

Cg-18/Doc. 5.2(2), APPROVED

Draft Manual

High Quality Global Data Management Framework for Climate

2019 Edition

EDITORIAL NOTE

The following typographical practice has been followed: standard practices and procedures have been printed in **bold**. Recommended practices and procedures have been printed in regular font. Notes have been printed in smaller type.

METEOTERM, the WMO terminology database, may be consulted at <http://public.wmo.int/en/resources/meteoterm>.

Readers who copy hyperlinks by selecting them in the text should be aware that there will be additional spaces immediately following <http://>, <https://>, <ftp://>, <mailto:>, and after slashes (/), dashes (-) and unbroken sequences of characters (letters and numbers). These spaces should be removed from the pasted URL. The correct URL is displayed when hovering over the link or when clicking on the link and then copying it from the browser.

WMO-No. XXXX

© World Meteorological Organization, 2019

The right of publication in print, electronic and any other form and in any language is reserved by WMO. Short extracts from WMO publications may be reproduced without authorization, provided that the complete source is clearly indicated. Editorial correspondence and requests to publish, reproduce or translate this publication in part or in whole should be addressed to:

Chairperson, Publications Board
World Meteorological Organization (WMO)
7 bis, avenue de la Paix, P.O. Box 2300,
CH-1211 Geneva 2, Switzerland E-mail: publications@wmo.int
Tel.: +41 (0) 22 730 84 03
Fax: +41 (0) 22 730 81 17
ISBN xxx-xx-xx-xxxxx-x

NOTE:

The designations employed in WMO publications and the presentation of material in this publication do not imply the expression of any opinion whatsoever on the part of WMO concerning the legal status of any country, territory, city or area, or of its authorities, or concerning the delimitation of its frontiers or boundaries. The mention of specific companies or products does not imply that they are endorsed or recommended by WMO in preference to others of a similar nature which are not mentioned or advertised.

CONTENT

INTRODUCTION

GENERAL PROVISIONS

PREAMBLE

PRINCIPLES

ENABLING MECHANISMS

PART I GENERAL DEFINITIONS, DATA SOURCES AND PRACTICES

1.1 DEFINITIONS

1.2 MANAGING DATA FROM VARIOUS SOURCES

1.2.1 GENERIC CLIMATE REQUIREMENTS FOR OBSERVATIONS AND DATA MANAGEMENT

1.2.2 CONVENTIONAL IN-SITU SOURCES OF CLIMATE DATA

1.2.3 AUTOMATIC WEATHER STATION DATA

1.2.4 REMOTE SENSING DATA

1.2.5 SPACE WEATHER DATA

1.2.6 AIRCRAFT DATA

1.2.7 OCEAN AND MARINE DATA

1.2.8 ATMOSPHERIC COMPOSITION DATA

1.2.9 HYDROLOGY DATA

1.2.10 CLIMATE MODEL DATA

1.2.11 EXTERNALLY-SOURCED DATA

1.2.12 SOCIO-ECONOMIC DATA

PART II REQUIREMENTS FOR CLIMATE DATA STEWARDSHIP

2.1. DATA MANAGEMENT REQUIREMENTS

2.1.1 GENERAL REQUIREMENTS

2.1.2 DISCOVERABILITY AND ACCESSIBILITY

2.1.3 USABILITY

2.1.4 QUALITY MANAGEMENT

2.1.5 STEWARDSHIP

2.1.6 DATA MANAGEMENT GOVERNANCE

2.1.7 CLIMATE DATA MANAGEMENT COMPETENCIES

2.1.8 ONGOING SKILL DEVELOPMENT

2.1.9 PLANNING AND DESIGN

2.2. DATASET MATURITY ASSESSMENT

2.2.1 PURPOSE

2.2.2 USE OF WMO MATURITY MATRIX FOR CLIMATE DATASETS

2.2.3 WMO CLIMATE DATA CATALOGUE

INTRODUCTION

The High Quality Global Data Management Framework for Climate (HQ-GDMFC) is a WMO collaborative initiative that enables the effective development and exchange of high quality climate data, based on a reliable integrated underpinning data infrastructure at the global, regional and national levels. The Framework establishes standards and recommended practices for sourcing, securing, managing, assessing, and cataloguing climate data, and for sharing infrastructure and responsibilities for, e.g., data exchange, analysis and data service provision. In this way the HQ-GDMFC provides a robust data foundation for the generation of climate products and the delivery of climate services through the Climate Services Information System (CSIS) of the Global Framework for Climate Services (GFCS).

An important outcome of the Framework is a Manual on HQ-GDMFC, which forms part of the WMO Technical Regulations. This represents the first time that the climate community has defined such regulatory requirements through WMO specifically for the management of climate data.

Collaboration within the Framework expands to several application areas such as Marine and Hydrological science, many of whose data-related activities should also be regarded as falling under the definition of climate data, to address their needs and requirements as generators of climate data, as well as for sharing related data for supporting national and international climate activities. The HQ-GDMFC data scope encompasses all of the Essential Climate Variables under the WMO auspices, as described in Resolution 60 (Cg-17). This includes observational data as well as data derived from climate analysis, reanalysis, prediction and projection. The procedures provided are also applicable to externally-sourced data and data on socio-economic impacts outside of the WMO auspices.

The HQ-GDMFC leverages relevant WMO systems such as the WMO Information System (WIS) for data exchange and data service sharing, the Global Climate Observing System (GCOS), the WMO Integrated Global Observing System (WIGOS) for observational standards and practices, and the global Data Processing and Forecasting System (GDPFS) for Model Data used in operational activities such as weather and climate analysis and forecasting. The framework of collaboration includes primarily the National Meteorological and Hydrological Services' Data Management units, Regional Climate Centers, international data centers, climate research bodies, certain Government agencies, academia and any other institution dealing with climate data archival, management, analysis and exchange.

The manual on HQ-GDMFC provides guidance and requirements on the development, provision, exchange and maintenance of high-quality climate datasets. The standards and recommended practices it describes are intended to ensure that the data made available for climate assessment, monitoring, applications and related services meet sustainably a minimum set of requirements with regard to quality, governance, accessibility and usability.

GENERAL PROVISIONS

1. The *Technical Regulations* (WMO-No. 49) of the World Meteorological Organization are presented in four volumes:

- Volume I – General meteorological standards and recommended practices
- Volume II – Meteorological service for international air navigation
- Volume III – Hydrology
- Volume IV – Quality management.

Purpose of the Technical Regulations

2. The Technical Regulations are determined by the World Meteorological Congress in accordance with Article 8 (d) of the Convention.

3. These Regulations are designed:

- (a) To facilitate cooperation in meteorology and hydrology among Members;
- (b) To meet, in the most effective manner, specific needs in the various fields of application of meteorology and operational hydrology in the international sphere;
- (c) To ensure adequate uniformity and standardization in the practices and procedures employed in achieving (a) and (b) above.

Types of Regulations

4. The Technical Regulations comprise *standard* practices and procedures and *recommended* practices and procedures.

5. The definitions of these two types of Regulations are as follows:

The *standard* practices and procedures:

- (a) Shall be the practices and procedures that Members are required to follow or implement;
- (b) Shall have the status of requirements in a technical resolution in respect of which Article 9(b) of the Convention is applicable;
- (c) Shall invariably be distinguished by the use of the term *shall* in the English text, and by suitable equivalent terms in the Arabic, Chinese, French, Russian and Spanish texts.

The *recommended* practices and procedures:

- (a) Shall be the practices and procedures with which Members are urged to comply;
- (b) Shall have the status of recommendations to Members, to which Article 9 (b) of the Convention shall not be applied;
- (c) Shall be distinguished by the use of the term *should* in the English text (except where otherwise provided by decision of Congress) and by suitable equivalent terms in the Arabic, Chinese, French, Russian and Spanish texts.

6. In accordance with the above definitions, Members shall do their utmost to implement the *standard* practices and procedures. In accordance with Article 9 (b) of the Convention and in conformity with Regulation 128 of the General Regulations, Members shall formally notify the Secretary-General, in writing, of their intention to apply the *standard* practices and procedures of the Technical Regulations, except those for which they have lodged a specific request for deviation. Members shall also inform the Secretary-General, at least three months in advance, of any change in the degree of their implementation of a *standard* practice or procedure as previously notified, and the effective date of the change.

7. Members are urged to comply with *recommended* practices and procedures, but it is not necessary to notify the Secretary-General of non-observance except with regard to practices and procedures contained in Volume II.

8. In order to clarify the status of the various Regulations, the *standard* practices and procedures are distinguished from the *recommended* practices and procedures by a difference in typographical practice, as indicated in the editorial note.

Status of annexes and appendices

9. The following annexes to the *Technical Regulations* (Volumes I to IV), also called Manuals, are published separately and contain regulatory material having the status of *standard* and/or *recommended* practices and procedures:

I *International Cloud Atlas* (WMO-No. 407), Volume I – Manual on the Observation of Clouds and Other Meteors, Part I; Part II: paragraphs II.1.1, II.1.4, II.1.5 and II.2.3; subparagraphs 1, 2, 3 and 4 of each paragraphs from II.3.1 to II.3.10; paragraphs II.8.2 and II.8.4; Part III: paragraph III.1 and the definitions (in italics) of paragraph III.2;

II *Manual on Codes* (WMO-No. 306), Volume I;

III *Manual on the Global Telecommunication System* (WMO-No. 386);

IV *Manual on the Global Data-processing and Forecasting System* (WMO-No. 485), Volume I;

V *Manual on the Global Observing System* (WMO-No. 544), Volume I;

VI *Manual on Marine Meteorological Services* (WMO-No. 558), Volume I;

VII *Manual on the WMO Information System* (WMO-No. 1060);

VIII *Manual on the WMO Integrated Global Observing System* (WMO-No. 1160).

These annexes (Manuals) are established by decision of Congress and are intended to facilitate the application of Technical Regulations to specific fields. Annexes may contain both *standard* and *recommended* practices and procedures.

10. Texts called appendices, appearing in the *Technical Regulations* or in an annex to the *Technical Regulations* have the same status as the Regulations to which they refer.

Status of notes and attachments

11. Certain notes (preceded by the indication "Note") are included in the *Technical Regulations* for explanatory purposes; they may, for instance, refer to relevant WMO Guides and publications. These notes do not have the status of Technical Regulations.

12. The *Technical Regulations* may also include attachments, which usually contain detailed guidelines related to *standard* and *recommended* practices and procedures. Attachments, however, do not have regulatory status.

