**GUIDANCE BOOKLET**

**WMO Stewardship Maturity Matrix for Climate Data for**

**National and Regional Purposes (SMM-CD\_NRP)**

Developed by the Stewardship Maturity Matrix Task Team for National Implementation (SMM-TTNI), consisting of members of the CCl Expert Team on Data Development and Stewardship (ET-DDS)1 & the International Expert Group on Climate Data Modernisation (IEG-CDM)2

Document ID: WMO-SMM-CD\_NRP-0001

Version: v01r00 20200805

**Members of the SMM-TTNI:**

1. Ali Eddenjal (Libyan National Meteorological Center)1
2. Christina Lief (WMO CCl)1,2
3. David Sinclair (BOM)1
4. Dubuisson Brigitte (MeteoFrance)1
5. Fatima Hdidou (Morocco Meteorological Service)1
6. Ge Peng (NCSU/CISESS, NOAA/NCEI)1,2
7. Jean-François Legeais (Collecte Localisation Satellites)2
8. José A. Guijarro (AEMET)1
9. Markus Donat (Barcelona Supercomputing Center)1
10. Markus Ziese (DWD)2
11. Omar Baddour (WMO)2
12. Peter Siegmund (KNMI)2
13. Reinaldo Silveira (Sistema Meteorológico do Paraná – SIMEPAR, Brazil)1
14. Robert Dunn (Met Office/Hadley Center)2
15. William Wright (WMO CCl)2
16. Xiaolan Wang (Environment and Climate Change Canada)1

**DOCUMENT HISTORY**

|  |  |  |
| --- | --- | --- |
| **Version** | **Contributors** | **Contribution/What is new** |
| v01r00 20200805 | The Stewardship Maturity Matrix Task Team for National Implementation (SMM-TTNI) | First baseline of a WMO Stewardship Maturity Matrix for Climate Data for National and Regional Purposes (SMM-CD\_NRP), incorporating this document (“Guidance Booklet”) |

**Disclaimer** This document is a guidance booklet to help facilitate the use of the Stewardship Maturity Matrix for Climate Data for National and Regional Purposes (SMM-CD\_NRP) of the World Meteorological Organization (WMO) by providing additional background information and resources. The views expressed herein do not necessarily reflect those of WMO or its Members. The document will be subjected to future evolution as knowledge improves and user requirements expand and/or change.

**TABLE OF CONTENT**

1. EXECUTIVE SUMMARY
2. BACKGROUND
3. SCOPE, RATIONALE, AND INTENDED AUDIENCE
4. CATEGORIES, ASPECTS AND MATURITY SCALE STRUCTURE
5. DEFINING CATEGORIES, ASSOCIATED ASPECTS AND MATURITY LEVELS
	1. Operational Data Management category
	2. Data Stewardship category
6. SELF-ASSESSMENT TEMPLATE
7. ACKNOWLEDGMENT
8. REFERENCES
9. DEFINITIONS
10. ACRONYMS
11. **EXECUTIVE SUMMARY**

The World Meteorological Organization (WMO) is a United Nations’ specialized agency with members from 193 countries and territories, in the field of weather, water, and climate. As part of its activities, WMO fosters international collaboration to develop technical guidance and standards for the collection, processing, management of data and forecast products. Under the technical advice of its Commission for Climatology (CCl), WMO has established a High-Quality Global Data Management Framework for Climate (HQ-GDMFC) initiative as an international collaborative initiative for developing standards and recommended practices for sourcing, securing and managing climate data, and for sharing infrastructure and responsibilities for, e.g. data exchange, analysis and data service provision. The goal of the initiative is to ensure mature data management and governance that adds to the already-established Global Climate Observing System (GCOS) climate data source reviewing process.

