The 'WMO Stewardship Maturity Matrix for Climate Data' Template

Document ID: WMO-SMM-CD-0003; Version: v03r01 20190227

Steps for carrying out a self-evaluation of data stewardship maturity of a dataset:

i) Download the latest **WMO** Stewardship Maturity Matrix for Climate Data **(SMM-CD)** template file from doi:10.6084/m9.figshare.7003709;
ii) Go over the whole file and read the disclaimer carefully before using the template;
iii) Enter dataset and relevant Point-Of-Contacts (POCs) information in the **WMO SMM-CD** metadata section;
iv) Read the content of **WMO SMM-CD**; for additional background information, download and read the guide for the SMM-CD from doi: 10.6084/m9.figshare.7002482;
v) Go through each Category with its relevant Aspects, identify the stewardship practices applied to the dataset, and document your rating and justifications;
vi) Obtain any additional information if necessary and, if any, restrictions should be stated;
vii) Review the results and fill in the matrix cells with the defined color scheme (provided in Table I);
viii) Capture the assessment results in the **WMO SMM-CD** metadata section and in the header **'Stewardship Maturity Ratings'** for each Aspect, along with the evidences.

Assumptions:

- Datasets are digital environmental data products that are publicly available online.
- Evaluators who use this template have a basic knowledge of or are able to obtain information about conventions or standards relevant to practices examined in each key component in the community that datasets are produced for or/and provided to.

WMO Stewardship Maturity Matrix for Climate Data (SMM-CD) Assessment as of <01/21/2020>
for <GISTEMP>

Dataset Title	**GISS Land-Ocean Temperature Index (GISTEMP)**
Dataset Information URL	**https://data.giss.nasa.gov/gistemp**
Data Provider POC (Name; E-mail; Affiliation)	Gavin A. Schmidt (gavin.a.schmidt@nasa.gov, NASA)
Dataset POC (Name; E-mail; Affiliation)	Reto Ruedy (reto.a.ruedy@nasa.gov, NASA & SciSpace LLC)
WMO SMM-CD Version (Document ID and Version Number)	WMO-SMM-CD-0001-v02r00-20180914
WMO SMM-CD POC (Name; E-mail; Affiliation)	Christina Lief; Christina.lief@gmail.com; ret.NOAA and Ge Peng; Ge.Peng@noaa.gov; North Carolina State University, Cooperative Institute for Climate and Satellites, North Carolina (CICS-NC) at NOAA's National Centers for Environmental Information (NCEI)
WMO SMM-CD Template Version (Document ID and Version Numbers)	WMO-SMM-CD-0003-v02r00-20180914
WMO SMM-CD Template POC (Name; E-mail; Affiliation)	Christina Lief; Christina.lief@gmail.com; ret.NOAA
Assessment Version (v<nn>r<mm>, e.g., v01r00)	V02r02
Assessment Date (MM/DD/YYYY)	(updated on 1/21/2020 with information on published paper on Uncertainty Analysis) (03/09/2019 CCl ET-DDS evaluation)
Assessment POC (Name; E-mail; Affiliation)	Robert Dunn, robert.dunn@metoffice.gov.uk, Met Office Hadley Centre
Stewardship Maturity Ratings (each category and associated aspects) (c1a1, c1a2; c2a1, c2a2, c2a3; c3a1, c3a2, c3a3, c3a4; c4a1, c4a2, c4a3)	**5, 5;** **4, 5, 5;** **5, 5, 5, 4;** **5, 5, 3.**
Original Assessment Date (MM/DD/YYYY)	10/10/2018
Original Assessment POC (Name; E-mail; Affiliation)	Robert Dunn, robert.dunn@metoffice.gov.uk, Met Office Hadley Centre