Updating of the Technical Regulations and their annexes (Manuals)

13. The *Technical Regulations* are updated, as necessary, in the light of developments in meteorology and hydrology and related techniques, and in the application of meteorology and operational hydrology. Certain principles previously agreed upon by Congress and applied in the selection of material for inclusion in the Technical Regulations, are reproduced below. These principles provide guidance for constituent bodies, in particular technical commissions, when dealing with matters pertaining to the Technical Regulations:

- (a) Technical commissions should not recommend that a Regulation be a *standard* practice unless it is supported by a strong majority;
- (b) Technical Regulations should contain appropriate instructions to Members regarding implementation of the provision in question;

- (c) No major changes should be made to the Technical Regulations without consulting the appropriate technical commissions;
- (d) Any amendments to the Technical Regulations submitted by Members or by constituent bodies should be communicated to all Members at least three months before they are submitted to Congress.

14. Amendments to the *Technical Regulations* – as a rule – are approved by Congress.

15. If a recommendation for an amendment is made by a session of the appropriate technical commission, and if the new regulation needs to be implemented before the next session of Congress, the Executive Council may, on behalf of the Organization, approve the amendment in accordance with Article 14 (c) of the Convention. Amendments to annexes to the *Technical Regulations* proposed by the appropriate technical commissions are normally approved by the Executive Council.

16. If a recommendation for an amendment is made by the appropriate technical commission and the implementation of the new regulation is urgent, the President of the Organization may, on behalf of the Executive Council, take action as provided by Regulation 9(5) of the General Regulations.

Note: A simple (fast-track) procedure may be used for amendments to technical specifications in Annexes II (*Manual on Codes* (WMO-No. 306)), III (*Manual on the Global Telecommunication System* (WMO-No. 386)), IV (*Manual on the Global Data-processing and Forecasting System* (WMO-No. 485)), V (*Manual on the Global Observing System* (WMO-No. 544)), VII (*Manual on the WMO Information System* (WMO-No. 1060)) and VIII (*Manual on the WMO Integrated Global Observing System* (WMO-No. 1160)).

17. After each session of Congress, a new edition of the *Technical Regulations*, including the amendments approved by Congress, is issued. With regard to the amendments between sessions of Congress, Volumes I, III and IV of the *Technical Regulations* are updated, as necessary, upon approval of changes thereto by the Executive Council. The *Technical Regulations* updated as a result of an approved amendment by the Executive Council are considered a new update of the current edition. The material in Volume II is prepared by the World Meteorological Organization and the International Civil Aviation Organization working in close cooperation, in accordance with the Working Arrangements agreed by these Organizations. In order to ensure consistency between Volume II and Annex 3 to the Convention on International Civil Aviation – *Meteorological Service for International Air Navigation*, the issuance of amendments to Volume II is synchronized with the respective amendments to Annex 3 by the International Civil Aviation Organization.

Note: Editions are identified by the year of the respective session of Congress whereas updates are identified by the year of approval by the Executive Council, for example "Updated in 2012".

WMO Guides

18. In addition to the *Technical Regulations*, appropriate Guides are published by the Organization. They describe practices, procedures and specifications which Members are invited to follow or implement in establishing and conducting their arrangements for compliance with the Technical Regulations, and in otherwise developing meteorological and hydrological services in their respective countries. The Guides are updated, as necessary, in the light of scientific and technological developments in hydrometeorology, climatology and their applications. The technical commissions are responsible for the selection of material to be included in the Guides. These Guides and their subsequent amendments shall be considered by the Executive Council.

PREAMBLE

With the increasing demand for High Quality data in support of climate services, disaster risk reduction and climate change adaptation and mitigation, it is important to ensure that a robust regulatory framework that defines standard and recommended practices and procedures for management of the data is defined.

The purpose of this Manual, which forms part of the WMO Technical Regulations, is to set out the high level standards and recommended practices for data stewardship that will ensure that data for climate purposes is reliable, accessible, sustainable, and as far as possible complete. Part I of the Manual provides a terminology of commonly-encountered definitions with regard to climate data, as well as general requirements for the management of data, along with more specific considerations of data from different sources. Part II provides greater detail on the required standards and recommended practices for managing climate data, and also introduces the concept of dataset maturity assessment as an objective tool for evaluating the level of compliance with these standards and recommended practices.

The provisions in this manual are based on universally agreed data management goals. Several attempts have been made to define such high-level goals, including those associated with data management principles of the Group on Earth Observation System of Systems (GEOSS). These goals relate essentially to the 12 aspects of the Stewardship Maturity Matrix for Climate Data (Section 2.2), which are as follows:

1. **Discoverability:** ensuring that data and relevant information about the dataset may readily be found, including visibility in online catalogues;
2. **Accessibility:** In the context of data management, accessibility refers to ensuring that the dataset is easily and conveniently downloadable by users;
3. **Portability:** ensuring that data are easily incorporated into users' working environment based on community standards;
4. **Documentation:** including all elements necessary to access, provide guidance for users, understand, and process the data, and including documentation on replication, reprocessing and updates;
5. **Usability and Usage:** Usability describes how easily the data product may be understood and used by users and incorporated into the users own working environment. It incorporates aspects of compatibility of the publication medium with community standards and supporting documentation. Usage refers to the degree of scientific credibility of the data among users, including through citations in peer-reviewed literature;
6. **Data quality control:** using community best practices, including ensuring that the results of quality control are also documented;
7. **Quality assessment of datasets:** ensuring that datasets and data products are routinely reviewed, and the results of the review are transparent;
8. **Uncertainty:** ensuring that uncertainty estimates are documented and made available;
9. **Data integrity:** ensuring that data are recorded, preserved and are free from corruption or loss when transferred between systems or in storage throughout the data life-cycle;
10. **Preservation:** ensuring that data and metadata are protected from loss and preserved for future use in line with well-specified retention policies;
11. **Metadata:** ensuring that information about data are publicly available, including full details of the origin and processing history of raw observations and derived products (provenance metadata), to ensure full traceability of the processing chain;
12. **Governance:** ensuring that accountability, responsibility and compliance mechanisms are well-defined and transparent.

PRINCIPLES

International collaboration within the HQ-GDMFC shall be based on the following principles:

1. Promoting adherence to the WMO relevant data policies including the provisions in Resolution 40 (Cg-XII) on WMO policy and practice for the exchange of meteorological and related data and products, Resolution 25 (Cg-XIII) on Exchange of Hydrological Data and Products, and Resolution 60 (Cg-17) on WMO policy for the international exchange of climate data and products to support the implementation of the Global Framework for Climate Services (GFCS);
2. Collaborating on registering datasets to be shared internationally through the WMO information System for use in climate studies, monitoring and applications;
3. Facilitating easy access to metadata and documentation underpinning the datasets;
4. Promoting preservation and management of all data that are used or may potentially be useful for climate change monitoring in duplicate repositories for the duration of their specified retention periods;
5. Collaborating on assessing and improving the maturity and quality of stewardship practices underpinning the datasets, cataloguing them for easy search, discovery and access, and promoting their use in informing policy relevant frameworks;
6. Promoting acquisition of user feedback on the quality, fitness for purpose and usability of shared datasets.

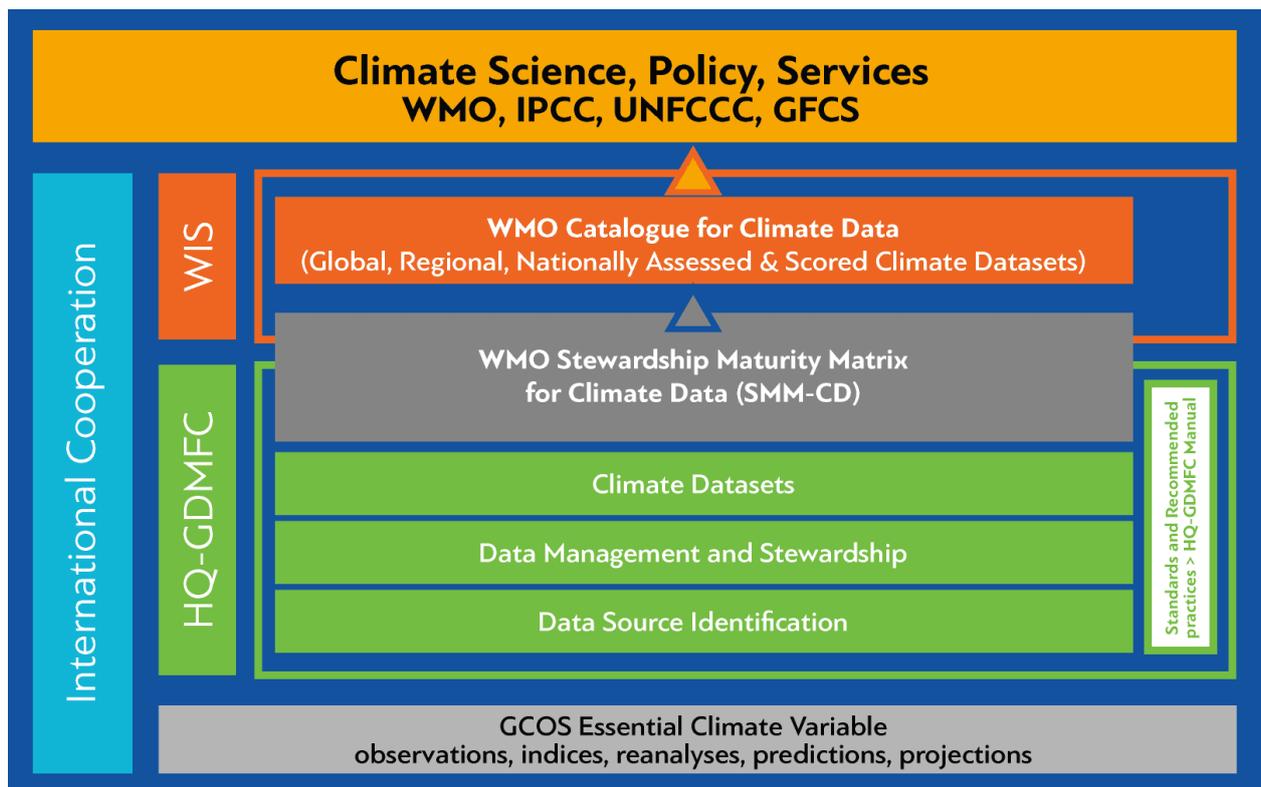


Figure 1: Architecture of HQ-GDMFC building on GCOS ECV science pillar, Data Management Standards, international cooperation. The WMO Information System (WIS) is used for registering climate data and ensures that maturity assessed data are promptly discoverable and made available for use in climate science, policy and services.