The HQ-GDMFC consists of three building blocks: (1) The Manual on the HQ-GDMFC, which is part of the WMO Regulatory material, and provides regulations and recommended best practice on climate data management. This was recently adopted by the eighteenth World Meteorological Congress (Cg-18), and has been published under WMO-No.1238 (WMO, 2019); (2) The WMO Stewardship Maturity Matrix for Climate Data (SMM-CD) was developed as a tool to enable dataset owners to self-assess and rate the stewardship of their datasets quantifiably, based on internationally validated data stewardship best practices. It constitutes a formal annex to the WMO-No.1238, and focuses on datasets that are of global use, such as for global climate monitoring and assessment. The process helps WMO in evaluating the stewardship maturity of the datasets which are to be submitted for inclusion in the third building block, (3) the WMO Catalogue for Climate Data. This online catalogue (<https://climatedata-catalogue.wmo.int/>) references easily discoverable and accessible high-quality climate datasets that are characterized by well documented methodologies and practices for their creation, management, stewardship, and governance. Such a catalogue provides an authoritative and trustworthy source of climate datasets useful for producing information on key climate indicators. Most prominent target users for the catalogue include, but are not limited to, the climate policy community in the context of the Paris Agreement Global Stocktake, which is expected to start in 2023. The initial climate datasets cover temperature, precipitation, ice sheets, sea ice, sea level, glaciers, and crowd-sourced data. The data included in the catalogue is intended to be discoverable and accessible in a prominent position through the WMO Information System (WIS) and internet search engines.

Recent activities have addressed the need to extend the self-assessment capability for datasets to the national and regional levels with the development of a targeted Stewardship Maturity Matrix for National and Regional Purposes (SMM-CD\_NRP) based on the SMM-CD which was developed for global scale datasets.

1. **BACKGROUND**

WMO is the United Nations’ authoritative voice on the state and behavior of the Earth's atmosphere, its interaction with the land and oceans, the weather and climate it produces, and the resulting distribution of water resources, snow, and ice.

There are several challenges regarding climate data that hamper full implementation of high quality climate services. Among these, much of the existing guidance on climate data management is out of date due to rapid recent advances in technologies and data management practices. The National Meteorological and Hydrological Services (NMHS) of many Member countries report lack of capacity in climate data management and there are significant gaps in stewardship, including a lack of standardization of terminology, processes, and practices due to an inadequate regulatory framework. On the other hand, there is an opportunity to make better use of the proliferation of new data sources and advances in technology that can support climate services.

The WMO initiative on the High Quality Global Data Management Framework for Climate (HQ-GDMFC) aims at making use of an extended range of climate data types needed to support the Climate Services Information System (CSIS) of the Global Framework for Climate Services (GFCS). It also aims to harmonize the definitions and processes that deal with climate data, and provide a more rigorous, standards-based regulatory framework around climate data practices. The Commission for Climatology (CCl)[[1]](#footnote-1), in collaboration with other Technical Commissions, submitted in June 2019 a Manual on the High Quality Global Data Management Framework for Climate at the eighteenth session of the WMO Congress, which welcomed the manual and issued resolution 22 (Cg18) adopting it as part of the WMO technical regulations.

The effort to develop an assessment tool for datasets focused initially on the global level, and the concept was developed during the WMO Workshop on Information Management (WWIM) which recommended the development of a project plan for identifying high quality climate datasets and providing access to them ([WWIM Meeting Report, October 2017](https://climatedata-catalogue.wmo.int/sites/default/files/WWIM-2017-FinalReport.pdf)). A key conclusion was that a process was needed to define how datasets can be endorsed by WMO. Essentially, in addition to being high-quality climate data products in themselves, how datasets are managed must also meet standards that demonstrate strong stewardship and governance of the data. The WWIM and following Commission for Basic Systems (CBS) Task Team on Information Management (TT-IM) developed a framework for a generic maturity model for information management, intended to be applied to all WMO domains.

The International Expert Group on Climate Data Modernisation (IEG-CDM) was then set up, and at a workshop held at KNMI, developed a climate data-specific version of a stewardship maturity assessment model for global datasets, to be used to assess and score how the individual climate datasets are managed and stewarded ([IEG-CDM Meeting Report, April 2018](https://climatedata-catalogue.wmo.int/sites/default/files/KNMI-Workshop-2018-FinalReport.pdf)):

The key points addressed during the KNMI meeting were:

* 1. Development of a WMO-wide Stewardship Maturity Matrix for Climate Data (SMM-CD), based on existing maturity assessment models. The SMM-CD assesses and scores various aspects of the management and stewardship of datasets. This model has been subjected to a broad review and road-testing process;
	2. Identifying an initial and provisional limited number of global climate related datasets with the aim that after proper evaluation, these could become part of the WMO Climate Data Catalogue;
	3. Development of a data discovery and access process for the catalogue through the WMO Information System (WIS) and major internet search engines, by recommending key metadata requirements and methods for optimising internet searches. The aim is that non-technical users can easily discover these high-quality datasets.

