An Introduction of the WMO Stewardship Maturity Matrix for Climate Data (SMM-CD)

Ge Peng, PhD

North Carolina State University, Cooperative Institute for Climate and Satellites–NC (CICS-NC) at NOAA's National Centers for Environmental information (NCEI)

October 22, 2018

WEATHER CLIMATE WATER
TEMPS CLIMAT EAU

WMO OMM

World Meteorological Organization
Organisation météorologique mondiale

The 46th Meeting of the Working Group on Information Systems & Services
German Aerospace Center (DLR)
Oberpfaffenhofen, Germany
October 22–25, 2018

WMO & WMO Information System (WIS):

- ***Specialized agency of the United Nations (weather, water, and climate): 191 member countries and territories;***
- ***Committed to free exchange of data and products;***
- ***Dedicated to ensuring the highest possible quality (data, information, and services) and providing effective access to authoritative, trusted datasets for science, policy and decision-making support.***

SMM-CD is developed by the SMM-CD Working Group

- ***Ge Peng (CICS-NC/NCEI, USA), lead;***
- ***William Wright (BOM, Australia), co-lead;***
- ***Christina Lief (WMO);***
- ***Omar Baddour (WMO);***
- ***Valentin Aich (GCOS)***

under the WMO High-Quality Global Data Management Framework for Climate (HQ-GDMFC), in collaboration with the members of an ad hoc International Expert Group on Climate Data Modernisation (IEG-CDM)

To Help Address Some of the Challenges Facing WMO & WIS

HQ-GDMFC

- *Inter-Programme Initiative Led by WMO CCl/CBS (Commission for Climatology/Commission for Basic Systems);*
- *Other Key Sponsors & Stakeholders:*
  - **WCRP** (World Climate Research Programme);
  - **JCOMM** (Joint WMO-IOC Technical Commission for Oceanography and Marine Meteorology);
  - **CHy** (WMO Commission for Hydrology);
  - **GCOS** (Global Climate Observing System)

HQ-GDFMC

A collaborative Framework that enables an effective development and exchange of high-quality climate data based on reliable underpinning infrastructure at the global, regional, and national levels.

Building blocks

1. Data Management Standards

Promoting data management standards and best practices for ensuring high quality datasets for use in climate policy and services

2. Data Maturity Assessment

Analyzing the maturity of the climate data management, identifying gaps, and fixing stewardship issues. SMM-CD: a mechanism for allowing compliance to WMO and internationally agreed stewardship standards

3. Access to High Quality Datasets

Enabling a quick discovery and access of high quality datasets using a federated cataloguing service compatible with the WMO Information System and international search engines

(Courtesy of Omar Baddour)

HQ-GDMFC

High Quality-Global Data Management Framework for Climate

Catalogue

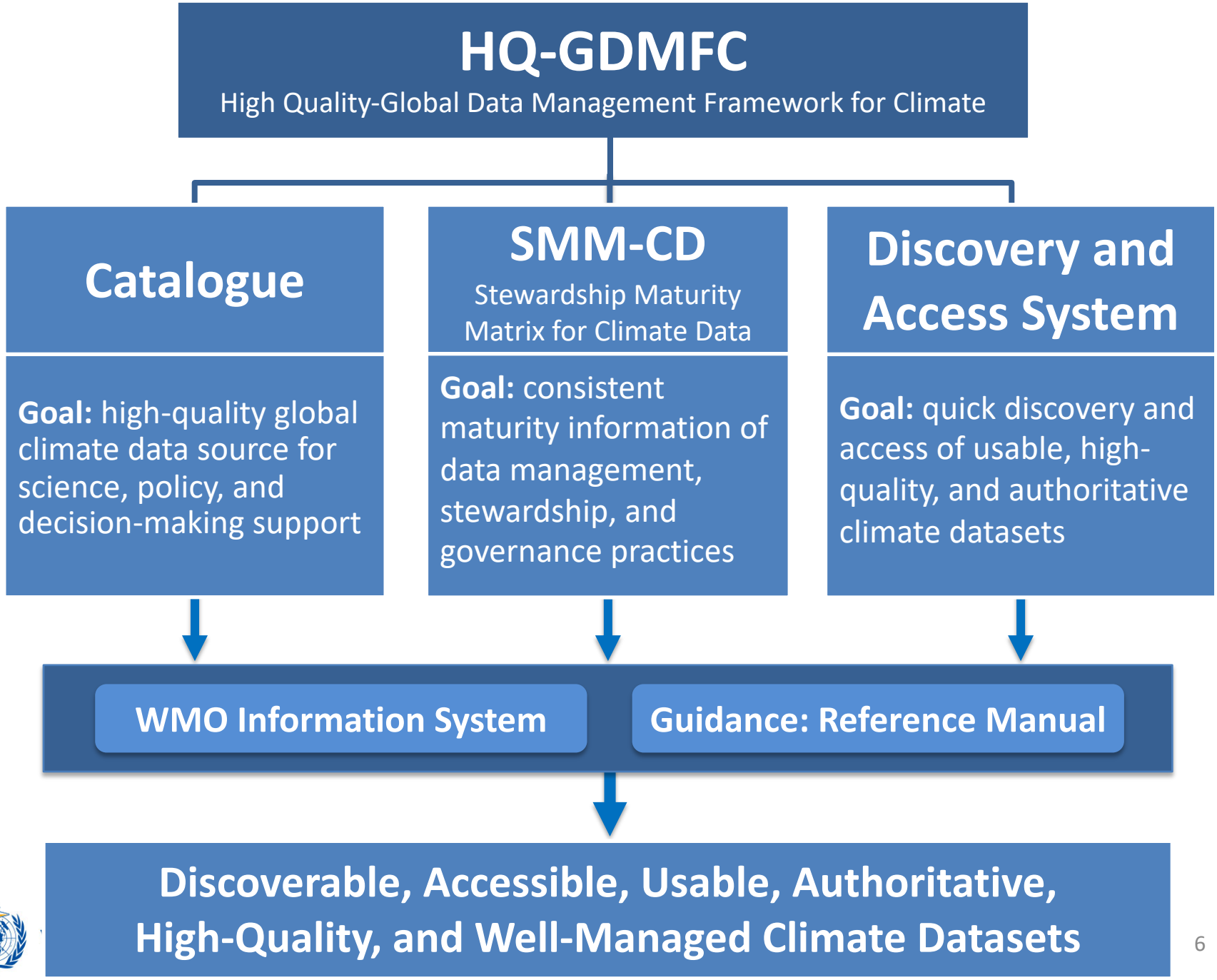
Goal: high-quality global climate data source for science, policy, and decision-making support