ENABLING MECHANISMS

Collaborating Entities

A Collaborating Entity is an organization that contributes in some way to climate data management which supports the goals and principles of the HQ-GDMFC. These organizations are referred to in this manual as an Entity in singular form or Entities in plural form. Entities fall into one of two types:

Type 1: Organizations that are contributing to one or several WMO programmes and/or co-sponsored programme activities, and have a mandate to collect, archive and process climate data, and generate a range of products and analyses based on the data; these Collaborating Entities could have a global, regional or national mandate that is recognized by a governmental or inter-governmental body. At national level these include National Meteorological and Hydrological Services (NMHSs), national archiving authorities, as well as climate related research institutes operating outside of the NMHS such as universities, cooperative research institutes etc.

Or,

Type 2: Collaborating Entities (public, private, governmental or non-governmental) outside the direct WMO auspices, which develop and/or maintain climate datasets and apply scientific methods for generating data products, such as by using statistical analysis or modelling approaches. These include most international data centres and space agencies, as well as government agencies with a climate or climate change focus, such as agricultural or water management agencies, some climate change authorities, and libraries.

Note: Some international data centres may also be classed as Type 1 entities.

The willingness of an organization to become a Collaborating Entity in HQ-GDMFC shall be notified with a simple letter to the WMO Secretary General. The letter shall be sent through the Permanent Representative of the Member State with WMO where the Collaborating Entity is located. The letter should express the Entity's willingness to adhere to the HQ-GDMFC principles and contribute to providing access to the datasets which it generates, along with the necessary documentation for producing, accessing and using them. If the Entity is an intergovernmental organization, the letter should be sent by the head of that organization.

Noting the importance of retaining and providing access to climate data from all possible sources, WMO Members through their Permanent Representatives should encourage NMHSs and all non-NMHS bodies as described under Types 1 and 2 above to register as Entities; or alternatively, to encourage them to provide their data to secure archiving facilities such as national archive centres, the NMHS itself, or international data centres.

The WMO Information System (WIS)

WIS responsibilities and functions are described in Annex VII to the WMO Technical Regulations. The WIS provides an overarching approach to data and information management for all WMO and related international programmes, leveraging the collaborative culture of WMO, as well as new technologies.

HQ-GDMFC builds on WIS infrastructure (Global Information System Centres (GISCs), Data Collection and Producing Centres (DCPC) and National Centers), as well as the Global Telecommunication System (GTS); and standards (metadata core profile, codes, services) for data collection and exchange, discovery, access and retrieval. Opportunities exist to improve efficiencies through emerging technological innovations. In order to fully benefit from the WIS, a climate data management infrastructure within an Entity contributing to HQ-GDMFC could be designated as part of the WIS infrastructure, in particular as a DCPC.

The WMO Observing Systems and Co-sponsored Systems

HQ-GDMFC shall leverage on data and related standards provided by WIGOS (Annex VIII to WMO Technical Regulations) observing capabilities (surface- and space-based components), including the Global Observing System of the World Weather Watch Programme, the observing component of the Global Atmosphere Watch Programme, the WMO Hydrological Observing System of the Hydrology and Water Resources Programme, and the observing component of the Global Cryosphere Watch. The Global Climate Observing System (GCOS) Essential Climate Variables concept and the GCOS Climate Monitoring Principles provide a strong scientific foundation for enabling high quality long-term datasets and their scientific integrity.

The WMO Global Data Processing and Forecasting Systems (GDPFS)

HQ-GDMFC shall make use of the GDPFS infrastructure and standards (Annex IV to WMO Technical Regulations) that are applied in the domain of climate data, monitoring and predictions. Regional Meteorological Specialized Centres (RSMCs), as components of the GDPFS, contribute to the framework by developing quality-controlled datasets, and providing or facilitating database and archiving services, as well as other highly recommended services such as data rescue and capacity building.

PART I GENERAL DEFINITIONS, DATA SOURCES AND PRACTICES

Note: The list of definitions as provided below is a living list that should be maintained and updated as science and applications and user requirements evolves. The focus is made on most needed definitions for critical aspect of data management for climate purpose.

1.1 DEFINITIONS

- 1.1.1 A **Climate archive** is a collection of archived climate records (cf. definition of *Climate record*), and also includes derived products such as analyses, summaries, maps, graphs, technical documentation, publications, and books. A Climate Archive can be comprised of digital and hard media records (maps, paper-based field observations, instrumentation tapes, etc.)
- 1.1.2 **Climate Change** (IPCC definition) Climate change refers to a change in the state of the climate that can be identified (e.g., by using statistical tests) by changes in the mean and/or the variability of its properties and that persists for an extended period, typically decades or longer. Climate change may be due to natural internal processes or external forcings such as modulations of the solar cycles, volcanic eruptions and persistent anthropogenic changes in the composition of the atmosphere or in land use.
- 1.1.3 **Climate data** encompasses long-term observational data (Essential Climate Variables (ECVs) and other related parameters such as evaporation, visibility, phenomena, etc.), variables derived from observations (such as mean sea-level pressure), as well as data products such as gridded data and numerical model (re)analyses, predictions and projections.
- 1.1.4 A **Climate Data Management System** (CDMS) is a system that facilitates the effective archival, management, analysis, delivery and utilization of a wide range of climate data. According to the WMO *Climate Data Management System Specifications* (WMO-No 1131), a CDMS combines both IT and governance components. It utilizes a set of inter-related modules that perform specific functions with the data. The purpose of a CDMS is to assist a Member, often through its National Meteorological Service, to manage its climate record securely and sustainably. Guidance on CDMS is provided in the publication (WMO-No. 1131).
- 1.1.5 **Climate Index** is a simple diagnosis quantity that is used to characterize an aspect of a geophysical system such as a circulation pattern. For example, to gain a uniform perspective on observed changes in weather and climate extremes, the joint CCI/WCRP-Clivar/JCOMM Expert Team on Climate Change Detection and Indices (ET-CCDI) has defined a core set of descriptive indices of extremes. The core set includes 27 extremes indices for temperature and precipitation (WMO, 2009), ranging from the frequency of occurrence above and below particular thresholds, warm/cold and dry/wet spell duration, among many others .
- 1.1.6 **Climate Indicator** is a clearly-defined indicator of climate change. For example, the WMO Global Climate Observing System (GCOS) has identified a core set of seven global climate indicators to be used as a basis for describing climate and for reporting climate change to the public (GCOS-206, 2017), these being surface air temperature, ocean heat content, atmospheric CO₂ content, ocean acidification, sea-level, glaciers, Arctic and Antarctic sea-ice extent.
- 1.1.7 **Climate information** is information derived from processing and analysing climate data and products. For example, a summary report on an extreme climate event at a given location; a synthesis of seasonal forecast products; a report on climate projections. Climate information is used in various climate applications for supporting a domain of activity such as agriculture, health, water management, energy, and transport.

- 1.1.8 **CLIMAT Messages**, sometimes referred to as CLIMAT Reports. CLIMAT and CLIMAT SHIP messages are coded messages that summarize and report monthly values of meteorological parameters from land and marine observing networks respectively. *Guidance on CLIMAT reporting is provided in the WMO handbook on CLIMAT and CLIMAT Temp Reporting, 2009 edition, (WMO/TD-No. 1188).* It should be noted that the monthly upper-air CLIMAT Temp message was discontinued in 2012 by the *Commission for Climatology in 2010 (WMO-No. 1054)* and by *WMO Executive Council – Sixty-fourth session (EC-LXIV 2012 - WMO-No. 1092).*
- 1.1.9 **Climate model.** A numerical representation of the *climate system* based on the physical, chemical and biological properties of its components, their interactions and *feedback* processes, which account for some of its known and observed properties. Climate models are applied as a research tool to study and simulate the Climate and for operational purposes, including monthly, seasonal, interannual and decadal climate predictions, as well as climate change projections.
- 1.1.10 **Climate predictions.** Computed future evolution of the climate system using deterministic climate model or statistical methods from a starting condition. Climate predictions can cover monthly, seasonal, interannual and decadal time scales.
- 1.1.11 **Climate products** are packaged information that includes data, summaries, tables, graphs, maps, reports and analyses (*Guide to Climatological Practices, WMO-No. 100*).
- 1.1.12 **Climate projection.** (IPCC definition) A climate projection is the simulated response of the climate system to a scenario of future emission or concentration of greenhouse gases and aerosols, generally derived using climate models. Climate projections are distinguished from climate predictions by their dependence on the emission/concentration/radiative forcing scenario used, which is in turn based on assumptions concerning, for example, future socio-economic and technological developments that may or may not be realized.
- 1.1.13 **Climate record** is a set of long-term observations with timeseries components covering any or all of the atmosphere, ocean, land and land water, cryosphere and space weather domains, with associated metadata, recorded according to standards, and archived for climate purpose. It includes observations recorded on paper, as well as scanned and digitized records. The metadata documents, at the minimum, the what, who, when, where and how an observation was made.
- 1.1.14 **Climate Service** Climate services provide climate information, products or activities that facilitate decision-making by individuals and organizations. "Service" is defined as the delivered product and the activities associated with the people, process and information technology required to deliver it. This definition and additional guidance is provided in the WMO Guide to Climatological Practices (WMO-No.100) The value chain for a climate service includes the development of climate data, information and products, through to their delivery to the users.
- 1.1.15 **Climate simulation** a description generated by climate models of the past/current and/or future based on a coherent and internally consistent set of assumptions about key driving forces and relationships.
- 1.1.16 **Climatological Standard Normals** are latest 30-year averages of climatological data updated every 10 years, commencing in a year ending with digit 1, i.e. the averages over for the following consecutive periods 1 January 1981 – 31 December 2010, 1 January 1991 –31 December 2020, and so forth. Additional guidance on the computational aspects of Climate Normals is provided in the *WMO Guidelines on the Calculation of Climate Normals, 2017 edition (WMO-No. 1203).*

Note: (1) If data are not continuous, or the time-series is of insufficient length, provisional Normals may be computed. (2) See also the definition of the Reference Period for Climate Change Monitoring at 1.1.63

- 1.1.17 **Data Access** refers to the ability to locate (Discoverability), and get (Accessibility), the dataset in question.
- 1.1.18 **Data Analysis** is the process of inspecting, cleansing, transforming, and modelling data with the goal of generating useful information. It is applied to investigating climate phenomena and processes, climate monitoring, and supporting decision-making in climate-sensitive applications.
- 1.1.19 **Data Archival** refers to activities that support the long-term secure storage of data and metadata, and the methods used to read or interpret them. It is part of the overall process of data preservation and applies to both hard copy records that need to be physically stored in a secure environment, and digital records.
- 1.1.20 **Data Archives- World Weather Records (WWR)** is an archive of digital meteorological data from locations around the world that include monthly mean values of pressure, temperature, precipitation, and where available, station metadata notes documenting observation practices and station configurations. World Weather Records are compiled on an annual basis (until 2010 it was a decadal basis), based on reports from Member countries according to a specific template.

