

**NOAA Technical Information Series NESDIS
DSMR-00227 Version 1.0**



doi: 10.25923/dybz-jn13

**Data Stewardship Maturity Report for Global Historical Climatology Network
- Daily (GHCN-Daily), Version 3**

Table 1 Legend				
Level 1	Level 2	Level 3	Level 4	Level 5
Ad Hoc	Minimal	Intermediate	Advanced	Optimal
Little or no management	Limited Management	Defined Management, partially implemented	Well-defined Management, fully implemented	Full Management, audited, measured, controlled

Table 1. Scores for the Nine DSMM Key Components at a Glance		
Preservability - 4.5	Accessibility - 4	Usability - 2.5
Production Sustainability - 4.5	Data Quality Assurance - 4.5	Data Quality Control/Monitoring - 5
Data Quality Assessment - 3	Transparency/Traceability - 2.75	Data Integrity - 2

NOAA National Centers for Environmental Information January 2020



U.S. DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
National Environmental Satellite, Data, and Information Service

Cover Image: Data Stewardship Rating Diagram for Global Historical Climatology Network - Daily (GHCN-Daily), Version 3

Shades of green are used to represent level 1 through level 5 ratings; denoting Ad Hoc, Minimal, Intermediate, Advanced, and Optimal stages for each of the nine key components, respectively. The dark green level indicates all the practices are completely satisfied. The lighter green levels indicate only some of the practices are satisfied. The lightest green level indicates none of the practices are satisfied.

The stewardship maturity of NCEI data product, Global Historical Climatology Network - Daily (GHCN-Daily), Version 3, is assessed based on a reference stewardship maturity framework. The current maturity ratings of Global Historical Climatology Network - Daily (GHCN-Daily), Version 3 are at Level 1 or higher for all nine key components with zero Level 1, three Level 2, one Level 3, four Level 4, and one Level 5 key components.

The National Environmental Satellite, Data, and Information Service (NESDIS) manages the Nation's civil Earth-observing satellite systems, as well as global national data bases for meteorology, oceanography, geophysics, and solar-terrestrial sciences. From these sources, it develops and disseminates environmental data and information products critical to the protection of life and property, national defense, and the national economy, energy development and distribution, global food supplies, and the development of natural resources.

Publication in the NOAA Technical Memorandum series does not preclude later publication in scientific journals in expanded or modified form. The NESDIS series of NOAA Technical Reports is a continuation of the former NESS and EDIS series of NOAA Technical Reports and the NESC and EDS series of Environmental Science Services Administration (ESSA) Technical Reports.

Copies of earlier reports may be available by contacting NESDIS Chief of Staff, NOAA/ NESDIS, 1335 East-West Highway, SSMC1, Silver Spring, MD 20910, (301) 713-3578.

ASSESSMENT REVISION HISTORY

Revision	Description	Date
V01r00	Initial Release	12/27/2021

NOAA Technical Information Series NESDIS DSMR-00227

Version 1.0

doi: 10.25923/dybz-jn13

Data Stewardship Maturity Report for Global Historical Climatology Network
- Daily (GHCN-Daily), Version 3

Raisa Ionin, Katy Luquire

NOAA's National Centers of Environmental Information (NCEI)

151 Patton Avenue, Asheville, NC 28801, (828) 271-4800

Recommended Citation

Raisa Ionin, Katy Luquire. (2021), Data Stewardship Maturity Report for Global Historical Climatology Network - Daily (GHCN-Daily), Version 3, NOAA Technical Information Series NESDIS DSMR-00227 Version 1.0, 22pp., doi: 10.25923/dybz-jn13

Table of Contents

List of Tables	7
Preface	8
1. Introduction	9
2. Results	10
3. Acknowledgment	16
4. References	17
Appendix I The Scientific Data Stewardship Maturity Matrix (DSMM)	19

List of Tables

Table 1. Scores for the Nine DSMM Key Components at a Glance	1
Table 2. Dataset and Data Stewardship Maturity Assessment Metadata	11
Table 3. Stewardship Maturity Levels and Detailed Justifications for Each of Nine DSMM Key Components for the Dataset	12

Preface

In response to the President's Open Government Initiative and related policies, NOAA has committed to providing improved public access to all of its environmental information, to enable research and commercial innovation through ease of data discovery and use [Casey, 2016].