After the development in 2019 of the SMM-CD, which is used to mainly assess global datasets, the need to assess national and regional datasets was recently addressed. The SMM Task Team for National Implementation (SMM-TTNI) was formed as a sub-team of the CCl Expert Team on Data Development and Stewardship (ET-DDS) and the International Expert Group on Climate Data Modernization (IEG-CDM). Members of the Task Team are from NMHSs and organizations representing the six WMO regions. The working group started work in March 2020 and developed an adapted version of the SMM-CD to assess national and regional datasets, termed the SMM-CD for National and Regional Purposes, hereafter SMM-CD\_NRP.

1. **SCOPE, RATIONALE, AND INTENDED AUDIENCE**

The SMM-CD\_NRP was developed to meet specific needs of the NMHS which are operational in nature and focus primarily on producing data for their national users. Because of the nature of the work of the NMHSs, publications are desirable but not necessarily part of a scored aspect. It has been included in the ‘Quality & Usage’ aspect in the ‘Highly Desirable’ level which is optional and allows referencing publication(s) for a dataset if available.

The following are some of the benefits of utilizing the self-assessment tools:

* + Allows NMHSs to assess their data management practices, identifying gaps and those aspects that would benefit most from process improvements;
	+ Allows Members to identify an appropriate level of process maturity that should be used for the data they are managing;
	+ Provides a reference model for helping prioritize cost planning, resource allocation and funding for data management;
	+ Improve ways for dissemination and sharing the NMHSs climate data set through the Web or integrated data archiving such as cloud based;
	+ Provides a roadmap, and a way of measuring progress towards, improving information management capability in support of WMO Programmes;
1. **CATEGORIES, ASPECTS AND MATURITY SCALE STRUCTURE**

The focus in the SMM-CD\_NRP was to retain two main categories, the Operational Data Management and Data Stewardship categories, which best inform the NMHSs on how to manage their data according to best management practices and standards. The goal is to provide NMHSs with a user-friendly self-assessment tool to determine gaps in their data management and stewardship, and provide a structure to move towards improved practices and standards to attain a satisfactory level of competence in data management and stewardship.

In Figure 1 below, the structure of the maturity scale of the SMM-CD\_NRP is presented. The scale can be summarized as follows:

* At Level 1, there are few or no procedures or processes defined or in place, or at least they are not reported or poorly documented. Often, information about what has been done to the dataset is not publicly available. The level is defined as ad hoc. For example, an individual researcher created a data file and stored it on their own hard disk.
* At Level 2, some efforts have been made to move the dataset to a managed state. The procedure or process is documented, but not fully, and is not compliant with an established national or international standard. The level shows limited management and only partially implemented stewardship. For example, datasets are available online but with limited information and metadata
* Level 3 is the highest level and can be regarded as a fully satisfactory level of management. Requirements or standards, procedures, and processes associated with that aspect are defined and compliant with national or international standards.
* There is an additional, optional, level which is labeled as ‘Highly Desirable’. The rating is a level 3 with additional features such as meeting ISO international standards and the netCDF interoperability format.

**Figure 1. Scale and Structure for the SMM-CD\_NRP**

|  |  |
| --- | --- |
| **Scale and Structure** | **Maturity Levels for Each Aspect** |
| **Categories/Aspects** | **Level 1** **(ad hoc)** | **Level 2 (medium)** | **Level 3** **(highest)** | **Highly Desirable** |
| **Ad Hoc** | **Medium** | **Highest** | **Level 3 +** |
| **Not Managed** | **Limited Managed** | **Managed** | **Level 3 +** |
| **Not implemented** | **Partially Implemented** | **Fully Implemented** | **Level 3 +**  |

The requirements or standard against which the maturity of a dataset is evaluated should be described in an assessment report prepared by the dataset point-of-contact or an evaluator. WMO defined requirements and standards are recommended where they are applicable. The ratings should be assessed at that level where all the descriptors in the current and lower levels are satisfied. A fraction may be used to indicate that one or more criteria may be satisfied at a level higher than the current. It should be noted that dataset maturity ratings are a snapshot of the current state which may evolve over time. Ideally the maturity ratings should be utilized to demonstrate the level of stewardship of individual datasets, and to identify priorities for improving stewardship quality.

1. **DEFINING CATEGORIES, ASSOCIATED ASPECTS AND MATURITY LEVELS**

The SMM-CD\_NRP identifies two main categories for assessment: Operational Data Management and Data Stewardship. The Operational Data Management category has five assessment aspects: Data Access, Data Portability, Data Preservation, Documentation and Data Integrity. The Data Stewardship category has 3 assessment aspects: Quality and Usage, Governance, and Metadata. Figure 2 below shows the matrix with categories and related aspects with definitions and examples.