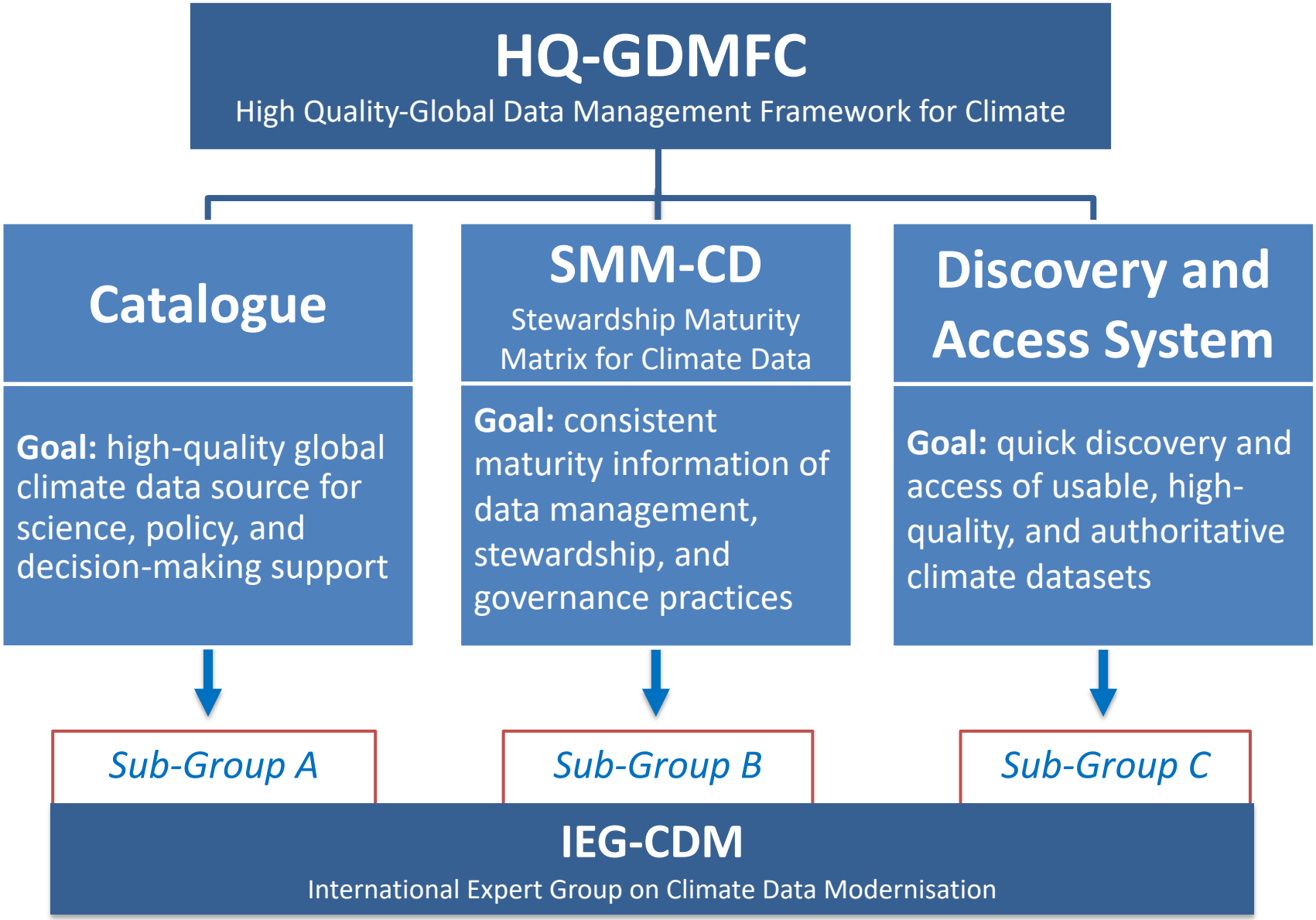
Matrix

Goal: consistent maturity information of data management, stewardship, and governance practices

Discovery and Access System

Goal: quick discovery and access of usable, high-quality, and authoritative climate datasets