Categories and Aspects	Level 1 Ad Hoc Not Managed	Level 2 Minimal Limit-Managed Not Defined	Level 3 Intermediate Managed Defined, Partially Implemented	Level 4 Advanced Well-Managed Well-Defined, Fully Implemented	Level 5 Optimal Level 4 + Measure, Controlled, Audited	The Data Access category refers to the ability to locate (Discoverability) and get to the dataset in question (Accessibility), with higher levels of maturity corresponding to the ease for a potential user to find and gain access to the dataset	
Data Access						**WMO SMM-CD Rating and Justification or Evidence**	**Comments**
Discoverability	By personal contact only; Dataset information not discoverable	Limited dataset information, such as scientific description of the methodology, in the literature	Minimal catalogue-level metadata; Dataset searchable online	Complete set of collection-level discovery metadata + minimal granular metadata	Level 4 + Available on an international catalogue, prominently displayed online and routinely updated	❖ **Level 5** Public domain dataset. Updated monthly, listed in all relevant catalogs Papers describing dataset (Hansen & Lebedeff, 1987, Hansen et al., 2010) Also via NCAR RDA: https://climatedataguide.ucar.edu/climate-data/global-surface-temperature-data-gistemp-nasa-goddard-institute-space-studies-giss and ESRL https://www.esrl.noaa.gov/psd/data/gridded/data.gistemp.html	
Accessibility	Data not available publicly; Person-to-person contact needed	Basic online services available for data access (e.g. FTP/HTTP direct download)	Non-standard data services	Standard-based interoperability data services	Level 4 + Full capability of sub-setting, aggregation and visualization	❖ **Level 5** HTTP access to files: https://data.giss.nasa.gov/gistemp/ Also via NCAR RDA: https://climatedataguide.ucar.edu/climate-data/global-surface-temperature-data-gistemp-nasa-goddard-institute-space-studies-giss and ESRL https://www.esrl.noaa.gov/psd/data/gridded/data.gistemp.html (providing THREDDS capability) Online data visualization, sub-settable means/trends/etc.	
Usability & Usage	**Level 1 Ad Hoc Not Managed**	**Level 2 Minimal Limit-Managed Not Defined**	**Level 3 Intermediate Managed Defined, Partially Implemented**	**Level 4 Advanced Well-Managed Well-Defined, Fully Implemented**	**Level 5 Optimal Level 4 + Measure, Controlled, Audited**	The Usability & Usage category describes how easily the data products may be understood and used by users and incorporated into the user's own working environment	
						WMO SSM-CD Rating and Justification or Evidence	**Comments**
Data Portability	Non-machine readable	Basic machine readable	Standards-based machine readable	Machine independent, self-describing, interoperable format	Level 4 + capability of providing user required format	❖ **Level 4** ASCII .txt and .csv as well as netCDF files available. Also at ESRL:	

						https://www.esrl.noaa.gov/psd/data/gridded/data.gistemp.html Checked using Puma CF checker. No Errors, 1 Warning, no messages Txt, csv, netcdf and native binary formats available.	
Documentation	Product information not publicly available online	Limited online documentation (e.g., User Guide)	Document on how the data product was created and how to use it, is available online	Full documentation based on a standard template and available online	Level 4 + Online tutorial on using and analyzing the dataset; Complete production system information available online	❖ **Level 5** Readme and example scripts (in Fortran) available to process files. Journal articles describing dataset (Hansen & Lebedeff, 1987, Hansen et al 2010) Source code available: https://data.giss.nasa.gov/gistemp/sources_v3/ All production data/code available. Limited tutorials on reproduction of data, multiple papers in scientific literature discussing processing. The web site contains an abundance of easy-to-use yet versatile tools to analyze the data including: • A custom plotter to graphically display time series of monthly, seasonal, annual means, of individual calendar months and seasons for means over the globe, hemispheres, various zones, ocean, land, the US with a user controlled smoothing to analyze the general trends, • A mapping utility that easily produces any anomaly or trend map, • A utility that displays zonal mean anomaly time series with a user controlled smoothing, as well as seasonal cycles • Animations showing time series of global maps	
Usage	No or weak citation in scientific publication in peer-review journal or as institutional reports	Intermediate citations + referenced in institutional climate assessment reports (e.g., by NOAA)	Strong citations + referenced in national climate assessment reports (e.g., by USGCRP)	Level 3 + referenced in international climate assessment reports (e.g., by IPCC)	Level 4 + referenced in international decision/policy making published reports (e.g., by UNFCCC, UN-ISDR, World Bank, etc.)	❖ **Level 5** Level 4 – NASA GISTEMP referenced in IPCC AR5 WG1, Ch 2 Level 5 – NASA GISTEMP used in https://unfccc.int/news/2017-was-among-top-three-hottest-years-on-record Used by IPCC/UNFCCC/WMO/USGRCP etc.	
Quality Management	**Level 1** **Ad Hoc**	**Level 2** **Minimal**	**Level 3** **Intermediate**	**Level 4** **Advanced** **Well-Managed**	**Level 5** **Optimal** **Level 4 +**	The Quality Management category encompasses quality assurance procedures including quality monitoring, quality control, and quality assessment and communication of reliability	