**5.1** **OPERATIONAL DATA MANAGEMENT CATEGORY:** addresses the assessment of operations that are required to enable access, portability, archival, documentation, and ensuring data integrity in the NMHS data holdings. **Data access** assesses the extent to which a dataset can be found and accessed. **Data Portability** ranges from inability to transfer data in computerized form to fully machine-readable and interoperable. **Data preservation** assesses the security of the data such as backup procedures and retention policies. **Documentation** assesses the extent and accessibility of information on how to use the dataset or product, and hence users’ ability to determine its fitness for purpose. **Data integrity** is an important factor of quality management which monitors dataflow and ingest processes. It would be highly recommended that NMHS would implement an alert system to ensure that circumstances in which the data can be corrupted or inadvertently overwritten or deleted cannot occur; or if they do occur, can rapidly be detected and corrected. There should also be automated processes for ensuring that data expected are received, that timestamping is accurate, and that ingested data are not unrealistic.

**5.2 DATA STEWARDSHIP CATEGORY:** The second category of the SMM-CD\_NRP provides a rating on how well a dataset is stewarded, by assessing the quality and usage, governance, and metadata aspects. The **Quality and Usage** aspect assesses the degree to which robust quality control is carried out on the data, along with quality flagging or error estimates, and the extent to which scientific peers trust the data in conducting research and compiling reports. **Governance** refers to the extent to which controls, accountabilities and compliance mechanisms are put in place, and their adherence with community best-practice. **Metadata** covers how detailed the metadata are, online availability of this information, and extent of compliance with standards such as the WIGOS Metadata standard It is also important for users is to have up to date contact information for the dataset.

**Figure 2. SMM-CD\_NRP Matrix**

|  |  |  |  |
| --- | --- | --- | --- |
|   |   | **Mandatory Maturity Assessment Criteria** | **Optional** |
| **Category** | **Aspect** | **Level 1** **(ad hoc)** | **Level 2 (medium)** | **Level 3** **(highest)** | **Highly Desirable** |
|  **Operational Data Management** | **Data Access** | Dataset information not discoverable; Data not available publicly; person to person contact. | Limited dataset information and discovery metadata; Basic online services available for data access (e.g. ftp). | Level 2 + comprehensive discovery metadata; Dataset searchable online; access provided with enhanced online data services. | Level 3 + interoperable data services; web services; conforming to international standards such as the ISO or WIS Core Metadata profile; auditable; ability to download subset of dataset. |
|  **Data Portability** | Non-machine readable. | Basic machine readable. (e.g. ASCII, data and metadata separate in one file such as [gpcc\_first\_guess\_01\_2020.gz](https://opendata.dwd.de/climate_environment/GPCC/first_guess/2020/)). | Standards-based machine readable. | Machine independent, self-describing, interoperable format (e.g. netCDF). |
| **Data Preservation** | Data not routinely backed up. | Backup copy of electronic data is routinely made. | Level 2 + follow institutional archival practices including off-site copy; basic retention policy defined and implemented. | Level 3 + Conforming to National archiving standards; Comprehensive retention policy defined and implemented.  |
| **Documentation** | Product information not publicly available online. | Limited online documentation (e.g. User Guide). | Comprehensive documentation available online.  | Level 3 + available in standard format template (e.g. [INSPIRE](https://inspire.ec.europa.eu/), [NASA ATBD's](https://eospso.nasa.gov/content/algorithm-theoretical-basis-documents)). |
| **Data Integrity** | Unknown or no data integrity check. | Random data integrity check. | Data integrity systematically verified through effective checks and rectification actions, so that data expected is data received into the climate database.  | Level 3 + verified with documented audit trails, and/or traceability; also for data services for users and applications. |
|  **Data Stewardship** | **Quality and Usage** | No QA/QC procedures defined or implemented and information not available online. | QA/QC procedures defined and available online but not systematically implemented. | Level 2 + data quality assessed and available online such as error or uncertainty estimates; dataset used in national report. | Level 3 + dataset citations in scientific publication in peer-reviewed journal. Used/cited by other well-known climate data centers (NOAA/NCEI, DWD, etc.). |
| **Governance** | Responsibility/accountability compliance mechanism is not defined; No contact information. | Responsibility/accountability compliance mechanism defined for data management operations; User support contact information available. | Level 2 + adheres to data management competency standards and includes written data management policies, procedures, and compliance mechanisms.  | Level 3 + compliance to community standardscontrolled and audited by ISO procedures; or following community best-practice guidance, such as the Data Management Body of Knowledge ([DAMA-DMBOK V2 Guide](https://dama.org/sites/default/files/download/DAMA-DMBOK2-Framework-V2-20140317-FINAL.pdf)) |
| **Metadata** | Collection-level metadata not publicly available and/or not usable. | Limited collection-level metadata available for the users. (e.g. model, network, granular and station metadata such as station latitude/longitude, elevation, instrument type and method) | Level 2 + metadata conforming to community standards and available online(e.g. for observations [WIGOS Metadata Standards](https://library.wmo.int/?lvl=notice_display&id=19925#.Xs0vp2hKhPY)). | Fully compliant with national and international standards; rich metadata content; support dataset provenance.  |