HQ-GDMFC
High Quality-Global Data Management Framework for Climate
Catalogue
Goal: high-quality global climate data source for science, policy, and decision-making support
SMM-CD
Stewardship Maturity Matrix for Climate Data
Goal: consistent maturity information of data management, stewardship, and governance practices
Discovery and Access System
Goal: quick discovery and access of usable, high-quality, and authoritative climate datasets
WMO Information System
Guidance: Reference Manual
Discoverable, Accessible, Usable, Authoritative, High-Quality, and Well-Managed Climate Datasets



The group met at KNMI (Royal Netherlands Meteorological Institute), 16–18 April 2018

WMO OMM

HQ-GDMFC

High Quality-Global Data Management Framework for Climate

Catalogue

Goal: high-quality global climate data source for science, policy, and decision-making support

Sub-Group A

Initial Set of 16 Datasets

SMM-CD

Stewardship Maturity Matrix for Climate Data

Goal: consistent maturity information of data management, stewardship, and governance practices

Sub-Group B

Initial SMM-CD Scope & Draft

Discovery and Access System

Goal: quick discovery and access of usable, high-quality, and authoritative climate datasets

Sub-Group C

Key Metadata Requirements

(https://wiswiki.wmo.int/tiki-index.php?page=WWIM-Data-2018-1)

HQ-GDMFC

High Quality-Global Data Management Framework for Climate

Catalogue

Goal: high-quality global climate data source for science, policy, and decision-making support

Sub-Group A

IEG-CDM

SMM-CD

Stewardship Maturity Matrix for Climate Data

Goal: consistent maturity information of data management, stewardship, and governance practices

Sub-Group B

IEG-CDM

The SMM-CD

Working Group

Discovery and Access System

Goal: quick discovery and access of usable, high-quality, and authoritative climate datasets

Sub-Group C

IEG-CDM

WMO OMM

Dataset Quality: Multi-dimensional

Data Product Lifecyle-Stage-Based Maturity Assessment Models

Science Maturity Matrix	**Product** Maturity Matrix	**Stewardship** Maturity Matrix	**Service** Maturity Matrix
Zhao et al. (2016) **Thorne et al. (2015)**	**Bates and Privette (2012)**	**Peng et al. (2015)**	**NCEI MM-Serv WG (under-development)**
	EUMETSAT (2013)		

Scientifically sound and utilized

Fully documented and transparent

Well-preserved and integrated

Readily obtainable and usable

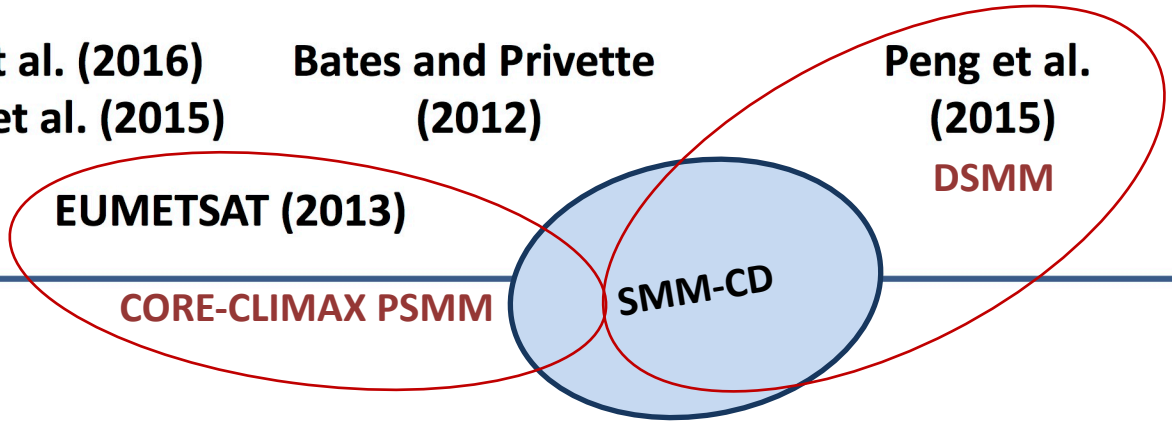
(Peng 2018, *Data Science Journal)*

WMO OMM

The Scope of SMM-CD

Data Product Lifecyle-Stage-Based Maturity Assessment Models

Define/Develop/Validate	Produce/Assess/Deliver	Maintain/Preserve/Disseminate	Use/User Service
Science	Product	Stewardship	Service
Science Maturity Matrix	**Product** Maturity Matrix	**Stewardship** Maturity Matrix	**Service** Maturity Matrix
Zhao et al. (2016) **Thorne et al. (2015)**	**Bates and Privette (2012)**	**Peng et al. (2015)** **DSMM**	**NCEI MM-Serv WG (under-development)**