Note: The WWR web site is maintained in collaboration with the World Data Centre for Meteorology at the National Centers for Environmental Information (NCEI), Asheville, USA.

- 1.1.21 A **Data Collection** consists of a set of targeted variables measured, gathered and accumulated in an established systematic fashion. For instance, a database containing temperature, rainfall and evaporation data could be referred to as a collection of atmospheric climate data.
- 1.1.22 **Data, Crowd- sourced** is a form of *Externally-sourced data* typically provided by interested companies, groups, or members of the public, such as Storm Chasers. Such sources may utilize custom-built systems. Data from these sources are generally understood to be less structured than data from National Meteorological and Hydrological Service (NMHS) observing systems or from contracted external partners, with standards of data quality, metadata and continuity of supply generally not prescribed and less well known. For these reasons they should be incorporated with great care into climate products and analyses.
- 1.1.23 **Data documentation** will ensure that your data will be understood and interpreted by users. It should explain how the data were created, and its context, structure and content.
- 1.1.24 **Data Exchange / Data Sharing** is the process of making available data collected or generated by an Entity, so it can be accessed and used by other Entities. In the meteorology and climate domains, Data exchange refers to using standard technical protocols and/or policies to enable data to be easily shared among organizations.
- 1.1.25 **Data, Externally-sourced.** In the NMHS context, externally-sourced data (also sometimes referred to as "third party" data, though this description is discouraged), are data obtained from providers outside the NMHS, such as private providers, other Government agencies, or international sources. The data may or may not conform to NMHS and WMO expectations for siting, instrumentation and metadata, but ideally these aspects should be well documented. A distinction may be drawn between "contracted" data from prescribed instrumentation and systems (with generally good metadata and compliance with NMHS observing standards), that might arise from partnerships with external providers, and "*Data – Crowd-sourced*" (see separate definition). Externally-sourced data are frequently sought from locations where the conventional NMHS observing network is sparse, for

incorporation into forecast products, climate analyses, etc., to provide greater local accuracy and reliability in the products.

- 1.1.26 **Data governance** includes standard procedure, policies, approval process, along with accountabilities and compliance mechanism for ensuring the data is secure, accessible and useable.
- 1.1.27 **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.
- 1.1.28 **Data Integration** is the technical process used to combine data from disparate sources into meaningful and valuable information. A complete data integration solution delivers data from a variety of sources.
- 1.1.29 **Data Inter-operability** is the ability of various components of a data/information infrastructure or systems to connect with, or "talk" to each other in order to enable data exchange collaboration between organizations. This may be done in various ways, e.g., through appropriate middleware, or by conforming with "standards", which may be thought of as a common language understood and spoken by different data systems that enable, for instance, data of different types and formats to be readily exchanged.
- 1.1.30 **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.
- 1.1.31 **Data Management Policy** Based on the *Data Management Association International* (DAMA) definition, a data management policy may be described as a short statement of management intent and requirements governing the management, security and integrity of data and information within an Entity. The purpose of a data management policy is to ensure that consistent rules surrounding data stewardship within the Entity are established and communicated, in order to harmonize processes across the Entity. As a high-level directive, it is expected that a Policy, once implemented, would only rarely be altered, whereas the specific details of how the Policy might be implemented would sit outside the Policy and be amended as needed. A typical data management policy will set out details of the context and scope of the policy, the specific requirements and who is responsible for them, compliance mechanisms and overall ownership of the policy, along with a statement of when or under what circumstances the Policy would be reviewed.
- 1.1.32 **Data Policy.** A Member country's data policy refers to *the use and limitations surrounding its data by the supervising organization* (italicized part of definition taken from Congress XVII (2015)).
- 1.1.33 **Data Policy-WMO Resolutions 40 and 25**, commits WMO Member countries to broadening and enhancing the free and unrestricted international exchange of a specific set of meteorological and related data and products (Resolution 40, (Cg-XII)), and the same for hydrological data and products (Resolution 25, Cg-XIII). In the context of this resolution "Free and unrestricted" means non- discriminatory and "Without charge", that is, provided at no more than the cost of reproduction and delivery, without charge for the data and products themselves.
- 1.1.34 **Data Policy-WMO Resolution 60 (Cg-17).** In addition to the exchange of data and products essential for operational weather analysis and forecasting as provided under Annex 1 to Resolution 40 (Cg-XII), and hydrological data and products required to support hydrological applications under Resolution 25 (Cg-XIII), there are specific data and information requirements needed to support the Global

Framework for Climate Services (GFCS). In addition to all data and products that are already available on a free and unrestricted basis, Resolution 60 states that the following types of data and products are necessary for the implementation of the GFCS:

- (i) Historical climate time-series from the Regional Basic Climate Networks (RBCNs), the GCOS Upper-Air Network and GCOS Surface Network at a temporal and spatial resolution necessary to resolve the statistics of climate, including trends and extremes;
- (ii) Essential climate variables for the ocean (full depth) (as defined by the GCOS Implementation Plan);
- (iii) Climate relevant coastal interface data, in particular sea level, waves and storm surges;
- (iv) Data on the composition of the atmosphere including aerosols;
- (v) Climate relevant satellite data and products;
- (vi) Climate relevant cryosphere data, in particular snow cover, snow depth, glacial monitoring, permafrost and lake and river ice.

1.1.35 **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.

1.1.36 **Data Product** refers to a product that facilitates an end goal through the analysis or use of data, and ideally is derived using a scientifically-sound algorithm or approach.

1.1.37 **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.

1.1.38 **Data Proxy** Proxy data are measurements of conditions that are indirectly related to climate, such as phenology, ice core samples, varves (annual sediment deposits), coral reefs and tree ring growth. These data are useful for climate change detection and assessment over centuries and millennial time-scales. Details on the constituents, methodologies and use of these data are available in the *Guide to Climatological Practices*, WMO-No. 100, third edition.

1.1.39 **Data Quality** in the context of data management and stewardship, Data Quality refers to the extent to which the data complies with the data management goals as defined in the preamble of this manual. High quality data means that the data provide accurate measurements of the true state of the atmosphere, that input and output processes are reliable; the data are well described (i.e. there are adequate metadata); the data are properly quality controlled to identify errors; and that the data are accessible in a timely manner.

1.1.40 **Dataset Quality Assessment** is the process of scientifically and statistically evaluating datasets and their level of stewardship to determine, based on appropriate documentation, whether they are of adequate quality.

1.1.41 **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 assurance is required across the whole data lifecycle, and should also include ensuring effective transmission and secure management of the data.

- 1.1.42 **Data Quality Control** is the process of ensuring that errors in the data are detected and flagged. It involves checking the data to assess representativeness in time, space and internal consistency, and flagging any potential errors or inconsistencies. The purpose of Quality Control is to ensure that meteorological and climate data available to potential users is sufficiently reliable to be used with confidence. Quality control is therefore part of the overall data quality assessment.
- 1.1.43 **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.
- 1.1.44 **Data Reanalyses** are atmospheric and oceanic analyses of temperature, wind, current, and other meteorological and oceanographic quantities, created by processing past meteorological and oceanographic data using fixed state-of-the-art weather forecasting models and data assimilation techniques.
- 1.1.45 **Data Repository** is a central location in which data is stored and managed. There are two types of data repositories:
- *Designated Data Repositories* are major archives that maintain, process and distribute current and historical environmental and geospatial data. The designated repositories may serve as Agency or Country Record Centres and are subject to all of the nationally or internationally accepted archive standards.
 - *Non-Designated Data Repositories* are facilities where extensive collections of environmental parameters are maintained because of individual research, institutional research or operational requirements. Non-Designated Data Repositories are typically not held to all of the nationally or international accepted archive standards, e.g. U.S. NARA or ISO, but must still adhere to basic good data management and stewardship practices, such as off-site data backup and maintenance of adequate environmental control and security for the holdings.
- 1.1.46 **Data Rescue** is the ongoing process of identifying and preserving all data, records and climate archives that are at risk of being lost, and of digitizing current and past data into computer-compatible form for easy access. The identification process also needs to search for data possibly contained in non-NMHS repositories such as universities, libraries, national archives. In some cases historical data may be held overseas. Data Rescue includes also migration from obsolete or computer corrupted media to modern media and readable formats. This definition applies to all meteorological and hydrological data.
- 1.1.47 **Data Rescue Portal (I-DARE)**. Overseen and implemented by the WMO Commission for Climatology, I-DARE provides a gateway for information exchange and coordination on climate DATA REscue (DARE) technologies and worldwide activities. It aims to enhance the visibility of existing DARE activities; stimulate new DARE activities and promote discoverability of all climatological data that has not yet been imaged and keyed.
- 1.1.48 **Dataset or Data-set** is a collection of data and/or related metadata that are made up of separate elements but can be managed as a unit by a computer.
- 1.1.49 **Data Stewardship** is the formal accountability for ensuring effective controls (policies and practices) around the management and use of data assets, in this case the Climate Record. Stewardship (or custodianship) assigns to an Entity rights and responsibilities for acquiring and managing climate data and information. The rights include the right to determine how the information will be managed and any access constraints, with accompanying responsibilities towards, maintenance, quality, security and enabling appropriate access to that information.

- 1.1.50 **Data Uncertainty in the observational data** 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. Observational uncertainty 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, such as when gridded analyses produce interpolated values that differ from the actual point value.
- 1.1.51 **Data Usability and Usage** is how easily the data product may be understood and used by users and incorporated into the user's own working environment.
- 1.1.52 **Dataset Maturity** provides information on the level of stewardship applied to a dataset, as assessed against a maturity or capability scale (*Maturity Matrix*).
- 1.1.53 **Digital Object Identifier (DOI)**. A DOI is a registered persistent object identifier for a digital object such as a dataset, which provides an ongoing link to the object, e.g. via a Web URL. It is typically used to denote the latest version of a dataset.
- 1.1.54 **Essential Climate Variable (ECV)** is a physical, chemical, or biological variable or a group of linked variables that critically contributes to the characterization of Earth's climate. It is these variables for which international exchange is required for both current and historical observations. Note: Guidance on ECVs list and climate monitoring principles are the mandate of the Global Climate Observing System (GCOS).
- 1.1.55 **Granular data** is detailed data. For example, at pixel level the granular satellite data record could be a brightness of that pixel as recorded by the satellite on-board radiometer. It is also used to describe the breaking down of a dataset into finer-detailed components or individual elements; for instance, the rainfall record of a particular station within an overall rainfall dataset.
- 1.1.56 **Homogenization**. The technique of making time-series homogeneous, by application of scientifically sound statistical methods to remove the effects of artificial biases, such as those caused by changes in observational practices, instrumentation, siting etc.
- 1.1.57 **Maturity Matrix-Stewardship** is a unified framework for measuring the level of stewardship practices applied to data. It generally defines measurable, level-progressive practices of key components of stewardship such as preservability, accessibility, and transparency/traceability, rating each component on a level scale from not managed to optimally managed.
- 1.1.58 **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.
- 1.1.59 **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 homogenization. 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, assumptions made around data generation and collection, 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. OSCAR/Surface is the WMO

official repository of contextual metadata for all surface-based observing stations and platforms.