1. **SELF ASSESSMENT TEMPLATE**

A SMM-CD\_NRP self-evaluation template (MS Word format) has been developed to facilitate the assessment process. The use of this template is encouraged to capture evidence when evaluating the stewardship maturity of a dataset. The template can be downloaded from [figshare](https://figshare.com/articles/online_resource/WMO_SMM-CD_NRP_v01r00_20200805_Assessment_Template_docx/13004018).

1. **ACKNOWLEDGEMENT**

The SMM-CD\_NRP was developed by the SMM-TTNI which was formed temporarily specifically for this task by members of the WMO CCl ET-DDS and the International Expert Group on Climate Data Modernisation (IEG-CDM). We thank the ET-DDS members for their contributions, reviews and beneficial comments, Christina Lief (USA, RA IV) Chair, Brigitte Dubuisson (France, RA VI), Ali Eddenjal (Libya, RA I), Lipeng Jiang (China, RA II), Reinaldo Silveira (Brazil, RA III), David Sinclair (Australia, RA V), José-Antonio Guijarro (Spain, RA VI), Xiaolan Wang (Canada, RA IV) as well as ET-DDS associated Experts, Fatima Zahra Hdidou (Morocco, RA I), Ge Peng (USA, RA IV), and Markus Donat (Australia, RA V) and IEG-CDM members, in particular, Markus Ziese (DWD), Robert Dunn (Met Office Hadley Centre), Jean-Francois Legeais (CLS), William Wright (WMO CCl), Peter Siegmund (KNMI), and Omar Baddour (WMO). Because of the present challenging times and travel restrictions, the work was entirely done through video calls meetings and email starting in March 2020. SMM-CD\_NRP version 1 was completed on August 5, 2020.