WGISS DMSMM is one of the references for the SMM-CD

WMO OMM

The Structure of SMM-CD

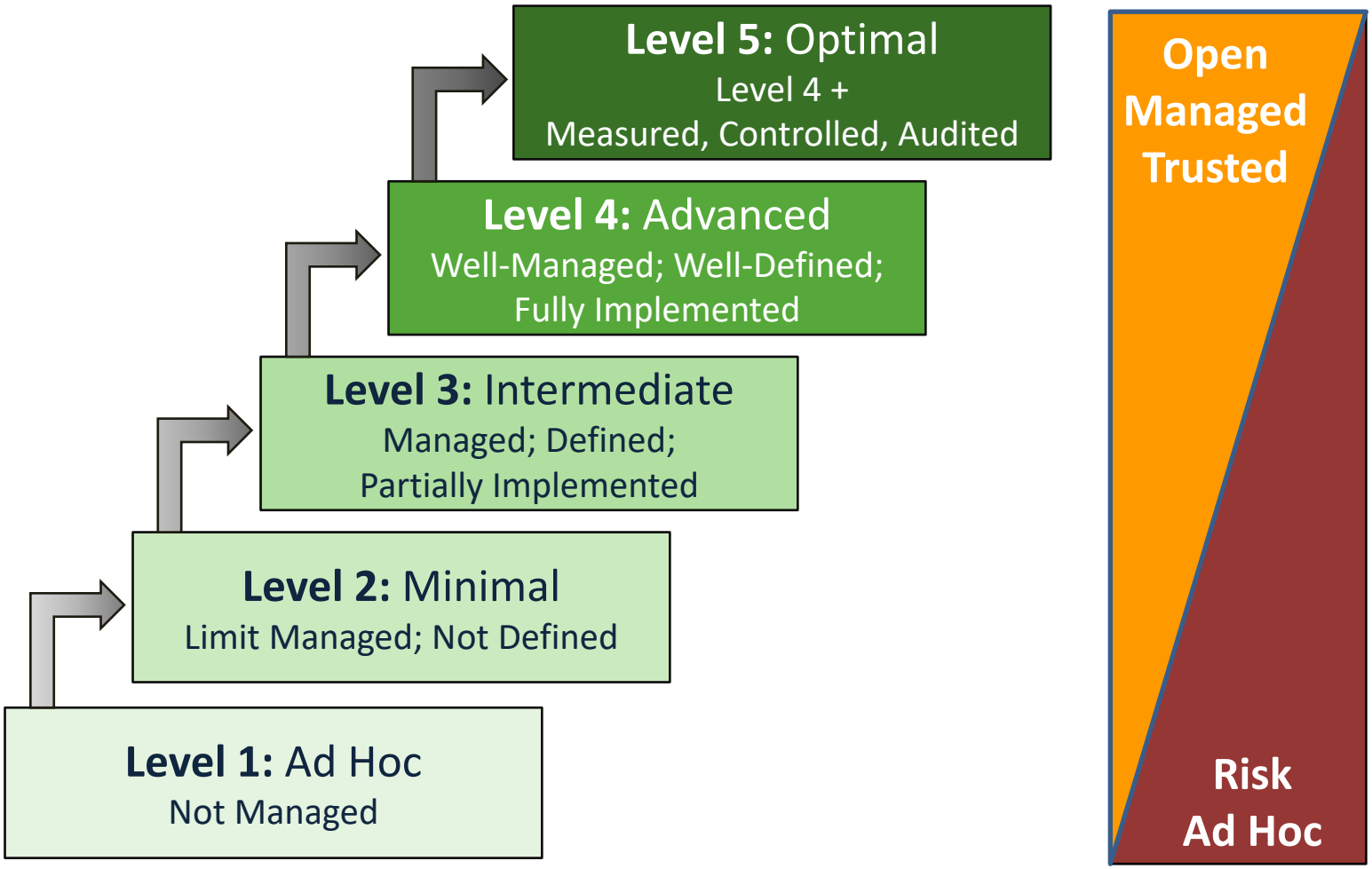
SMM-CD Category

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

Aspect

The state or ability to locate (Discoverability) and get to the dataset (Accessibility)	*How easily the data product may be understood and integrated by users; the usage and impact of the dataset*	*The state of quality assurance, control, and assessment; data uncertainty and reliability, and data fixation*	*The state of the dataset preservation, metadata completeness, and governance practices*