- 1.1.60 **Metadata, Discovery** is metadata which enables a user to query or search a catalogue to determine what information is held, where it's 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.
- 1.1.61 **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 across a network. It is important again to document the time, location and details of any such changes. Moreover to support effective homogenization, simultaneous changes across an entire network should be avoided.
- 1.1.62 **Metadata, Provenance.** Apart from the need to know what changes to observation siting, practices, etc. have been made over time (an essential step in homogenization 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 a particular analysis or product was derived. Provenance should therefore include details of any quality control or homogenization 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 particular version of the data on which they are based.
- 1.1.63 **Reference period for climate change monitoring.** In addition to the standard periods for computing Standard Climatological Normals, a fixed reference 30-year period is defined specifically for climate change monitoring purposes. This is currently 1961-1990, which will be maintained until there is a compelling scientific reason to change it. Note: the 1961-1990 reference period which can be easily applied for conventional data, may not be applicable to more recent datasets such as those derived from satellite. Other reference periods can be defined for other purposes such as for climate projection scenarios.

1.2 **MANAGING DATA FROM VARIOUS SOURCES**

- 1.2.1 **Generic Climate Requirements for observations and data management**
- 1.2.1.1 Entities making, or processing observations shall put in place appropriate quality control and quality assurance procedures.
- 1.2.1.2 Entities making, or processing observations should include independent external audit of their procedures as part of their quality assurance.
- 1.2.1.3 Entities making observations shall use appropriate standards for making and recording the observations, and document the processes used to make the observations and the conditions under which the observations were made.
- 1.2.1.4 Entities providing observations shall implement and sustain arrangements for the long-term preservation of their observations and appropriate metadata.
- 1.2.1.5 Entities preserving observations and appropriate metadata shall manage the observations and metadata so as to ensure traceability, transparency, interoperability, discoverability and accessibility.

- 1.2.1.6 Entities providing observations should arrange for those observations to be shared in accordance with relevant national, regional and international data sharing policies.
- 1.2.1.7 Members should ensure their observation networks can provide long time series of homogenous observations that meet their own needs and the minimum climate requirements of the international programmes.
- 1.2.1.8 Members should take account of potential availability of data from all sources, such as the National Meteorological and Hydrological Service, other government agencies, and the private sector as potential contributions to the Climate Record.
- 1.2.1.9 Entities should use standard terminology for their data management activities as defined in this Manual to ensure consistency of data management practices and interoperability of data management systems.
- 1.2.1.10 Entities shall incorporate Quality Control and Quality Assurance as an essential part of their climate data management practices.
- 1.2.1.11 Entities which generate datasets for use in climate science and services shall ensure that the various types of Metadata as defined in this Manual are well described, maintained, archived and made available.
- 1.2.1.12 Entities should establish a Quality Management System around their climate data management processes, including process monitoring, documentation and improvement mechanisms, as far as possible compliant with relevant WMO *Quality Management in Climate Services* guidelines as summarized in WMO-No. 1221.
- 1.2.1.13 Entities should ensure adequate versioning of their datasets, in line with the definition of "Provenance Metadata" outlined above.

Note:

- (1) In support of climate activities, HQ-GDMFC is intended to enable the development and updates of best practices and standards for managing data from all sources that are useful or have the potential to become useful for climate science, applications and services.
- (2) The guidelines described in WMO No 1221 are based on the ISO 9001: 2015 *Quality Management System – Requirements* standard.

- 1.2.1.14 WMO requires NMHSs to routinely provide basic climate-related data in agreed standard formats, as follows:
 - (i) NMHSs shall provide Climatological Standard Normals and monthly CLIMAT reports;
 - (ii) NMHSs should provide daily CLIMAT and annual updates of World Weather Records datasets;
 - (iii) NMHSs should exchange National Climate Monitoring Products;
 - (iv) NMHSs should exchange all other fields requested under Resolution 60;
 - (v) All climate datasets should be assigned a Digital Object Identifier (DOI), which provides an ongoing linkage to the latest version of a dataset.

1.2.2 Conventional in-situ sources of climate data

Note:

- (1) In-situ climate data are derived from conventional observing networks. The design of these networks can be based on a tiered structure that places differing technical constraints on observing stations in different tiers of the network. In addition to improving the quality and utility of observations used for climate purposes, this design will also lead to improvements in the understanding of the quality of the observations.

- (2) WMO-No. 1185 provides guidelines on observation networks and systems for ensuring that the data collected are adequate for the needs of climate services, applications and research, based on the 10 GCOS Principles for climate monitoring.

The typology of climate stations which contribute to climate observations includes:

- 1.2.2.1 **Ordinary** climatological stations, which provide the basic land area requirements for observing daily maximum and minimum temperature, and amount of precipitation.
- 1.2.2.2 **Principal climatological stations** usually provide a broader range of observations of weather, wind, cloud characteristics, humidity, temperature, atmospheric pressure, precipitation, snow cover, sunshine and solar radiation.
- 1.2.2.3 **Reference climatological stations** provide long-term, homogeneous data for the purpose of determining climatic trends. Each Member should establish and maintain at least one reference climatological station, and ideally sufficient stations to adequately characterize its country's climate. Such stations need to provide more than 30 years of homogeneous records and should be situated where non-climate related environmental changes have been, and are expected to remain, at a minimum. Reference climatological stations should be calibrated to SI or community-accepted traceable standards with fully quantified uncertainties, have the highest level of robustness (e.g. duplicate or triplicate sensors of key variables such as temperature and precipitation), be well sited in locations least likely to be affected by urbanization and other non-climatic influences, have regular maintenance and replacement cycling of instruments, the highest standard of metadata collection including photo documentation, have continuous monitoring of system performance to resolve instrument and environmental issues as they arise, with calibration checks performed on a regular basis, and inter-comparisons carried out when sensors change.
- 1.2.2.4 **Upper-air stations** provide observations of the vertical structure of the atmosphere above the ground, including temperature, humidity and wind. The longest record of upper-air observations has been obtained from radiosonde measurements. A database of atmospheric variables dates back to the 1930s, although coverage is generally poor before 1957.
- 1.2.2.5 The **GCOS Surface Network (GSN)** and the **GCOS Upper-Air Network (GUAN)** are in-situ networks of observing stations that provide high quality data for global climate monitoring and have been established as Global Climate Observing System (GCOS) Baseline Networks. As of 1 April 2017, the GSN had 1023 stations and the GUAN had 177 stations. The Essential Climate elements from upper-air observations are listed in Table 1 of the Second Report on the *Adequacy of the Global Observing Systems for Climate in Support of the UNFCCC* (WMO/TD-No. 1143) and Table 1 of the *Global Observing System for Climate: Implementation needs* (GCOS-200, 2016).
- 1.2.2.6 The **GCOS Reference Upper-Air Network (GRUAN)** is an international reference observing network of sites measuring ECVs above the Earth's surface, designed to fill an important gap in the current global observing system. GRUAN measurements provide long-term, high-quality climate data records from the surface, through the troposphere, and into the stratosphere. These are being used to determine climate trends, to constrain and calibrate data from more spatially-comprehensive observing systems (including satellites and current radiosonde networks), and to provide appropriate data for studying atmospheric processes. GRUAN is envisaged as a global network of eventually 30-40 stations from existing observational networks (WIGOS Technical Report No. 2013-02; GCOS-170).

1.2.2.7 **Urban stations** refer to weather observation sites in urban areas. They usually observe the same elements as principal climatological stations, with the addition of air pollution data such as low-level ozone and other chemicals and particulate matter, and should also observe visibility, ultraviolet radiation levels, and variables used in the calculation of heat-stress related parameters such as Wet Bulb Global Temperature. In general, siting exposure and other features may not be compatible with WMO standards for monitoring the free atmosphere, and this should be taken into account when utilizing these data for climate purposes.

1.2.3 Automatic weather station data

1.2.3.1 Entities shall ensure that effective change management processes are implemented when moving from manual to automated observations, or from one automated system to another, or where an AWS is relocated.

1.2.3.2 Entities shall ensure that where an AWS replaces a manual station, full metadata information from the previous station is retained.

1.2.3.3 Entities shall ensure that data ingested and used for climate change monitoring purpose meets the required quality standards, including unbiased, homogeneous, consistent and complete climate time series.

1.2.3.4 Entities shall ensure that appropriate communication infrastructure, data ingestion procedures and other facilities necessary to support automated weather stations are in place and maintained.

1.2.3.5 Entities operating automatic weather stations for climate purposes shall perform regular, effective and sustained maintenance and calibration of these types of stations.

Note: Guidance on these challenges and recommendations are provided in WMO guidance on "Challenges in the Transition from Conventional to Automatic Meteorological Observing Networks for Long-term Climate Records", 2017 edition (WMO-No. 1202).

1.2.4 Remote sensing data

1.2.4.1 Members should ensure that key data management activities extend to remote sensing sources, including dataset lifecycle management, metadata standards and quality control, assessment and assurance. This requirement is critical for ensuring that remotely-sensed climate records are complete, fit for purpose and that uncertainty is well documented.

1.2.4.2 Entities managing remotely-sensed data should ensure that good data stewardship practices are followed that enable data discovery, sharing and re-use of data, as well as reducing unnecessary redundancy. The practices should include developing and implementing data management strategies, among them long term retention plans, and preserving data and metadata in multiple locations.

