Aspect	Level 1	Level 2	Level 3	Level 4	Level 5
Preservation	Any storage location; Data only; Data not backed up	Non-designated repository; A backup copy of electronic data is made	Designated archive; Basic retention policy defined. Routine backups made, including offsite copy.	Level 3 + conforming to community archiving standards. (including compliant with national archival policy); Comprehensive retention policy defined and executed.	Level 4 + archiving process performance controlled, measure and audited Future archiving standard changes planned
Metadata	Metadata not publicly available and/or not usable	Limited Metadata publicly available; Conforming to community-standard; Basic characteristics of dataset	Level 2 + conforming to international standards in most aspects; limited quality and provenance metadata	Fully compliant with international standards; Rich metadata content; Basic granular-level metadata; Support dataset provenance.	Level 4 + complete granular-level metadata; Metadata QC-ed and Regularly updated
Governance	Responsibility is not defined; No person is assigned.	Responsible entity is identified; Accountability and competency are not well-defined.	Responsibility, accountability, and compliance mechanisms are defined; Good competency; Processes established conforming to community standards	Level 3 + competency defined; Confirming to international standards; auditable	Level 4 + accountability and responsibility well-defined and fully compliant with international standards; transparent; Monitored and audited

SMM-CD: SUMMARY

- There has been an avalanche of climate data available on the web. Users are having a hard time getting information on the quality of the data.
- The SMM-CD provides a means to get quantifying information by scoring the quality of the data management of datasets.
- WMO is producing a catalog of highly scored datasets for the key climate indicator datasets and having these datasets appear prominently in the WMO Information System (WIS) data access feature as well as in the main search engines.
- The SMM-CD Working Group is interested in ESIP reviewing and commenting on the SMM-CD to ensure and improve its quality.
- A Poster of the SMM-CD is available for comment markup. You can also email comments to christina.lief@gmail.com or gpeng@ncsu.edu

WEATHER CLIMATE WATER
TEMPS CLIMAT EAU

Thank you

Christina de Groot-Lief
Christina.lief@gmail.com

WMO OMM

World Meteorological Organization
Organisation météorologique mondiale