	Not Managed	Limit-Managed Not Defined	Managed Defined, Partially Implemented	Well-Defined, Fully Implemented	Measure, Controlled, Audited	WMO SSM-CD Rating and Justification or Evidence	Comments
Quality Assurance & Control	Ad hoc or no data quality assurance (QA) & control (QC) procedure or information unknown	QA/QC procedure are defined, documented, and partially implemented	QA/QC procedure are well-defined, according to community best practices, documented and fully applied	Level 3 + provision of error statistics published or tracked with results made available online and commutated to data providers; Procedure for user feedback, improvement prioritization in place	Level 4 + Detailed analysis of errors and gaps at space-time unit level: (Station, grid points, daily, monthly and or annual time-scale, etc.) QA/QC procedure monitored.	❖ **Level 5** Papers describing the QC (Hansen et al, 1999, 2001, 2010). It is a derived dataset, so QC information will exists for parent datasets	
Quality Assessment	Product quality assessment not done or done internally and information not available	Assessed by Principal Investigator (PI) or data producer; Assessment results available online	Level 2 + Product validation and evaluation done by PI published in peer-reviewed journal	Level 3 + Independent product validation and evaluation published in peer-reviewed journal	Level 4 + The complete product provenance is captured and publicly available	❖ **Level 5** Dataset is completely and independently reproducible from available raw data links and code. Independent validation has been provided (via comparisons to reanalyses, satellite remote sensing, independent processing).	
Uncertainty Analysis	Uncertainty estimates not available	Uncertainty estimates presented without explanation	Uncertainty estimates presented with partial explanation	Full uncertainty budget available with all assumptions; Estimates of accuracy of trend available	Full uncertainty assessment published in peer reviewed journal	❖ **Level 5** The definitions for the Uncertainty Analysis aspect have been updated. The corresponding has been published and is available at https://data.giss.nasa.gov/gistemp/uncertainty.	
Data Integrity	Unknown or no data integrity check	Random data integrity check	Data integrity verified systematically but methodology not commonly known	Data integrity verified systematically but methodology not commonly known	All steps in data integrity check systematically verified and adhering to well-known practices and reported	❖ **Level 4** The data we use passed several quality checks performed by NCEI and described on https://www.ncdc.noaa.gov/data-access/land-based-station-data/land-based-datasets/global-historical-climatology-network-monthly-version-3. Additionally, all outliers and reports that differ substantially from their neighbors' reports - discovered by a visual inspection of land data maps - are compared to reports from other independent sources; unless confirmed by those sources, the data are rejected.	
Data Management	**Level 1** **Ad Hoc** **Not Managed**	**Level 2** **Minimal** **Limit-Managed** **Not Defined**	**Level 3** **Intermediate** **Managed** **Defined, Partially Implemented**	**Level 4** **Advanced** **Well-Managed** **Well-Defined, Fully Implemented**	**Level 5** **Optimal** **Level 4 +** **Measure, Controlled, Audited**	The Data Management category refers to the processes undertaken to ensure the data and the contextual metadata are securely archived. It covers not just the preservation of the data and metadata with appropriate safeguards, but well defined and informed governance processes to endure that the right procedures are followed at the right times by the right people.	
						WMO SSM-CD Rating and Justification or Evidence	**Comments**