SMM-CD: Maturity Scale Structure

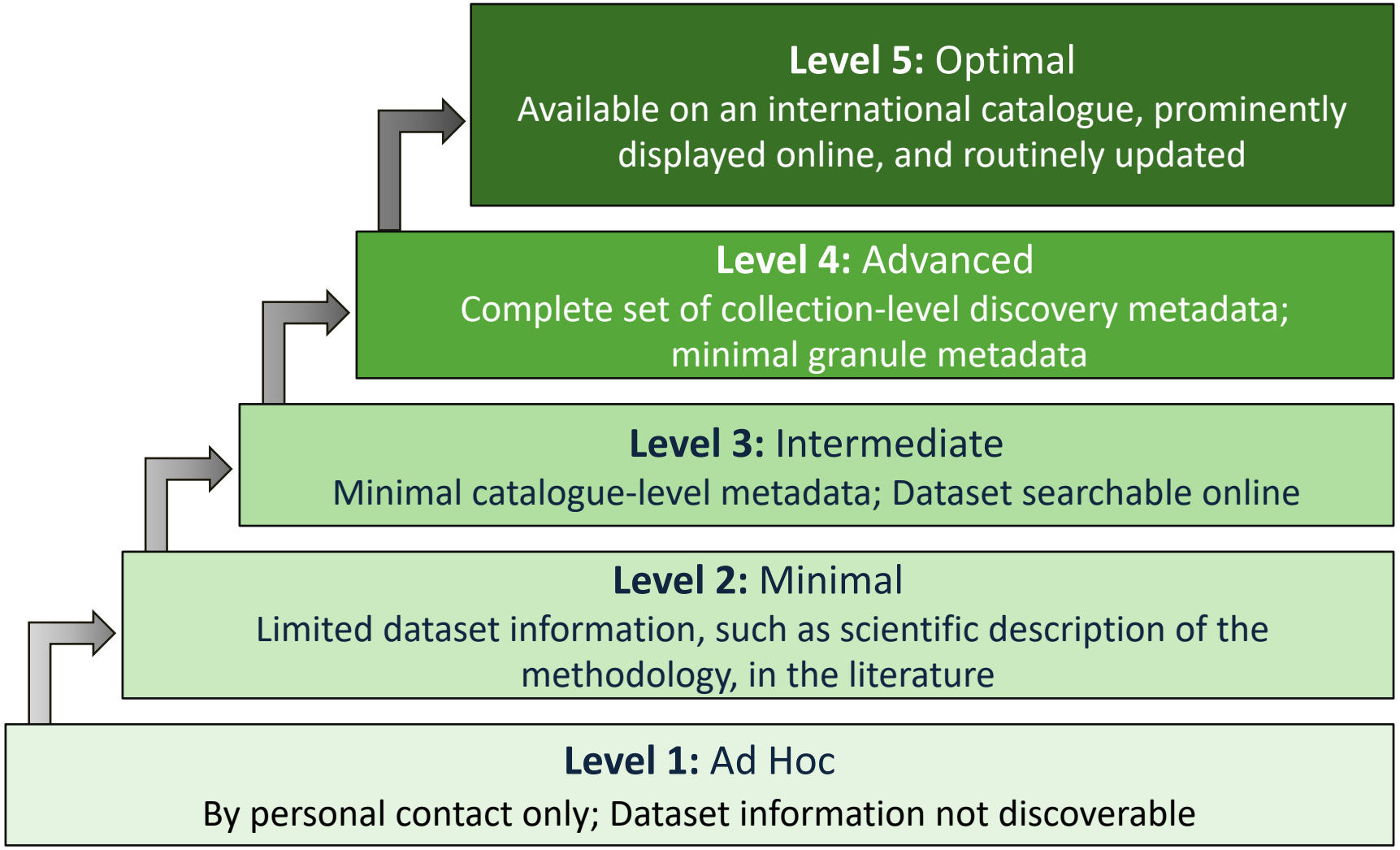


Reference Maturity Level Structure

- Capability Maturity Model Integration (CMMI)
- Levels of Maturity of Digital Repository