1. **REFERENCES**
	* CCSDS (The Consultative Committee for Space Data Systems), 2012a: Reference Model for an Open Archival Information System (OAIS), Recommended Practices, Issue 2. Version: CCSDS 650.0-M-2. 135 pp. Available at: [http://public.ccsds.org/publications/archive/650x0m2.pdf](https://library.wmo.int/doc_num.php?explnum_id=10197)
	* CCSDS, 2012b: Audit and certification of trustworthy digital repositories - Recommended Practices. Version: CCSDS 652.0-M-1. 77 pp. Available online at: [https://public.ccsds.org/pubs/652x0m1.pdf](https://library.wmo.int/doc_num.php?explnum_id=10197)
	* CORE-CLIMAX, 2015: European ECV CDR Capacity Assessment Report. COordinating Earth observation data validation for RE-analysis for CLIMAte ServiceS, Grant Agreement No: 313085, Deliverable D2.25, pp 72. Available at: [https://www.eumetsat.int/web- site/wcm/idc/idcplg?IdcService=GET\_FILE&dDocName=PDF\_CORE\_CLIMAX\_CAP\_ASSESS&RevisionSelectionMethod=LatestReleased&Rendition=Web](https://library.wmo.int/doc_num.php?explnum_id=10197)
	* Edmunds, R., H. L’Hours, L. Rickards, P. Trilsbeek, and M. Vardigan, 2016: Core Trustworthy Data Repository Requirements. DOI:10.5281/zenodo.168411 [Available online at: [https://zenodo.org/record/168411 - .WV\_NQBPyuSN]](https://library.wmo.int/doc_num.php?explnum_id=10197)
	* GCOS-143, 2010: Guideline for the Generation of Datasets and Products Meeting GCOS Requirements. GCOS. Geneva, Switzerland. Available at: [https://library.wmo.int/index.php?lvl=notice\_display&id=12884.](https://library.wmo.int/doc_num.php?explnum_id=10197)
	* GCOS-200, 2016: The Global Observing System for Climate: Implementation Needs. GCOS. Geneva, Switzerland. pp. 315, Available at: [https://library.wmo.int/opac/index.php?lvl=no- tice\_display&id=19838.](https://library.wmo.int/doc_num.php?explnum_id=10197)
	* GCOS-206, 2017: Indicators of Climate Change – Outcome of a meeting held at WMO 3 February 2017. GCOS. Geneva, Switzerland. pp. 29. Available at: [https://library.wmo.int/doc\_num.php?explnum\_id=3418](https://library.wmo.int/doc_num.php?explnum_id=10197)
	* ICSU WDS (World Data System, 2016: Core trustworthy data repositories requirements. Version: v01.11 201611. Available at: [https://www.coretrustseal.org/wp-content/up-loads/2017/01/Core\_Trustworthy\_Data\_Repositories\_Requirements\_01\_00.pdf](https://library.wmo.int/doc_num.php?explnum_id=10197)
	* ISO 14721, 2012: Space data and information transfer systems -- Open archival information system (OAIS) – Reference model. Version: ISO 14721:2010. Geneva, Switzerland.
	* ISO 16363, 2012: Space data and information transfer systems -- Audit and certification of trust- worthy digital repositories. Version: ISO 16363:2012. Geneva, Switzerland.
	* ISO 26324, 2012: Information and documentation – Digital object identifier system. Version ISO 26324:2012. Geneva, Switzerland.
	* Lief, C. and G. Peng, 2020: The WMO Stewardship Maturity Matrix for Climate Data (SMM-CD) Template. Document ID: WMO-SMM-CD-0003. Version: v04r00–20200417. Figshare: [doi:10.6084/m9.figshare.7003709](https://library.wmo.int/doc_num.php?explnum_id=10197).
	* Peng, G., J.L. Privette, E.J. Kearns, N.A. Ritchey, and S. Ansari, 2015: A unified framework for measuring stewardship practices applied to digital environmental datasets. Data Science Journal, 13, 231 - 253. [doi:10.2481/dsj.14-049](https://library.wmo.int/doc_num.php?explnum_id=10197).
	* Peng, G., 2015: The scientific data stewardship maturity assessment model template. Version: NCDC-CICS-SMM-0001-Rev.1 v4.0 06/23/2015. [Figshare: doi:10.6084/m9.figshare.1211954](https://library.wmo.int/doc_num.php?explnum_id=10197)
	* Peng, G., 2018: The state of assessing data stewardship maturity – An overview. Data Science Journal. 17, [doi:10.5334/dsj-2018-007](https://library.wmo.int/doc_num.php?explnum_id=10197).
	* Peng, G., J.L. Privette, C. Tilmes, S. Bristol, T. Maycock, J.J. Bates, S. Hausman, O. Brown, and E. J. Kearns, 2018: A Conceptual Enterprise Framework for Managing Scientific Data Steward- ship. Data Science Journal. [doi:10.5334/dsj-2018-015](https://library.wmo.int/doc_num.php?explnum_id=10197).
	* UNFCCC, 2007: Reporting global observing systems for climate. Decision 11/CP.13. Conference of the Parties, Thirteenth Session. Bali, Indonesian, 3–15 December 2007. The final report is available online at: [https://unfccc.int/resource/docs/2007/cop13/eng/06a02.pdf](https://library.wmo.int/doc_num.php?explnum_id=10197)
	* WMO, 2003: Resolution 9 (Cg-XIV) — GCOS Climate Monitoring Principles
	* WMO Climate Data and Monitoring, 2009: Guidelines on analysis of extremes in a changing cli- mate support for informed decisions for adaptation. WCDMP-No. 72/WMO-TD No. 1500, pp. 55, WMO, Geneva, Switzerland. Available at: [https://www.ecad.eu/documents/WCDMP\_72\_TD\_1500\_en\_1.pdf](https://library.wmo.int/doc_num.php?explnum_id=10197)
	* WMO Resolution 25 (Cg-XIII), 1999: Exchange of hydrological data and products. Thirteenth WMO Congress, Geneva, Switzerland, May 1999. Available at: [http://www.wmo.int/pages/prog/hwrp/documents/Resolution\_25.pdf](https://library.wmo.int/doc_num.php?explnum_id=10197)
	* WMO Resolutions (40, 60), 2016: WMO data policy (resolutions 40, 60) and update on CBS discussions. Consultative Meetings on High-Level Policy on Satellite Matters, Thirteenth Session, Geneva, Switzerland, 28–20 January 2016. Available at:

[http://www.wmo.int/pages/prog/sat/meetings/documents/CM-13\_Doc\_03-01\_WMO-Data-Policies.pdf](https://library.wmo.int/doc_num.php?explnum_id=10197)

* + WMO, 2017: WIGOS Metadata Standard. WMO-No. 1192, pp. 46. WMO: Geneva, Switzerland. Available at: [https://library.wmo.int/doc\_num.php?explnum\_id=3653](https://library.wmo.int/doc_num.php?explnum_id=10197)
	+ WMO, 2019: Manual for High Quality Global Data Management Framework for Climate. WMO-No. 1238, pp. 32. WMO: Geneva, Switzerland. Available at: [https://library.wmo.int/doc\_num.php?explnum\_id=10197](https://library.wmo.int/doc_num.php?explnum_id=10197%20)
1. **DEFINITIONS**
	* **Data Access** refers to the ability to locate (Discoverability) and get (Accessibility) the dataset in question.
	* **Data Documentation** will ensure that your data will be understood and interpreted by any user. It will explain how your data was created, what the context is for the data, structure of the data and its contents.
	* **Data Governance** includes standard procedure, policies, approval process, along with accountabilities and compliance mechanism for ensuring the data are secure, accessible, and useable.
	* **Data Integrity** refers to the extent to which data are recorded, preserved and used exactly as intended, and that data are free from corruption or loss when transferred between systems or in storage throughout the data life cycle. It is a critical aspect to the design, implementation and usage of any system which stores, processes, or retrieves data. Data integrity is the opposite of data corruption.
	* **Data Management** is the set of operations, procedures, protocols, and policies required to organize, archive, quality control, secure and enable access to an organization’s data holdings.
	* **Data Portability** is a concept to protect users from having their data stored in closed platforms that are incompatible with one another. Data portability requires common technical standards to facilitate the transfer from one data controller to another, thus promoting interoperability.
	* **Data Preservation** means ensuring data remains accessible and usable for as long as it is required for operational, research, business evidentiary or historical purposes. It includes securing the data and making provision for obsolescence of storage media used to store the data; the hardware used to access the data; and the software and hardware required to access the data.
	* **Data Quality Assurance** refers to the processes for maintaining a desired level of quality in a dataset or collection. Data verification, quality control and validation are important steps in supporting defensible products and decisions.
	* **Data Quality Control** is the process of ensuring that errors in the data are detected, flagged, and corrected. It involves checking the data to assess representativeness in time, space, and internal consistency, and flagging any potential inconsistencies. The purpose of Quality Control (and broader Quality Management) is to ensure that meteorological and climate data available to potential users is sufficiently reliable to be used with confidence.
	* **Data Quality Assessment** is the process of scientifically and statistically evaluating datasets and their level of stewardship, to determine whether they are of sufficient quality to reliably support their intended use.
	* **Data Quality Management** is the process of overseeing the activities, tasks and policies required to ensure that data maintain a required standard of excellence. Quality Management involves quality planning, the establishment and continued operation of a quality assurance system, including adequate quality control, and quality assessment and improvement processes.
	* **Data Uncertainty** is a measure of noise that deviates from the correct, intended or original values. All measurements of an observed phenomenon have a degree of uncertainty regardless of precision and accuracy. This is caused by two factors, the limitation of the measuring instrument (systematic error) and the skill of the observer making the measurements (random error). Further uncertainty can arise when, for instance, values are rounded, interpolated, or extrapolated.
	* **Data Usage** measures how much a dataset has been utilized based on relevant scientific literature. Note that citations refer not just to citations in peer-reviewed journals, but also in widely accepted, authoritative institutional reports.
	* **Metadata** is information about data and sometimes referred to as “data about data”. It is important to distinguish between a number of different types of metadata, as described below. To ensure that data are fit for purpose for climate services and research, Entities which produce data for climate purposes are required to create and maintain all of the following types of metadata.
	* **Metadata, Collection-level** is metadata describing a data collection. A collection is a grouping of environmental data that share common characteristics, which often time refers to a dataset. A dataset may have one or many data files or contain one or many variables.
	* **Metadata, Contextual** is information about how the data were collected or generated, featuring the who, how, when and where a measurement was made. This information is required to establish fitness for purpose, as well as providing indispensable information for operations such as homogenisation. In the case of meteorological data, it includes such details as where and when the measurement was made, with what instrumentation, by whom, under what siting conditions, what changes to the above have occurred, quality control status, intellectual property information. If the data/information were created by processing or analysis methods, details of the algorithms and methodology used are also required.
	* **Metadata, Discovery** is metadata which enables a user to query or search a catalogue to determine what information is held, where it is held and by whom, along with some details about the data/information set. There is a considerable body of knowledge about the requirements for such metadata, with the internationally accepted standard for what metadata should be maintained referred to as ISO 19115.
	* **Metadata, Granular-level** is metadata describing the smallest aggregation of data that can be independently managed (described, inventoried, and retrieved), for example, for an individual variable or station.
	* **Metadata, Network**. Changes to the way climate variables are measured apply not only at the individual station level, but to whole networks of stations. An example might be when manual observations are replaced by Automatic Weather Stations (AWS), or when a network of AWSs are progressively replaced by a model with a different central processing unit, or when new sensors are introduced. It is important again to document the time, location, and details of any such changes. Moreover, to support the effective homogenisation simultaneous changes across an entire network should be avoided.
	* **Metadata, Provenance.** Apart from the need to know what changes to observation siting, practices, etc. have been made over time (an essential step in homogenisation procedures), it is important to know about changes to the versions of a dataset. This is because of the need for traceability – being able to identify the version of a dataset from which an analysis or product was derived. Provenance should therefore include details of any quality control or homogenisation processes, details of disaggregation or infilling, or any other changes made to the dataset. Climate products and services need to contain a link to the version of the data on which they are based.
	* **Metadata, Rich:** Refers to how detailed and complete contextual metadata has been captured.
2. **ACRONYMS**