Note:

- (1) There are a number of climate assessment and services activities that utilize remotely-sensed data. These include monitoring temporal changes and trends in climate variables at the surface and through the atmosphere, as well as analysing extreme weather and climate events. Remotely-sensed data are of paramount importance for enhancing climate monitoring and assessment by bridging the existing gaps in the in-situ observation systems, where coverage is lacking in many parts of the earth. Examples of remotely-sensed data sources include satellite data, radar data, lightning detection systems, Lidar and Sodar, etc.
- (2) Notwithstanding efforts within some entities, standardization in radar data formats and metadata is currently still lacking.

1.2.5 Space weather data

Note:

- (1) Space Weather encompasses the conditions on the Sun, the solar wind, the magnetosphere, the ionosphere, and the atmosphere that can influence the performance and reliability of space-borne and ground-based technological systems and endanger human life or health.
- (2) Guidance on space weather phenomena, measurements and data exchange and use can be found in *the potential role of WMO in Space Weather* (WMO/TD-No. 1482, 2009).

1.2.5.1 The data acquired to support Space Weather shall be exchanged based on a standard internationally agreed format which are most currently used by Space Weather agencies.

1.2.5.2 Space Weather data shall be properly identified in Catalogues and described by metadata in accordance with WIS agreed standards.

1.2.6 Aircraft data *(to be reviewed by AMDAR office)*

1.2.6.1 Entities should ensure, in addition to common data management practices, the following two specific requirements:

- (i) Documenting the potentially large number of error sources contributing to aircraft measurement uncertainty;
- (ii) Understanding and publishing the influence of sampling interval and averaging time on statistical properties of the data. (Variation of air speed for a single aircraft, and between different aircraft types, alters the sampling distances and varies the wavelengths filtered).

Note:

- (1) Entities refer here to Aircraft data providers, or relevant dataset developers, when these data are also used or being potentially useful for climate change monitoring.
- (2) Many long-distance aircraft are fitted with automatic recording systems that regularly report temperature and wind, and in some cases humidity, while en-route. Some aircraft record and report frequent observations during take-off and descent, which significantly augment the conventional radiosonde data, at least throughout the troposphere. Such data are assimilated into operational meteorological analysis systems and, through programs of reanalysis, ultimately contribute substantially to the broader climate record. Aircraft Meteorological DATA-Relay systems (AMDAR) operate on aircraft that are equipped with navigation and other sensing systems. While not as precise and accurate as most ground observing systems, aircraft data can provide useful supplemental information to meteorological databases which can also be a good source for climate applications.

1.2.7 Ocean and marine data

1.2.7.1 Entities shall ensure that data from mobile platforms is traceable back to a specific ship, buoy, etc. whilst ensuring that any security requirements associated with the platform are honored.

1.2.7.2 Entities providing observations and data, and contributing to the MCDS shall share the data on a free, open and unrestricted basis through the appropriate international data centres and systems, including the WIS and relevant IOC systems.

1.2.7.3 Entities operating as an MCDS centre shall operate according to the appropriate terms of reference, scope, capabilities, functions and tasks listed in WMO-No. 558, WMO-No. 471 and in their initial application to become an MCDS centre.

1.2.7.4 Entities operating as an MCDS centre shall adhere to the regulations with respect to marine meteorological and oceanographic observations and data management given in the WMO-No. 471, and WMO-No. 558, 2018 Editions. These include requirements for minimum quality control at the Data Acquisition Centres (DAC) and higher-level quality control for Centres for Marine Meteorological and Oceanographic Climate Data (CMOC) respectively.

Note:

- (1) Ocean and Marine data can generally be classified into physical-dynamical and biochemical elements. The physical-dynamical elements (such as wind, temperature, salinity, wind and swell waves, sea ice, ocean currents and sea level) play an active role in causing the marine system to change. The biochemical elements (such as dissolved oxygen, nutrients and phytoplankton biomass) are generally not active in the physical dynamical processes, except perhaps at long timescales, and thus are called passive elements.
- (2) The marine components of the WMO Global Data Processing and Forecasting (GDPFS) are addressed by the Infrastructure Commission in consultation with the Joint Collaborative Board of WMO and the Inter-Governmental Oceanographic Commission (IOC). The regulations covering the collection of observations, the data flows and archiving of climate data apply to its newly established Marine Climate Data System (MCDS).
- (3) The MCDS provides standards and recommended practices and procedures, together with non-regulatory guidance on the collection, rescue, digitization, exchange, data processing, quality control (QC), value adding and data flow of marine meteorological and oceanographic climate data and products from various sources. Real-time (RT) and delayed-mode (DM) data are collected through a network of specialized centres, and ultimately aggregated at the Centres for Marine Meteorological and Oceanographic Climate Data (CMOCs), which provide higher-level QC and deliver the consistent data and products needed for a wide range of marine climatological applications.
- (4) Basic sources of data include in situ observations, for example from ships, moored and drifting data buoys, tide gauges, expendable bathythermographs (XBTs), profiling floats, surface and subsurface gliders, as well as remote sensing data from satellites, aircrafts and a few other specialized sensing systems such as land-based HF radars.
- (5) Programmes for the training and competency assessment of marine meteorologists and technical support personnel are developed and applied in the system.
- (6) Guidance on the marine meteorological and oceanographic observations, data and services can be found in the *Guide to Marine Meteorological Services* (WMO-No. 471, 2018 Edition).
- (7) Guidance on quality management systems can be found in the *Guide to the Implementation of a Quality Management System for National Meteorological and Hydrological Services* (WMO-No. 1100, 2017 Edition). This includes a description of how the principles of the framework could be applied.
- (8) The application and evaluation procedure to become an MCDS centre is described in the *Manual on Marine Meteorological Services* (WMO-No. 558).

1.2.8 Atmospheric composition data (to be consulted with GAW)

Note: Collaboratively there is a need to collect, maintain and exchange data on significant climate-impacting atmospheric components such as Greenhouse Gases (specifically the GCOS ECVs for atmospheric composition such as CO₂, methane etc.), pollutants such as black carbon, Sulphur oxides, aerosols, and ozone-depleting substances such as those listed as "controlled substances" under the Montreal Protocol.

[Recommended practices will be defined in collaboration with GCOS (implementation plan) and GAW recommendations]

1.2.9 Hydrology data

- 1.2.9.1 Entities should share data including through WHOS and relevant Global Data Centers.
- 1.2.9.2 Entities should apply appropriate climate Data Management and Stewardship standards and recommended practices when hydrological data are used in climate change monitoring
- 1.2.9.3 Entities responsible for water monitoring should ensure that long historical time series are continuously updated, with appropriate quality assessment, and shared, with the support of existing mechanisms within WMO (e.g., HydroHub-WHYCOS, WHOS, Global data centers).

Note:

- (1) Hydrological data such as river level, river discharge or groundwater level are collected and managed by NMHSs or collaborating agencies. However, many low- and middle-income Countries have insufficient water monitoring systems due to constraints such as budget, lack of qualified staff or insufficient integration of the National Hydrological Services (NHSs) with key national partners and collaborating agencies.
- (2) The WMO HydroHub is the WMO Global Hydrometry Support that includes WHYCOS, or the High-Level Panel on World Water Data Initiative. Shareable data can be discovered through the WHOS portal (including real time data) and through the three WMO global hydrological data centers for historical data: GRDC (Global Runoff Data Center hosted by Germany), IGRAC (International Groundwater Assessment Center, hosted by the Netherlands) and HYDROLARE (Lake and Reservoirs Data Center, hosted by the Russian Federation).

1.2.9.4 Entities should rely on WMO Regulatory material for ensuring data quality, such as the *Technical Regulations Volume III (Hydrology, WMO 49)*, the *Guide to Hydrological Practices volumes I and II (WMO 168)*, the *Manual on Stream Gauging (WMO 1044)* and the *Guidelines for Hydrological Data Rescue (WMO 1146)*, the latter being especially useful for safeguarding long term series for climate analysis.

Note: Entities referred to here, include hydrological data providers, or relevant dataset developers, where these data are used or are potentially useful for climate change monitoring. An NMHS which is not responsible for Hydrology data should make necessary arrangements with providers of such data for making these data available for their activities in climate change monitoring.

1.2.10 Climate Model data

1.2.10.1 Entities using or providing datasets derived using climate models should ensure appropriate data management-related practices for the datasets, and specifically:

- (i) A close integration of observations, reanalyses and climate simulations so as to support model development, evaluation and inter-comparison efforts;
- (ii) Full model metadata documentation, including identifying computational algorithms and parametrizations used, versioning and citation (with associated Digital Object Identifier (DOI) preferably), and guidance on usage in science and climate related applications;
- (iii) Traceability for quality control, assurance, dataset maturity and uncertainty estimates;
- (iv) Support an open data policy that ensures data access to various user communities, so as to broaden the customer base and end-user feedback;
- (v) Interoperable data formats, such as NetCDF, its CF convention and associated CMIP and similar standards;
- (vi) Distributed data archiving infrastructure with replications (such as the Earth System Grid Federation), to share risk and avoid access issues;
- (vii) Appropriate accountability and disclaimers reflecting the data context, including whether it is research data or data supporting operational services.

Note: Climate model data incorporates a broad set of records generated using some sort of dynamical and/or statistical tools. Climate model data are used to improve our understanding of the climate system's variability and change, to support assessment reports such as those from the Intergovernmental Panel on Climate Change, to improve climate services including seasonal prediction and others specified within the Global Framework for Climate Services. Model data are also employed in the provision of climate information and services in support of adaptation and mitigation policies, and to enhance resilience to climate related disasters. The requirements in this section apply to the following categories of model data:

- (1) Climate analysis/monitoring and reanalysis products, derived from observations or a mix of observations and models, and used to represent an historical record of the climate;

- (2) Operational climate predictions, which are initialized with observations, and rely on dynamical or statistical techniques to predict the climate on monthly, seasonal or decadal time scales;
- (3) Climate projections on multi-decadal to centennial time scales, based on broad assumptions about future societal scenarios; and
- (4) Downscaled climate simulations, which are relevant to any of the above categories, and focus on smaller domains with higher resolution dynamical or statistical models, aiming at added value to the coarser resolution models.

1.2.11 Externally-sourced data (also, non-NMHS or “third party” data)

- 1.2.11.1 NMHSs should assess the opportunities offered by externally-sourced data and establish, as needed, arrangements in line with WMO Integrated Global Observing System (WIGOS) guidance for using these data in climate applications and integrating them into their data management systems (see *Guidance on WIGOS Data Partnerships*, manuscript in preparation, 2018).
- 1.2.11.2 Entities responsible for managing externally-sourced data should ensure that procedures are set up to differentiate these data from data from NMHSs networks, as different quality and metadata standards may apply.