OneStop supports NOAA's efforts by leveraging existing access technologies and infusing specific innovations to provide improved discover, access, and visualization services for NOAA's data. Also, OneStop is viewed by a NESDIS as a pathfinder effort with an initial focus on selected high-priority datasets from NESDIS and other program data meeting OneStop standards, but eventually scalable across NOAA's data. Lastly, OneStop is implementing the USGEO Common Framework for Earth Observation Data and leveraging/supporting the NOAA Big Data Project (BDP) and Big Earth Data Initiative (BEDI) [Casey, 2016].

As with any process of improvement planning, agencies need to find out where they are in terms of their compliance to the federal regulations and what they need to do if any areas of non-compliance are identified. To this end, a unified framework would be beneficial for assessing the current stage of stewardship practices applied to individual datasets and for providing a road map that will guide future investments towards enhanced stewardship of environmental datasets. The value and quality of a dataset depends in part on the stewardship practices applied after its development and production. Therefore, a unified framework providing a holistic view of the quality of stewardship practices applied to individual datasets is beneficial to data stewards and users [Casey, 2016].

The Data Stewardship Maturity Matrix (DSMM), jointly developed by domain (data management, technology, and science) subject matter experts from NOAA's National Centers for Environmental Information (NCEI) and Cooperative Institute for Climate and Satellites – North Carolina (CICS-NC), provides such a consistent framework [Peng *et al.*, 2016]. The DSMM, leveraging institutional knowledge and community practices and standards, defines a graduated maturity scale for each of nine key components of scientific data stewardship to enable a consistent assessment of the measureable stewardship practices applied to a given data set or product.

The NOAA Data Stewardship Maturity Technical Series captures stewardship maturity assessment results for individual datasets, provides consistent representation and citable documents of those assessments, ensures transparency, and allows better data quality information integration and content-based search and discovery of NOAA data.

Data Stewardship Maturity Report for Global Historical Climatology Network - Daily (GHCN-Daily), Version 3

1. Introduction

1.1 Purpose

The purpose of this document is to describe the results of stewardship maturity assessment for NOAA Climate Data Record for Mean Layer Temperature (Upper Troposphere & Lower Stratosphere from UCAR, Version 2, utilizing the Scientific Data Stewardship Maturity Matrix or DSMM [Peng, et al, 2016]. DSMM defines levels of stewardship maturity stages for Preservability, Accessibility, Usability, Production Sustainability, Data Quality Assurance, Data Quality Control/Monitoring, Data Quality Assessment, Transparency/Traceability, and Data Integrity key components. Each of these components is ranked from ‘Ad hoc’ to ‘Optimal’ (see Appendix I). This report is based on evaluation performed by NOAA OneStop metadata specialists working with Subject Matter Experts and utilizing the DSMM template [Peng, 2016].

1.2 Scope

Assessing stewardship maturity - the current state of how datasets are documented, preserved, stewarded, and made accessible publicly, is a critical step towards meeting U.S. federal regulations, organizational requirements, and user needs [Peng et al., 2016]. The goal of this document is to provide consistent and transparent stewardship maturity information to data users and decision-makers.

1.3 Dataset Abstract

The Global Historical Climatology Network - Daily (GHCN-Daily) dataset integrates daily climate observations from approximately 30 different data sources. Version 3 was released in September 2012 with the addition of data from two additional station networks. Changes to the processing system associated with the version 3 release also allowed for updates to occur 7 days a week rather than only on most weekdays. Version 3 contains station-based measurements from well over 90,000 land-based stations worldwide, about two thirds of which are for precipitation measurement only. Other meteorological elements include, but are not limited to, daily maximum and minimum temperature, temperature at the time of observation, snowfall and snow depth. Over 25,000 stations are regularly updated with observations from within roughly the last month. The dataset is also routinely reconstructed (usually every week) from its roughly 30 data sources to ensure that GHCN-Daily is generally in sync with its growing list of constituent sources.