Preservation	Any storage location; Data only; Data not backed up	Non-designated repository; a backup copy of electronic data is made	Designated archive; Basic retention policy publicly defined. Routine backups made, including offsite copy.	Level 3 + Conforming to community archiving standards. Comprehensive retention policy defined and implemented	Level 4 + Archiving process performance controlled, measured, and audited; Future archiving standard changes planned	❖ **Level 5** All data on which our displays are based are routinely archived and backed up and available on demand.	
Metadata	Metadata not publicly available and/or not usable	Limited Metadata publicly available; Conforming to community-standard; Basic characteristic of dataset	Level 2 + Conforming to international standards in most aspects; limited quality and provenance metadata	Fully compliant with international standards; Rich metadata content; Basic granular-level metadata; Support dataset provenance	Level 4 + complete granular-level metadata; Metadata QC-ed and regularly update	❖ **Level 5** CF-1.6 compliant netCDF files – so a community standard – or is this sufficient to be international? All metadata for our products are provided. The metadata for GHCN v3 is offered by the providing organization (NOAA/NCEI). Metadata also available on the website under the "metadata" tab: https://climatedataguide.ucar.edu/climate-data/global-surface-temperature-data-gistemp-nasa-goddard-institute-space-studies-giss	
Governance	Responsibility is not defined; No person is assigned	Responsible entity is identified; Accountability and competency are not well-defined	Responsibility/accountability and compliance mechanism are defined; Good competency; Process established conforming to community standards	Level 3 + competency defined; conforming to international standards; auditable	Level 4 + accountability and responsibility well-defined and fully compliant with international standards; Transparent; Monitored and audited	❖ **Level 3** Responsible Entity (NASA-GISS) and person (Reto Ruedy) well defined and noted on website as well as metadata page	

Recommended citation for this document:

Lief, C. and G. Peng, 2019: The WMO Stewardship Maturity Matrix for Climate Data (SMM-CD) Template. Document ID: WMO-SMM-CD-0003. Version: v03r01 20190227. Figshare. doi:10.6084/m9.figshare.7003709.

References:

The SMM-CD Working Group, 2019: The Guidance Booklet on the WMO-Wide Stewardship Maturity Matrix for Climate Data. Document ID: WMO-SMM-CD-0002. Version: v03r00 20190131. Figshare. doi:10.6084/m9.figshare.7002482.

The SMM-CD Working Group, 2019: A WMO-Wide Stewardship Maturity Matrix for Climate Data. Document ID: WMO-SMM-CD-0001. Version: v03r00 20190128. Figshare. doi: 10.6084/m9.figshare.7006028.

Peng, G., 2015: The Scientific Data Stewardship Maturity Assessment Model Template, Version: NCDC-CICS-SMM-0001-Rev.1 v4.0 6/23/2015. doi:10.6084/m9.figshare.1211954.

Acknowledgement: The WMO SMM-CD was developed by the SMM-CD Working Group under the WMO Commission for Climatology High Quality Global Data Management Framework for Climate (HQ-GDMFC) initiative. We thank the members of the International Expert Group for Climate Data Modernisation (IEG-CDM) for contributions to both development of the SMM-

CD and internal review. In particular, we thank Markus Ziese, Lydia Gates, Axel Andersson, and Kate Roberts for beneficial comments during the internal review. We thank Iolanda Maggio, Peter Thorne, Simon Eggleston, Darren Ghent, Jörg Schulz, Nancy Ritchey, Ken Kehoe, Imke Durre, Carolin Richter, and Ruth Duerr for beneficial comments during the external review. We would also like to thank the following participants in our external review session at the Earth Science Information Partners (ESIP) summer meeting: John Graybeal, Wen Hoa Li, Robert Downs, Luann Dahlman, Joseph Conran, Ronalda Tsosie, H.K. 'Rama' Ramapriyan, Nancy Ritchey, David Moroni, Bob Sandusky, Ken Kehoe, and Sophie Hou.

Table I: RGB Color Scheme
for the WMO-Wide Stewardship Maturity Matrix for Climate Data

Maturity Scale	Color Code	R	G	B	Color
Level 1	Lighter Green	229	244	224	
Level 2	Light Green	203	234	192	
Level 3	Green	176	223	161	
Level 4	Dark Green	85	168	57	
Level 5	Darker Green	56	112	38	