Note: Externally sourced data include data provided by volunteer observing networks, observations based on arrangements between the Meteorological services and other Government agencies or private providers, measurements recorded by sensors attached to vehicles or roadways, observations taken by enthusiastic amateurs such as storm chasers, often via social media, and other sources. Externally-sourced data also encompasses the wealth of observational and model data generated by the broad research community, such as those coordinated by the World Climate Research Programme (WCRP), Reanalyses, and climate predictions and projections. These sources support the WMO mandate, but also present interoperability, sustainability and resourcing challenges.

1.2.12 Socio-economic data

- 1.2.12.1 NMHSs should liaise with national statistical offices, emergency and disaster management authorities and any other authoritative sources for such data to enable effective sharing of these data.
- 1.2.12.2 Entities responsible for sharing climate data should make those data available in formats that can be integrated with socio-economic data.
- 1.2.12.3 Entities responsible for collection and/or development of socio-economic data should apply international data management standards and recommended practices when these data are used for climate change monitoring (such as climate change impact data)

Note: Synergies and linkages with socio-economic data providers are essential to enable best practices and standards to be employed in collecting and providing information on climate related socio-economic impacts. Examples of such data include statistics on diseases and mortality, crop monitoring, loss and damage, population displacement and migration, energy production, and consumption and economic growth. This requires Entities responsible for climate data to work with providers of socio-economic data to agree on data supply and interoperability mechanisms.

PART II REQUIREMENTS FOR CLIMATE DATA STEWARDSHIP

2.1 DATA MANAGEMENT REQUIREMENTS

2.1.1 General requirement

- 2.1.1.1** Entities shall establish and document data management capabilities and practices for their climate data.

Note: Guidance on requirements and functionalities for computer-based Climate Data Management Systems is provided in *Climate Data Management Systems Specifications* WMO-No. 1131.

2.1.2 Discoverability and Accessibility

- 2.1.2.1 Data and all associated metadata shall be easily discoverable and accessible by the user, and data access and use conditions, including licenses, shall be clearly indicated.

2.1.3 Usability

- 2.1.3.1 Data should be structured using interoperable international standards for encoding, exchange and retrieval.
- 2.1.3.2 Data shall be comprehensively documented, including all elements necessary to ensure access, proper use, and efficient processing, using international or community-based metadata standards.

2.1.4 Quality Management

Provenance

- 2.1.4.1 Entities should document and retain information on data provenance, designated by the term Provenance Metadata. Such metadata should document, at a minimum, the origin and processing history (including any changes) of raw observations and derived data, such as from satellite and models, and links to the documented methodology used in creating products to ensure full traceability of the product chain.

Note: For example, a time-series of tropospheric temperature derived from satellite data should include details of the mix of satellites used, how the data were combined, as well as details of any corrections for orbital or calibration drift, changes in channel or sensor, and details of overlap of atmospheric bands.

- 2.1.4.2 Where information on changes to observing practices has been collected, Entities shall provide this information as part of the provenance record.
- 2.1.4.3 Entities shall ensure that results of any inter-comparison trials are archived, as part of the dataset metadata.
- 2.1.4.4 Entities shall ensure that for critical climate datasets such as those used to monitor climate change, full details of any changes to the dataset are kept and versioned.
- 2.1.4.5 Entities should ensure that a snapshot of the available data should be captured at regular intervals, so that, should data for any reason be lost, it is possible to restore the missing data.

Quality Assurance

- 2.1.4.6 Entities shall subject all climate datasets to quality control and record the result of the quality control in the metadata for the dataset.

Note: Guidance on quality control and procedures for observational data are provided in the *WMO Guide to Climatological Practices* (WMO-No. 100), and the *Guidelines on surface station data quality assurance for climate applications*, http://www.wmo.int/pages/prog/wcp/wcdmp/hq-gdmfc/documents/QC_QAguidelines-April2019.pdf.

- 2.1.4.7 Entities shall assign quality flags, or indicators, to climate data, indicating whether or not the data have been QCed.
- 2.1.4.8 Entities should indicate the assessed quality of the data following QC, at each stage of the QC process.
- 2.1.4.9 Climate quality assurance processes shall include monitoring procedures. The results of QC processes shall be analyzed, and systemic errors or problems revealed by the analysis shall be fed back to the data providers for rectification.
- 2.1.4.10 Entities shall record, as part of the provenance metadata, any homogenization and reprocessing carried out on the data, and provide links to documentation of how the homogenization and reprocessing were performed.

Documentation

- 2.1.4.11 Entities shall maintain, revise and regularly update documentation of their data management processes, including details of responsibilities and accountabilities. Such documentation should be easily retrievable for all relevant personnel.

Note: Such documentation should contain reference to more detailed material such as guideline documents and procedural manuals that describe the procedures and include identification of various intervention levels for queries and problem solving.

- 2.1.4.12 Work involving changes to the Data Management System and associated software shall be communicated in a timely fashion to stakeholders likely to be affected by the changes, and action taken to mitigate any adverse effects.

2.1.5 Stewardship

Protecting

- 2.1.5.1. Members shall ensure that a business continuity plan to mitigate risks associated with disruption of operations to its database is developed and maintained. Such a plan should incorporate provision for routine backup, and procedures for timely restoration of the database and associated infrastructure.
- 2.1.5.2 Entities shall ensure that original and irreproducible climate records and their associated metadata are permanently retained .
- 2.1.5.3 Members shall ensure that a robust retention policy is in place, such that all other climate data are assigned retention periods, based on an assessment of their operational, research, legal, evidentiary, business and wider social and historical value, considering also the cost of regenerating versus storing the assets. All decisions on retention of climate-related data, including permanent retention for original climate archives , or temporary retention (such as for model outputs) shall be planned, authorized, accountable, and recorded by the responsible Entities.
- 2.1.5.4 Members should ensure management and preservation of all data in duplicate repositories for the duration of their retention periods.
- 2.1.5.5 Entities shall protect data and metadata from loss or corruption and preserve it for future use. This includes defining and enforcing retention schedules and plans for technology migration.
- 2.1.5.6 Data and associated metadata held in data management systems should be periodically monitored to ensure integrity.
- 2.1.5.7 An Entity planning, or entering into, a project or initiative concerning climate data should, as part of its documentation plan, document arrangements for managing the data, including appropriate resourcing.

Note: It is recommended that Entities develop specific data management plans for this purpose.

Enabling reprocessing

- 2.1.5.8 Entities should manage climate data in such a way as to enable corrections, updates and reprocessing where needed.
- 2.1.5.9 Entities should assign a permanent identifier for each published dataset in order to allow it to be cited.

Note:

- (1) The citation should ensure that the origin of the datasets is duly acknowledged.
- (2) Entities creating climate data by combining climate data from other sources should document the origins of the constituent datasets in the provenance metadata.

Securing

In order to secure data:

- 2.1.5.10 Entities shall implement rigorous authorization processes for modification or deletion of data;
- 2.1.5.11 Entities should ensure that systems for physical- and cyber-security of the Climate Record and Archive are in place and regularly reviewed.
- 2.1.5.12 Entities providing public access to the Climate record or Climate archive should provide that access through a copy of the data and not allow public access to the original database.
- 2.1.5.13 Entities shall back up Climate Records at regular prescribed intervals, including scheduled backups at offsite locations.
- 2.1.5.14 Entities shall test their ability to recover climate records from a backup and take actions necessary to ensure that they would be able to restore the climate record database.

Creating Discovery Metadata

- 2.1.5.15 Entities shall create, maintain and manage metadata to facilitate the discoverability of, access to and use of climate data. The metadata shall indicate at the minimum information on data owner, contact point, name of the ECVs included in the dataset and the date of dataset publication.

Note: Other information that would be helpful includes Data Source: in-situ, remote sensed, Reanalyses, prediction, projection; geographical scale: Global, Regional, National; Time scale: time period covered by the dataset; Title and Version of the Dataset; DOI if exists; how to cite the dataset in references; Data Sharing Policy, including provision of reference, link or contact to inform on the policy for access and use of the data.

2.1.6 Data Management Governance

Note: Structurally, governance includes standard procedures, policies, approval processes, along with accountabilities and compliance mechanisms for ensuring the data is secure, accessible and useable. The scope of data management governance will be influenced by an Entity's strategy, the user requirements as well as national and international requirements.

Governance mechanism

- 2.1.6.1 Entities should implement and maintain a governance mechanism to oversee the various governance elements, coordinated by a governance body.
 - (i) The governance body should comprise stakeholders representing the data service providers and subject matter experts in data management.
 - (ii) The governance body should have the authority to adopt key decisions on data management aspects and endorse related policies and procedures.

- (iii) The governance mechanism should include processes for monitoring compliance with such policies and procedures.

2.1.6.2 Entities should consider relevant national legislation and international obligations with regard to data management, in particular to take into account WMO Resolutions 25 (Cg-XIII), 40 (Cg-XII) and 60 (Cg-17) on data policy.

Data Management Policy

2.1.6.3 Entities should have a data management policy-suite that provides a consistent, enforceable and well communicated set of formal statements of requirements and improves the efficiency and commonality of data-related processes. A data management policy should be a clear and concise statement of requirements around data management.

Note:

- (1) The data policy may need to include appendices for describing the procedures and standards.
- (2) A template illustrating the structure and content of a data management policy document along with an example is provided <http://www.wmo.int/pages/prog/wcp/wcdmp/hq-gdmfc/documents/AttachmentB.GuidanceonDataManagementPolicy-06022019.pdf>.

Data preservation

2.1.6.4 Entities shall develop a data rescue strategy that includes:

- (i) Undertaking data rescue activities in a systematic way to ensure preservation and recovery of all non-digital and superseded media archives;
- (ii) Establishing plans to digitize non-digital data into Climate Data Management Systems;
- (iii) Putting in place a migration strategy for media storage of climate archives. Media storage should be reviewed at regular intervals to ensure modernization and sustainability;
- (iv) Mobilizing resources, including financial and staff, for the maintenance and sustainability of the climate record as described in the preceding regulations; and
- (v) Developing and maintaining inventories of all data that are held in the Climate Record and Climate Archive.

Note:

- (1) Guidance on the preservation, rescue and digitization of climate data is provided in the *WMO Guidelines on Best Practices for Climate Data Rescue* (WMO-No. 1182).
- (2) Information on international efforts for documenting Data Rescue projects and activities worldwide, is hosted in the WMO international Data Rescue Portal (I-DARE), <https://www.idare-portal.org/>.

2.1.7 Climate Data Management competencies

Note: Competency refers to the knowledge, skills and behaviours required to perform specific tasks in the fulfilment of a job responsibility - *Basic Documents, 2. Technical Regulations, Volume I - General Meteorological Standards and Recommended Practices* (WMO-No. 49).