AOPC Atmospheric Observation Panel for Climate

BOM The Bureau of Meteorology

CCl Commission for Climatology

CDR Climate Data Record

CEOS Committee on Earth Observation Satellites

CF Climate and Forecast

CICS-NC NOAA’s Cooperative Institute for Climate and Satellites – North Carolina

CISESS NOAA’s Cooperative Institute for Satellite Earth System Studies

CORE-CLIMAX European Union Framework 7 Project in the context of Climate Services (www.coreclimax.eu)

CSIS Climate Services Information System

DSMM Data Stewardship Maturity Matrix

DWD The Deutscher Wetterdienst

ECMWF European Centre for Medium-range Weather Forecasts

ECV Essential Climate Variable

ESA European Space Agency

EUMETSAT European Organisation for the Exploitation of Meteorological Satellites

GFCS Global Framework for Climate Services

GCOS Global Climate Observing System

GPCC Global Precipitation Climatology Centre

HQ-GDMFC High-Quality Global Data Management Framework for Climate

IEG-CDM International Expert Group for Climate Data Modernization

ICOADS International Comprehensive Ocean-Atmosphere Data Set

IPCC Intergovernmental Panel on Climate Change

ISO International Organization for Standardization

JAXA Japan Aerospace Exploration Agency

JMA Japanese Meteorological Agency

KNMI Royal Netherlands Meteorological Institute

LEGOS Laboratoire d’Etudes en Géophysique et Océanographie Spatiales

Met Office United Kingdom's national weather service

NARA National Archives and Records Administration

NASA National Aeronautics and Space Administration

NCEI National Centers for Environmental Information

NCSU North Carolina State University

NMHS National Meteorological and Hydrological Services

NOAA National Oceanic and Atmospheric Administration

NSIDC National Snow and Ice Data Center

OAIS Open Archival Information System

OAIS RM Open Archival Information System Reference Model

OOPC Ocean Observing Panel for Climate

RCC Regional Climate Center

SMM-CD Stewardship Maturity Matrix for Climate Data

SMM-CD\_NRP Stewardship Maturity Matrix for Climate Data for National and Regional Purposes

TDS THREDDS Data Server

THREDDS Thematic Real-time Environmental Distributed Data Services

TOPC Terrestrial Observation Panel for Climate

UNFCCC United Nations Framework Convention on Climate Change

USGCRP U.S. Global Change Research Program

WCRP World Climate Research Programme

WGISS CEOS Working Group on Information Systems and Services

WIGOS WMO Integrated Global Observing System

WMO World Meteorological Organization

WIS WMO Information System

1. The 18th WMO Congress (June 2019) established the restructuring of WMO, including the Commission for Climatology, into the Infrastructure and Application Commissions and a Research Board. [↑](#footnote-ref-1)