During this process, quality assurance checks are applied to the full dataset. Where possible, GHCN-Daily station data are also updated daily from a variety of data streams. Station values for each daily update also undergo a suite of quality checks.

1.4 Document Maintenance

This document is generated and maintained by NOAA's National Centers for Environmental Information. More on policy is available at <https://www.ncei.noaa.gov/>.

2. Results

The data stewardship maturity assessment information is summarized in Table 1. Each component is displayed along with its corresponding score in a color-coded table.

Table 2. Dataset and Data Stewardship Maturity Assessment Metadata	
Dataset Title	Global Historical Climatology Network - Daily (GHCN-Daily), Version 3
Dataset Information URL	http://doi.org/10.7289/V5D21VHZ
Data Provider POC (Name; Email; Affiliation)	National Centers for Environmental Information, NESDIS, NOAA, U.S. Department of Commerce 301-713-3277 NCEI.Info@noaa.gov
Dataset POC (Name; Email; Affiliation)	GHCN-Daily Operations Team; ncdc.ghcnd@noaa.gov; DOC/NOAA/NESDIS/NCEI > National Centers for Environmental Information, NESDIS, NOAA, U.S. Department of Commerce
SMM Version (Document ID and Version Number)	NCDC-CICS-SMM_0001_Rev.1 12/09/2014
SMM POC (Name; E-mail; Affiliation)	Ge Peng, ge.peng@uah.edu, University of Alabama-Huntsville
SMM Template Version (Document ID and Version Numbers)	NCDC-CICS-SMM_0001_Rev.1 v4.0 06/23/2015
SMM Template POC	Ge Peng, ge.peng@uah.edu, University of Alabama-Huntsville
SMM Assessment Version (v<nn>r<mm>, e.g., v01r00)	v00r10
SMM Assessment Date (MM/DD/YYYY)	07/14/2017
SMM Assessment POC (Name; E-mail; Affiliation)	Raisa Ionin, raisa.ionin@noaa.gov, Earth Resources Technology, Inc.
Stewardship Maturity Ratings (each key component) (kc1/kc2/kc3/kc4/kc5/kc6/kc7/kc8/kc9)	4.5 /4 /2.5 /4.5 /4.5 /5 /3 /2.75 /2
SMM Original Assessment Date (MM/DD/YYYY)	07/14/2017
SMM Original Assessment POC (Name; E-mail; Affiliation)	Raisa Ionin, raisa.ionin@noaa.gov, Earth Resources Technology, Inc.
SMM Last Modified Date (MM/DD/YYYY)	12/17/2021
SMM Last Modification POC (Name; E-mail; Affiliation)	Katy Luquire, catherine.luquire@noaa.gov , CASE Consultants International
SMM Modified Date (MM/DD/YYYY)	07/14/2017
SMM Modification POC (Name; E-mail; Affiliation)	Raisa Ionin, raisa.ionin@noaa.gov, Earth Resources Technology, Inc.

Table 3. Stewardship Maturity Levels and Detailed Justifications for Each of Nine DSMM Key Components for the Dataset.