2.1.7.1 Entities should adopt clear definition of the competencies, roles and responsibilities required for staff involved in the stewardship of climate data.

2.1.7.2 Entities should develop plans to meet recruitment, capacity development and training needs to ensure that they have access to people with the competencies required.

Note: guidance on Responsibilities and Roles in climate data management is available at http://www.wmo.int/pages/prog/wcp/wcdmp/hq-gdmfc/documents/Attachment_A-Guidance-on-Responsibilities-Roles-in-Climate-Data-Management-06022019.pdf.

2.1.8 Ongoing skill development

2.1.8.1 Entities should assess their capabilities and processes against internationally recognized good practices in climate data and its management and analysis.

Note: Good practices are described in the *WMO Guide to Climatological Practices (WMO-No. 100)*, the *Climate Data Management System specifications (WMO-No. 1131)*, and other relevant WMO Guidelines, in particular dealing with data stewardship, data rescue, assessment and quality control/assurance.

2.1.8.2 In order to maintain the continuity and quality standards of climate data, Entities should:

- (i) Ensure that the required competencies for climate data management keep pace with evolving methodologies and new technology;
- (ii) Ensure that when new methodology or technology is acquired, there is sufficient, efficient and effective knowledge transfer and onsite training to ensure the implementation of the new tools is effective and sustainable;
- (iii) Provide updated training to their trainers at intervals that allow changes in good practice and feedback from quality assessments to be acted on;
- (iv) Ensure climate data management staff are trained and certified according to relevant WMO standards, as an ongoing process that ensures that there are adequate staff at all times to meet requirements and allow seamless succession planning.

Note: the required practices and training intervals will depend on the national and local situation; simply replicating an approach used somewhere else may be counter-productive.

2.1.9 Planning and design

2.1.9.1 Within the appropriate national, regional and global plans, Entities should ensure that when planning or designing a climate data management system, the mandatory functions of such systems comply as far as possible with international requirements. The goal is to make sure that the Climate Record and Archive of an Entity is reliably and sustainably maintained, and that high-quality climate data required by the World Climate Programme, such as CLIMAT messages, World Weather Records, Climatological Standard Normals and climate indices, are delivered in a timely and effective manner.

Note: *The WMO Climate Data Management Systems Specifications (WMO-No. 1131)* is recommended as a primary source of such requirements. HQ-GDMFC promotes active participation in international initiatives in the domains of Data Rescue and CDMS development and implementation.

2.2 CLIMATE DATASET MATURITY ASSESSMENT

2.2.1 Purpose

2.2.1.1 A Maturity Assessment methodology applying to Climate Datasets shall be established by WMO as a tool for a consistent data stewardship review process that evaluates how well climate datasets are managed. This methodology is termed a Stewardship Maturity Matrix for Climate Data. Its purpose is:

- (i) Enhancing the discoverability, access and use of high-quality climate data by developing and maintaining a WMO Catalogue of datasets covering all climate parameters;
- (ii) Ensuring users of climate datasets have transparent information about the quality of stewardship of the data and the underlying metadata;
- (iii) Helping identify and prioritize areas where the stewardship of Climate Data may be improved;

- (iv) Fostering international collaboration on data development, management and stewardship;
- (v) Providing an internationally coordinated platform for consistently assessing the stewardship maturity of global, regional and national climate datasets; and
- (vi) Making maturity-assessed datasets prominently visible and accessible through Internet search engines and the WMO Information System.

2.2.2 Use of WMO Maturity Matrix for Climate Datasets

2.2.2.1 The maturity assessment used by an Entity to assess climate datasets should be based on internationally recognized practices for stewardship maturity, such as those reflected in the WMO Stewardship Maturity Matrix for Climate Data (SMM-CD).

Note:

- (1) Maturity matrix approaches can be applied to other aspects of the end to end data lifecycle – for instance evaluating the maturity of the climate observing systems.
- (2) Guidance on WMO SMM-CD is available at https://figshare.com/articles/The_manual_for_the_WMO-Wide_Stewardship_Maturity_Matrix_for_Climate_Data/7002482.

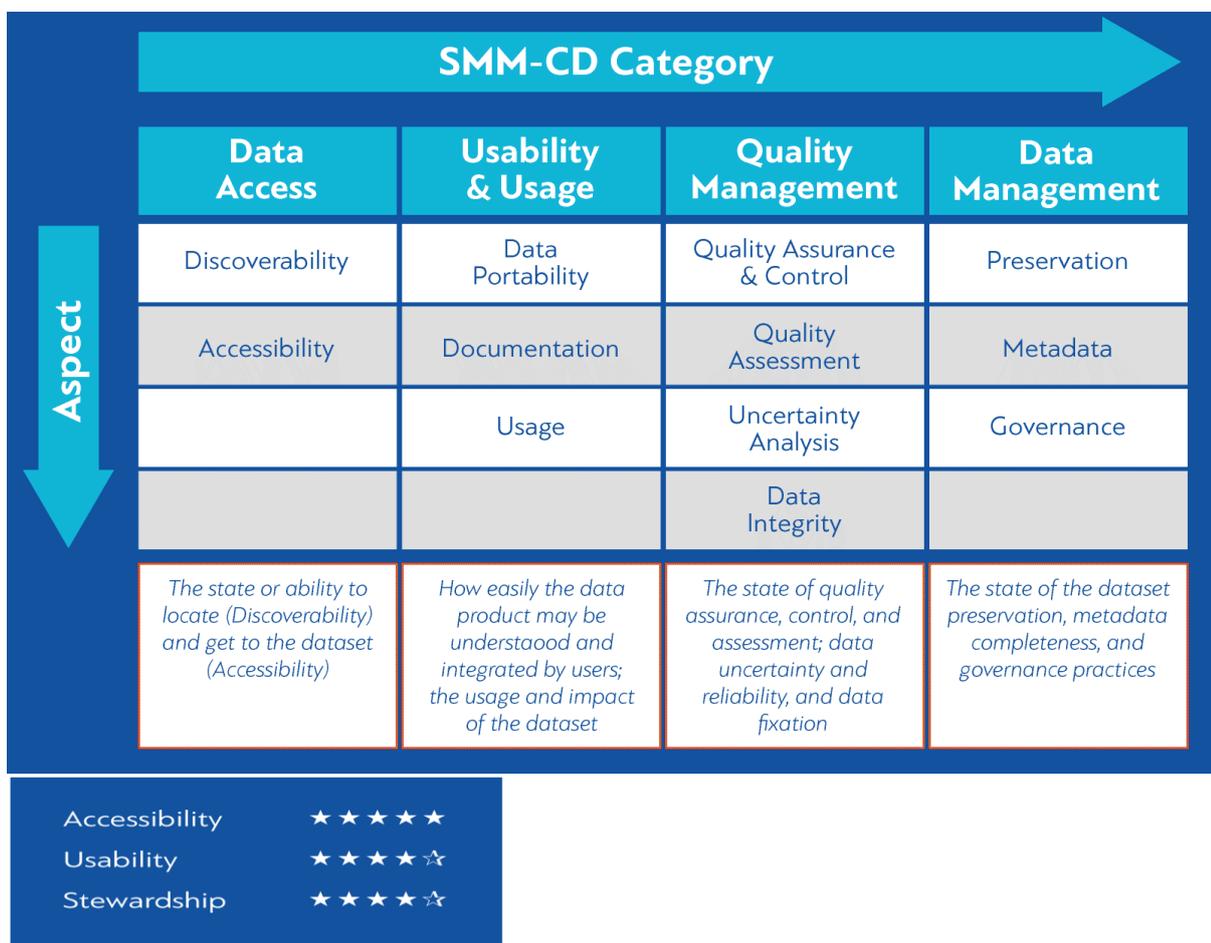


Figure 2: WMO Stewardship Maturity Matrix for Climate Data (SMM-CD) enables data providers to assess and rate their datasets quantifiably based on internationally validated data stewardship best practices. The four categories defined for SMM-CD are: Data Access, Usability & Usage, Quality Management, and Data Management. For each category, there are two to four sub-categories, referred to as “Aspects”. There are twelve categories/aspects in total for which maturity can be rated from the lowest level 1 to the highest level 5. In this example, the rating shown in 5 stars is an example of results of maturity assessment. Rating of stewardship is an average of rating of quality management and data management categories.

2.2.3 WMO Climate Data Catalogue

Note:

- (1) Datasets of key ECVs are needed to describe climate variability and change. Scientists and Decision- and Policy-makers are looking for trusted datasets that are easily discoverable, accessible and retrievable.
- (2) A climate data catalogue in the context of climate change monitoring is intended to provide a living list of datasets of the GCOS Essential Climate Variables with a primary focus on climate indicators such as Temperature, Precipitation, Sea Level, Sea Ice, Ice Sheet and Glaciers, Climate Extreme indices, Hydrology and others. All such datasets should be maturity assessed and a maturity rating will provide the user with information on the level of documentation, archival, access, data quality assurance, data integrity and more, for each of the datasets.

- 2.2.3.1 A WMO Climate Data Catalogue which includes maturity-assessed climate datasets shall be established to document the maturity of the datasets according to the criteria described in the Guidance on WMO SMM-CD (see 2.2.2.1 above), and to facilitate their documentation, discovery, access and retrieval. The Catalogue will be maintained and updated under WMO auspices.
 - 2.2.3.2 The Catalogue shall provide information about access to, and ensure updates where appropriate for, a list of datasets for which Maturity has been assessed and maturity scores assigned.
 - 2.2.3.3 Global climate datasets proposed to be included in the WMO Climate Data Catalogue should first be recommended by one of the GCOS science panels or other authoritative sources who will base their recommendation on an established scientific foundation and quality criteria such as observation uncertainty, calibrations, metadata completeness, etc.
 - 2.2.3.4 Climate datasets having regional, national or a limited (non-global) geographical domain of coverage can be included in the Catalogue if a relevant GCOS science panel, or an authoritative scientific body at national or regional level considers them of high relevancy to global climate science efforts. In this case the same process for maturity assessment and cataloguing shall be followed as for the global datasets.
 - 2.2.3.5 An Entity which desires to include a Climate dataset in the WMO Climate Data Catalogue shall conduct a self-assessment of the dataset against the WMO Stewardship Maturity Matrix for Climate Datasets (see 2.2.2.1) and submit the Dataset proposal to the relevant WMO technical committee responsible for climate data management.
 - 2.2.3.6 For each dataset included in the Catalogue, the documents providing the results of the maturity assessment should be made available to users upon request.
-