WMO OMM

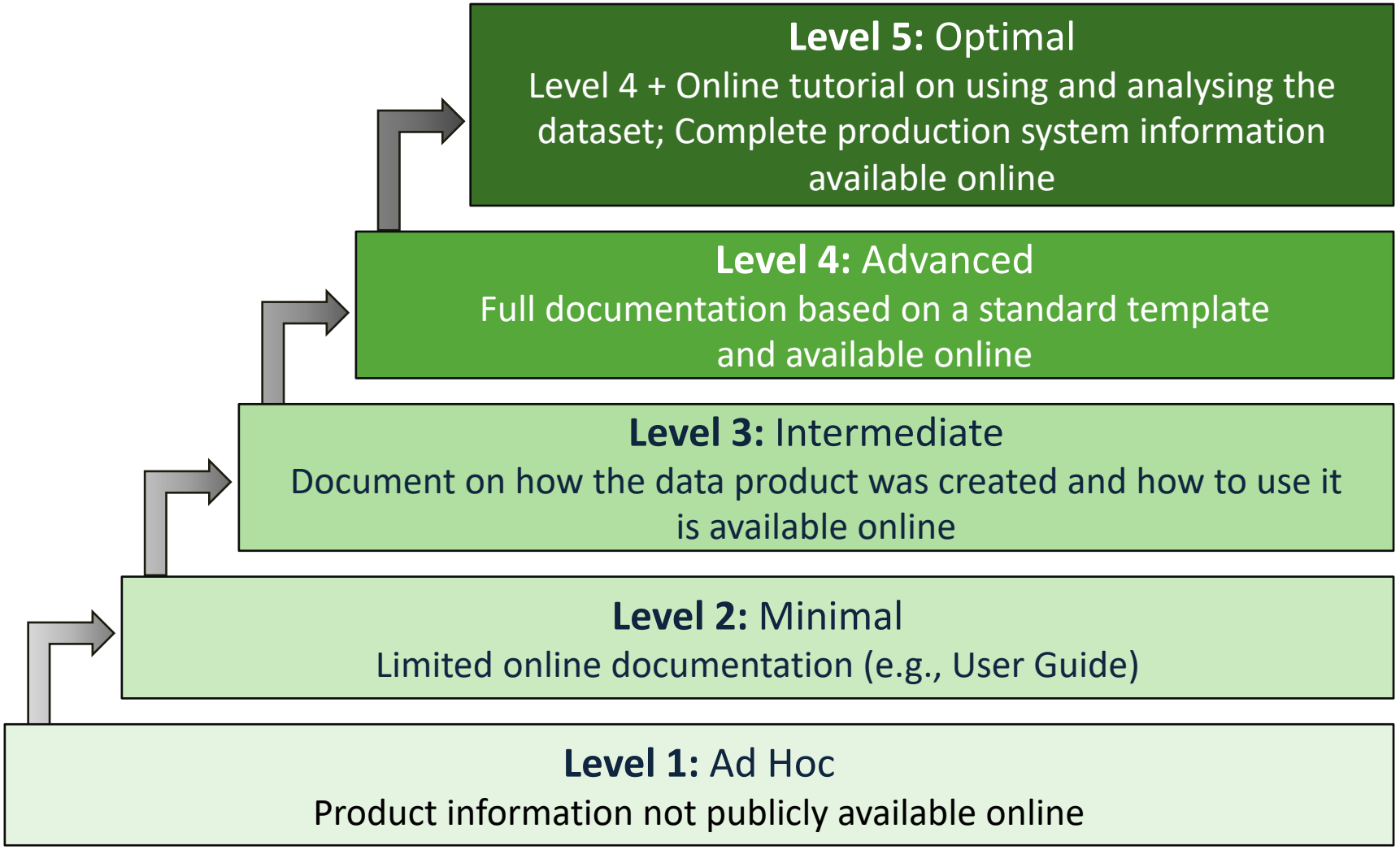
SMM-CD: Maturity Level Definitions



Aspect: Discoverability

Category: Data Access

SMM-CD: Maturity Level Definitions

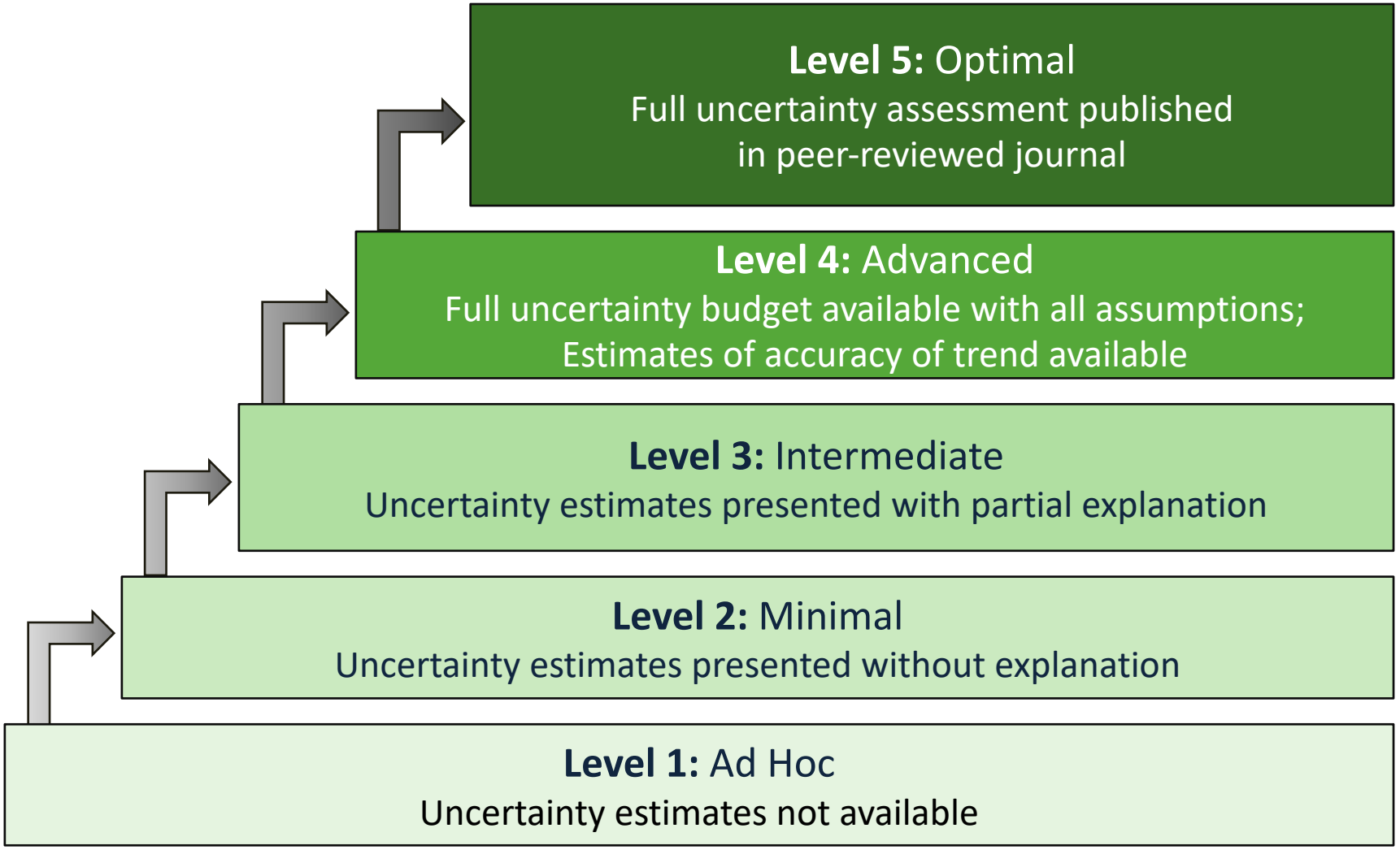


Aspect: Documentation

Category: Usability and Usage

WMO OMM

SMM-CD: Maturity Level Definitions

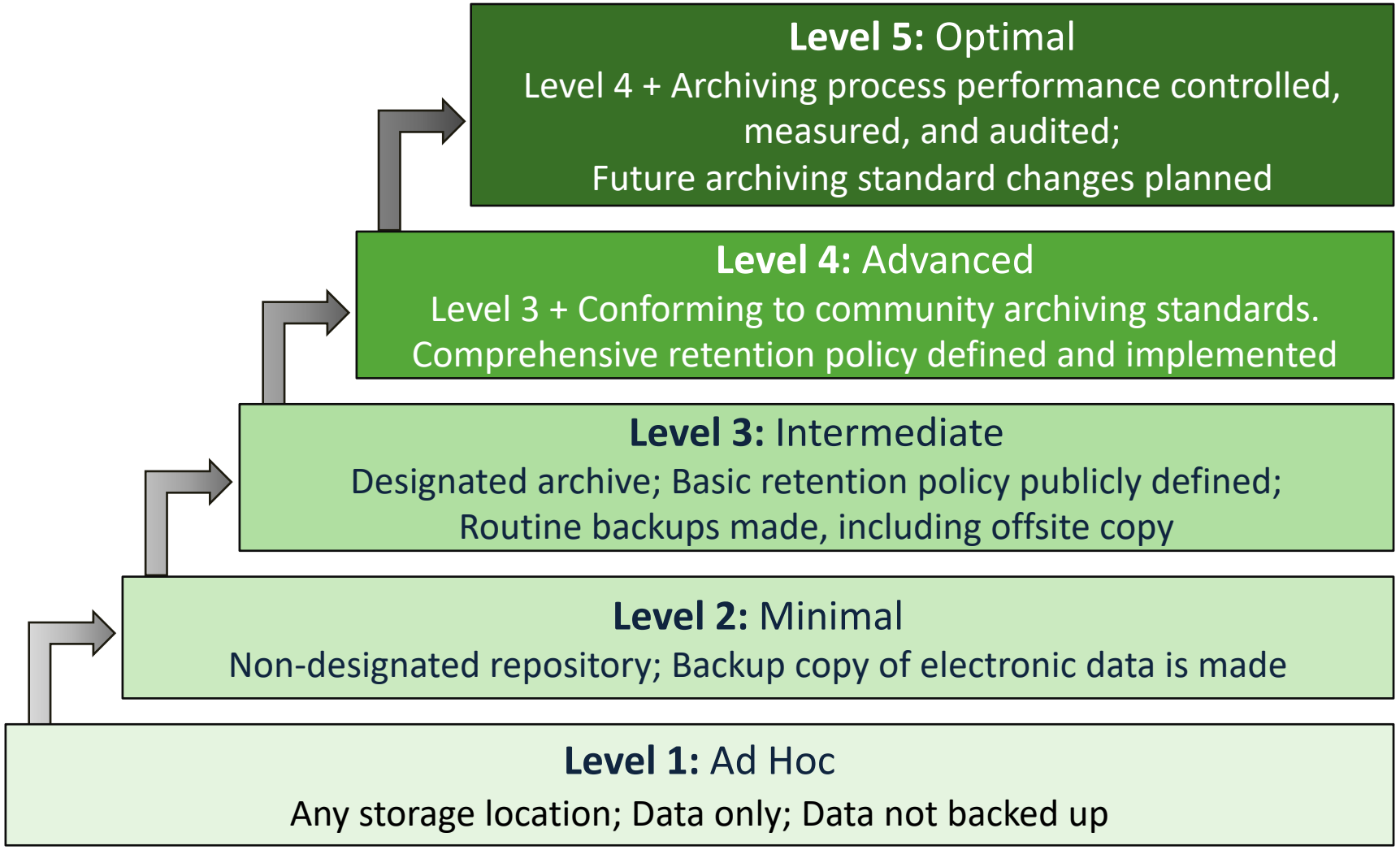


Aspect: Uncertainty Analysis

Category: Quality Management

WMO OMM

SMM-CD: Maturity Level Definitions



Aspect: Preservation

Category: Data Management

Outcomes

- **A Matrix and A Guidance Booklet**
  - **Internal IEG-CDM team review**
  - **External community-wide reviews:**
    - ✓ **Invited international domain experts (science, data management, and stewardship);**
    - ✓ **GCOS secretariat;**
    - ✓ **ESIP (Earth Science Information Partner) community – a working session at the ESIP 2018 summer meeting in July**
- **An Evaluation Template**

WMO SMM-CD Modification POC (Name; E-mail; Affiliation)						
Categories and Aspects	**Level 1 Ad Hoc Not Managed**	**Level 2 Minimal Limit-Managed Not Defined**	**Level 3 Intermediate Managed Defined, Partially Implemented**	**Level 4 Advanced Well-Managed Well-Defined, Fully Implemented**	**Level 5 Optimal Level 4 + Measure, Controlled, Audited**	The Data Access category refers to the ability to locate (Discoverability) and get to the dataset in question (Accessibility), with higher levels of maturity corresponding to the ease for a potential user to find and gain access to the dataset
Data Access						**WMO SMM-CD Rating and Justification or Evidence** / **Comments**
Discoverability	By personal contact only; Dataset information not discoverable	Limited dataset information, such as scientific description of the methodology, in the literature	Minimal catalogue-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	❖ Level /
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 services	Standard-based interoperability data services	Level 4 + Full capability of sub-setting, aggregation and visualization	❖ Level /
Usability & Usage	**Level 1 Ad Hoc Not Managed**	**Level 2 Minimal Limit-Managed Not Defined**	**Level 3 Intermediate Managed Defined, Partially Implemented**	**Level 4 Advanced Well-Managed Well-Defined, Fully Implemented**	**Level 5 Optimal Level 4 + Measure, Controlled, Audited**	The Usability & Usage category describes how easily the data products may be understood and used by users and incorporated into the user's own working environment
						WMO SSM-CD Rating and Justification or Evidence / **Comments**
Data Portability	Non-machine readable	Basic machine readable	Standards-based machine readable	Machine independent, self-describing, interoperable format	Level 4 + capability of providing user required format	❖ Level /
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 analyzing the dataset; Complete production system information available online	❖ Level /
Usage	No or weak citation 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.)	❖ Level /