DSMM Key Component	Stewardship Maturity Rating, Justification, and Comments
<p>Preservability</p>	<p>Level 4.5</p> <ul style="list-style-type: none"> ▪ Archived at NOAA NCEI-NC which is a NOAA designated archive with built in redundancy. ▪ Archive follows OAIS RM. ▪ Conforms to ISO 19115-2 metadata standard. ▪ Plans to transition ISO metadata to newer 19115-1 standard. ▪ No known audits performed on the archive at this time. <p>Comments:</p>
<p>Accessibility</p>	<p>Level 4</p> <ul style="list-style-type: none"> ▪ Collection level searchable online ▪ Granule level is searchable online ▪ Additional search options available from collection level site ▪ Direct file download available from ▪ LAS: ▪ THREDDS: ▪ OPeNDAP: ▪ HTTP: ▪ FTP: ▪ More search options are available: https://www.ncdc.noaa.gov/cdo-web/search?datasetid=GHCND ▪ https://gis.ncdc.noaa.gov/maps/ncei/summaries/daily ▪ Monthly dissemination reports are available but only internally. Future technology changes are planned <p>Comments:</p> <p>If dissemination become reports publicly available, then the rating should be changed to 4.5</p>
<p>Usability</p>	<p>Level 2.5</p> <ul style="list-style-type: none"> ▪ The format is NOT YET interoperable: No NetCDF or nc.gz/nc.bz for granules ▪ Project website contains Source description: https://www.ncei.noaa.gov/products/land-based-station/global-historical-climatology-network-daily ▪ Documentation: ▪ Link to the Stations Inventory List: https://www1.ncdc.noaa.gov/pub/data/ghcn/daily/ghcnd-stations.txt ▪ Link to Daily Status Reports: https://www1.ncdc.noaa.gov/pub/data/ghcn/daily/status.txt ▪ [NOAA] is available online https://www1.ncdc.noaa.gov/pub/data/cdo/documentation/GHCND_documentation.pdf ▪ Error estimates and algorithm information are described in [Durre, 2010] and available online https://doi.org/10.1175/2010JAMC2375.1 ▪ No community metrics of data ▪ No external ranking ▪ No enhanced data visualization. <p>Comments:</p> <p>If the format becomes interoperable, then rate at 3</p>

Table 3. Stewardship Maturity Levels and Detailed Justifications for Each of Nine DSMM Key Components for the Dataset.

DSMM Key Component	Stewardship Maturity Rating, Justification, and Comments
<p>Production Sustainability</p>	<p>Level 4.5</p> <ul style="list-style-type: none"> ▪ The data set is currently operational and supported by NCEI. The GHCN-D product web site: http://doi.org/10.7289/V5D21VHZ ▪ GHCN-D has international commitment – US contribution to World Meteorological Organization Global Climate Observing System (WMO GCOS) program. ▪ Changes for technology are planned. <p>Comments: If changes for technology are available, then rate at 4.5.</p>
<p>Data Quality Assurance</p>	<p>Level 4.5</p> <ul style="list-style-type: none"> ▪ DQA procedure is monitored and reported ▪ File level quality flags exist which can be considered Limited Data Quality Assurance Metadata. ▪ Sample Record of Climatological Observations [NCEI, NOAA, 2015] is available online https://www.ncdc.noaa.gov/cdo-web/search?datasetid=GHCND ▪ Conforms to community quality metadata and standards: ISO metadata standards. Metadata XML contains Data Quality <gmd:DQ_DataQuality> section ▪ Limited Data Quality Assurance Metadata exists in txt format: ▪ https://www1.ncdc.noaa.gov/pub/data/ghcn/daily/status.txt ▪ All information above is based on the following publications: ▪ [Durre, 2008] is available online https://doi.org/10.1175/2007JAMC1706.1 ▪ [Durre, 2010] is available online https://doi.org/10.1175/2010JAMC2375.1 ▪ [Menne, 2012] is available online https://doi.org/10.1175/JTECH-D-11-00103.1 ▪ Also based on the project website https://www.ncdc.noaa.gov/ghcn-daily-description ▪ No external review. <p>Comments:</p>

Table 3. Stewardship Maturity Levels and Detailed Justifications for Each of Nine DSMM Key Components for the Dataset.