Current Status

- **SMM-CD documents have been baselined;**
- **Use case of 16 global datasets identified by IEG-CDM Sub-Group A is underway**

  (Datasets: http://www.wmo.int/pages/prog/wcp/ccl/opace/opace1/meetings/documents/DraftMeetingReport.pdf)

  - Five assessments completed;
  - A couple more near completion.

The latest unofficial version of three SMM-CD documents are available at Figshare.com.

The short URLs:

- **Matrix:** bit.ly/WMO-SMM-CD
- **Guidance Booklet**: bit.ly/SMM-CD-Manual
- **Template:** bit.ly/SMM-CD-Template

WEATHER CLIMATE WATER
TEMPS CLIMAT EAU

WMO OMM

World Meteorological Organization
Organisation météorologique mondiale

Acknowledgement

The members of IEG-CDM (in alphabetical order)

1. AICH, Valentin (GCOS)
2. BADDOUR, Omar (WMO)
3. BERGERON, Cedric (ECMWF)
4. BEROD, Dominique (WMO)
5. BUSSELBERG, Thorsten (DWD)
6. CAZENAVE, Anny (LEGOS)
7. DUNN, Robert (Met Office/Hadley Center)
8. GALLAHER, David (NSIDC)
9. GATES, Lydia (DWD)
10. LIEF, Christina (WMO, lead)
11. MILAN, Anna (NOAA/NCEI)
12. PENG, Ge (NCSU/CICS-NC, NOAA/NCEI)
13. ROBERTS, Kate (BOM)
14. SIEGMUND, Peter (KNMI)
15. VERVER, Ge (KNMI)
16. WRIGHT, William (BOM)
17. ZIESE, Markus (DWD)

WEATHER CLIMATE WATER
TEMPS CLIMAT EAU

WMO OMM

World Meteorological Organization
Organisation météorologique mondiale

Acknowledgement

Discussions with and feedback from the following domain SMEs are beneficial:

Iolanda Maggio, Peter Thorne, Simon Eggleston, Darren Ghent, Jörg Schulz, Nancy Ritchey, Kenneth Kehoe, Imke Durre, Carolin Richter, Ruth Duerr

Special THANKS to

Christina Lief, William Wright, Omar Baddour, and Valentin Aich;

Management of NCEI's Center for Weather and Climate;

Management of CICS-NC;

NCEI Communication Team for copyediting the slides

WEATHER CLIMATE WATER
TEMPS CLIMAT EAU

WMO OMM

World Meteorological Organization
Organisation météorologique mondiale

References

- **CORE-CLIMAX Production System Maturity Matrix (PSMM):** EUMETSAT 2013 CORE-CLIMAX Climate Data Record Assessment Instruction Manual. Version 2, 25 November 2013.
- **NCEI/CICS-NC Scientific Data Stewardship Maturity Matrix (DSMM):** Peng, G, Privette, JL, Kearns, EJ, Ritchey, NA, and Ansari, A 2015 A unified framework for measuring stewardship practices applied to digital environmental datasets. *Data Science Journal*, 13. doi:10.2481/dsj.14-049.

WEATHER CLIMATE WATER
TEMPS CLIMAT EAU

WMO OMM

World Meteorological Organization
Organisation météorologique mondiale

An EGU 2019 Session

For science data centers and repositories:

- Establishing trustworthiness and fitness for purpose, i.e., suitability, at the level of individual data products and services

For end-users:

- Finding content-rich, interoperable, and accessible quality descriptive information

Call for approaches, frameworks, workflows, best practices, tools, etc., that are under development or being implemented towards:

- systematically evaluating quality attributes of individual data products and services,
- automatically generating content-rich quality descriptive information that is interoperable and discoverable.

https://meetingorganizer.copernicus.org/EGU2019/session/30950

WEATHER CLIMATE WATER
TEMPS CLIMAT EAU

Thank you!

Any Comments or Suggestions?

Ge.Peng@noaa.gov

WMO OMM

World Meteorological Organization
Organisation météorologique mondiale

To Cite This Presentation:

Peng, G., 2018: An Introduction of the WMO Stewardship Maturity Matrix for Climate Data (SMM-CD). Version: v01r00 20181020. *The 46th Meeting of the CEOS Working Group on Information Systems and Services.* 22–25 October 2018, Oberpfaffenhofen, Germany. doi: https://doi.org/10.6084/m9.figshare.7289738.

WEATHER CLIMATE WATER
TEMPS CLIMAT EAU

WMO OMM

World Meteorological Organization
Organisation météorologique mondiale