DSMM Key Component	Stewardship Maturity Rating, Justification, and Comments
<p>Data Quality Control/ Monitoring</p>	<p>Level 5</p> <ul style="list-style-type: none"> ▪ Sampling and analysis are frequent, systematic, and automatic. ▪ Anomaly detection procedure is well documented and reported ▪ Qualifies for Limited quality monitoring metadata ▪ Conforms to community quality metadata and standards: ▪ ISO metadata standards and contains Data Quality <gmd:DQ_DataQuality> section ▪ Limited Quality Monitoring Metadata also exists in txt format: https://www1.ncdc.noaa.gov/pub/data/ghcn/daily/status.txt ▪ Procedure documented and available online ▪ Physical consistency check - (GHEN-D) provides a complete list of multiple consistencies checks in the Quality Control paragraph on this web page https://www.ncdc.noaa.gov/ghcn-daily-methods ▪ Quality checks of the data are done on a daily basis and statistics are calculated weekly ▪ Cross-validation of temporal & spatial characteristics is present– based on publications ▪ All information above is based on the following publications: ▪ [Durre, 2008] is available online https://doi.org/10.1175/2007JAMC1706.1 ▪ [Durre, 2010] is available online https://doi.org/10.1175/2010JAMC2375.1 ▪ [Menne, 2012] is available online https://doi.org/10.1175/JTECH-D-11-00103.1 ▪ And also based on the project website https://www.ncdc.noaa.gov/ghcn-daily-description <p>▪ Dynamic providers/users’ feedback in place – email and phone number are provided at the bottom of the landing page, “Dataset Point of Contact” section.</p> <p>Comments:</p>
<p>Data Quality Assessment</p>	<p>Level 3</p> <ul style="list-style-type: none"> ▪ Algorithm information is available: ▪ [Menne, 2012] is available online https://doi.org/10.1175/JTECH-D-11-00103.1 ▪ [Durre, 2010] is available online https://doi.org/10.1175/2010JAMC2375.1 ▪ Research and Operational products are assessed based on the following publications: ▪ [Durre, 2008] is available online https://doi.org/10.1175/2007JAMC1706.1 ▪ [Durre, 2010] is available online https://doi.org/10.1175/2010JAMC2375.1 ▪ [Menne, 2012] is available online https://doi.org/10.1175/JTECH-D-11-00103.1 ▪ No external ranking <p>Comments:</p>
<p>Transparency / Traceability</p>	<p>Level 2.75</p> <ul style="list-style-type: none"> ▪ Product information is available in literature: ▪ [Durre, 2008] is available online https://doi.org/10.1175/2007JAMC1706.1 ▪ [Durre, 2010] is available online https://doi.org/10.1175/2010JAMC2375.1 ▪ [Menne, 2012] is available online https://doi.org/10.1175/JTECH-D-11-00103.1 ▪ GHEN-D is not currently under Configuration Management. ▪ DOI assigned:10.7289/V5D21VHZ ▪ OID, Unique Object Identifier, is assigned: NCEI DSI 9101_01 ▪ Additional OID, Unique Object Identifier, is assigned by NCEI: gov.noaa.ncdc:C00861 <p>Comments:</p>

Table 3. Stewardship Maturity Levels and Detailed Justifications for Each of Nine DSMM Key Components for the Dataset.	
DSMM Key Component	Stewardship Maturity Rating, Justification, and Comments
Data Integrity	<p>Level 2</p> <ul style="list-style-type: none"> ▪ Data ingest integrity verifiable - a checksum is stored in the archive catalog for each data file in the archive. <p>Comments: This checksum can be provided upon request, however, neither this nor other checksums are put online with the data.</p>

3. Acknowledgment

This work is supported by the NOAA OneStop Project.

We thank the dataset POCs for their valuable input, as well as the collaborative efforts of the OneStop teams, especially the Metadata team. We would also like to show appreciation to Ge Peng for her contributions.

The draft of this data stewardship maturity report is systematically generated by a tool created by Kieran Hodnett and populated with the stewardship maturity assessment done by the author(s) of this report. The tool was developed based on a Word template created collaboratively by Robert Partee II, Raisa Ionin, Paul Lemieux III, Ge Peng, Don Collins, and Sonny Zinn with helpful input from the NOAA Central Library and the NCEI Communication Team.

4. References

Casey, K. (2016), The NOAA OneStop data discover and access framework project, Version: June 3, 2016. <https://cdn.ioos.noaa.gov/media/2017/12/OneStop-IOOS-DMAC-03-June-2016.pdf>

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.

Peng, G., J.L. Privette, E.J. Kearns, N.A. Ritchey, and S. Ansari (2015), A unified framework for measuring stewardship practices applied to digital environmental datasets, *Data Science Journal*, 13, 231-253, doi: 10.2481/dsj.14-049.

Peng, G., J. Lawrimore, V. Toner, C. Lief, R. Baldwin, N. Ritchey, D. Brinegar, and S. A. Delgreco (2016) assessing stewardship maturity of the global historical climatology network-monthly (GHCN-M) dataset: use case study and lessons learned, *D-Lib Magazine*, 22, doi:10.1045/november2016-peng.

Global Historical Climatology Network - Daily (GHCN-Daily Data Documentation https://www1.ncdc.noaa.gov/pub/data/cdo/documentation/GHCND_documentation.pdf (Accessed 18 July, 2017)

Durre, I., M. J. Menne, B. E. Gleason, T. G. Houston, and R. S. Vose (2010), Comprehensive Automated Quality Assurance of Daily Surface Observations, *Journal of Applied Meteorology and Climatology*, 49(8), 1615–1633, doi:10.1175/2010jamc2375.1

National Centers for Environmental Information (NCEI), National Oceanic and Atmospheric Administration (NOAA) (2015), Record of Climatological Observations, retrieved online <https://www.ncdc.noaa.gov/cdo-web/search?datasetid=GHCND>. (Accessed 17 December, 2021)

Durre, I., M. J. Menne, and R. S. Vose (2008), Strategies for Evaluating Quality Assurance Procedures, *Journal of Applied Meteorology and Climatology*, 47(6), 1785–1791, doi:10.1175/2007jamc1706.1

Durre, I., M. J. Menne, B. E. Gleason, T. G. Houston, and R. S. Vose (2010), Comprehensive Automated Quality Assurance of Daily Surface Observations, *Journal of Applied Meteorology and Climatology*, 49(8), 1615–1633, doi:10.1175/2010jamc2375.1

Menne, M. J., I. Durre, R. S. Vose, B. E. Gleason, and T. G. Houston (2012), An Overview of the Global Historical Climatology Network-Daily Database, *Journal of Atmospheric and Oceanic Technology*, 29(7), 897–910, doi:10.1175/jtech-d-11-00103.1

Appendix I: The Scientific Data Stewardship Maturity Matrix (DSMM)

Table A1: This matrix (Version: NCDC-CICS-SMM-0001-Rev.1. 12/09/2014) describes the criterion used to evaluate data stewardship maturity for each of the nine DSMM key components [Peng *et al.*, 2015].

DSMM Component	Level 1 <i>Ad hoc</i> Little or no management	Level 2 <i>Minimal</i> Limited management	Level 3 <i>Intermediate</i> Defined management, partially implemented	Level 4 <i>Advanced</i> Well-defined management, fully implemented	Level 5 <i>Optimal</i> Full management, audited, measured, controlled
<i>Preservability</i> <i>(The state of being preservable)</i>	Any storage location Data only	Non-designated repository Redundancy Limited archiving metadata	Designated archive Redundancy Community-standard archiving metadata Conforming to limited archiving standards	Level 3 + Conforming to community archiving standards	Level 4 + Archiving process performance controlled, measured, and audited Future archiving standard changes planned
<i>Accessibility</i> <i>(The state of being searchable and accessible publicly)</i>	Not publically available person-to-person	Publically available direct file download (e.g., via anonymous FTP server) Collection or dataset level searchable online	Level 2 + Non-standard data service Limited data server performance Granule/file level searchable Limited search metrics	Level 3 + Community-standard data service Enhanced data server performance Conforming to community search metrics Dissemination report metrics defined and implemented internally	Level 4 + Dissemination reports available online Future technology and standard changes planned

<p>Usability</p> <p><i>(The state of being easy to use)</i></p>	<p>Extensive product-specific knowledge required</p> <p>No documentation online</p>	<p>Non-standard data format</p> <p>Limited documentation (e.g., user's guide online)</p>	<p>Community standard-based interoperable format & metadata</p> <p>Documentation (e.g. source code, product algorithm document, processing or/and data flow diagram) online</p>	<p>Level 3 +</p> <p>Basic capability (e.g., subsetting, aggregating) & data characterization overall/global,</p> <p>e.g., climatology, error estimates) available online</p>	<p>Level 4 +</p> <p>Enhanced online capability (e.g., visualization, multiple data formats)</p> <p>Community metrics of data characterization (regional/cell) online</p> <p>External ranking</p>
<p>Production Sustainability</p> <p><i>(The state of data production being sustainable and extendable)</i></p>	<p>Ad Hoc or Not applicable</p> <p>To obligation or deliverable requirement</p>	<p>Short-term</p> <p>Individual PI's commitment (grant obligations)</p>	<p>Medium-term</p> <p>Institutional commitment (contractual deliverables with specs and schedule defined)</p>	<p>Long-term Institutional commitment</p> <p>Product improvement process in place</p>	<p>Level 4 +</p> <p>National or international commitment</p> <p>Changes for echnology planned</p>
<p>Data Quality Assurance</p> <p><i>(The state of data quality being assured)</i></p>	<p>Data quality assurance (DQA) procedure unknown or none</p>	<p>Ad Hoc and random</p> <p>QA procedure not defined and documented</p>	<p>DQA procedure defined and documented and partially implemented</p>	<p>DQA procedure well documented, fully implemented and available online with master reference data</p> <p>Limited data quality assurance metadata</p>	<p>Level 4 +</p> <p>DQA procedure monitored and reported</p> <p>Conforming to community quality metadata & standards</p> <p>External review</p>

<p>Data Quality Control/Monitoring</p> <p><i>The state of data quality being controlled and monitored</i></p>	<p>None or Sampling unknown or spotty</p> <p>Analysis unknown or random in time</p>	<p>Sampling and analysis are regular in time and space</p> <p>Limited product-specific metrics defined & implemented</p>	<p>Level 2 +</p> <p>Sampling and analysis are frequent and systematic but not automatic</p> <p>Community metrics defined and partially implemented</p> <p>Procedure documented and available online</p>	<p>Level 3 +</p> <p>Anomaly detection procedure well-documented and fully implemented using community metrics, automatic, tracked and reported</p> <p>Limited quality monitoring metadata</p>	<p>Level 4 +</p> <p>Cross-validation of temporal & spatial characteristics</p> <p>Physical consistency check</p> <p>Conforming to community quality metadata & standards</p>
<p>Data Quality Assessment</p> <p><i>(The state of data quality being assessed)</i></p>	<p>Algorithm/method/model</p> <p>Theoretical basis assessed (methods and results online)</p>	<p>Level 1 +</p> <p>Research product assessed (methods and results online)</p>	<p>Level 2 +</p> <p>Operational product assessed (methods and results online)</p>	<p>Level 3 +</p> <p>Quality metadata assessed</p> <p>Limited quality assessment metadata</p>	<p>Level 4 +</p> <p>Assessment performed on a recurring basis</p> <p>Conforming to community quality metadata & standards</p> <p>External ranking</p>
<p>Transparency/Traceability</p> <p><i>(The state of being transparent, trackable, and traceable)</i></p>	<p>Limited product information available</p> <p>Person-to-person</p>	<p>Product information available in literature</p>	<p>Algorithm Theoretical Basis Document (ATBD) & source code online</p> <p>Dataset configuration managed (CM)</p> <p>Unique Object Identifier (OID) assigned (dataset, documentation, source code)</p> <p>Data citation tracked (e.g., utilizing Digital Object Identifier</p>	<p>Level 3 +</p> <p>Operational Algorithm Description (OAD) online, OID assigned, and under CM</p>	<p>Level 4 +</p> <p>System information online</p> <p>Complete data provenance online</p>

<p>Data Integrity</p> <p><i>(The state of data integrity being verifiable)</i></p>	Unknown or no data ingest integrity check	Data ingest integrity verifiable (e.g., checksum technology)	(DOI) system) Level 2 + Data archive integrity verifiable	Level 3 + Data access integrity verifiable Conforming to community data integrity technology standard	Level 4 + Data authenticity verifiable (e.g., data signature technology) Performance of data integrity check monitored and reported