Stewardship Maturity Matrix (SMM) as of <09/04/2015>
for <GHCN-M-v3>

Dataset Title	Global Historical Climatology Network-Monthly, Version 3
Dataset Information URL	http://www.ncdc.noaa.gov/ghcnm/v3.php; http://dx.doi.org/10.7289/V5X34VDR
Data Provider POC (Name; E-mail; Affiliation)	Jay Lawrimore; Jay.Lawrimore@noaa.gov; NCEI/CWC/CSD/DSB
Dataset POC (Name; E-mail; Affiliation)	Jay Lawrimore; Jay.Lawrimore@noaa.gov; NCEI/CWC/CSD/DSB
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@noaa.gov; Cooperative Institute for Climate and Satellites, North Carolina (CICS-NC), North Carolina State University (NCSU) & NOAA's National Centers for Environmental Information (NCEI) [1]
SMM Template Version (Document ID and Version Numbers)	NCDC-CICS-SMM_0001_Rev.1 v4.0 06/23/2015
SMM Template POC (Name; E-mail; Affiliation)	Ge Peng; Ge.Peng@noaa.gov; Cooperative Institute for Climate and Satellites, North Carolina (CICS-NC), North Carolina State University (NCSU) & NOAA's National Centers for Environmental Information (NCEI)
SMM Assessment Version (v<nn>r<mm>, e.g., v01r00)	v01r03
SMM Assessment Date (MM/DD/YYYY)	09/04/2015
SMM Assessment POC (Name; E-mail; Affiliation)	Jay Lawrimore; Jay.Lawrimore@noaa.gov; NCEI/CWC/CSD/DSB; Valerie Toner; valerie.toner@noaa.gov; NCEI/DSD/AB; Christina Lief; Christina.Lief@noaa.gov; NCEI/DSD/AB; Ge Peng; Ge.Peng@noaa.gov; NCEI/CWC/CSD/PRB & CICS-NC; Rich Baldwin; Rich.Baldwin@noaa.gov; NCEI/DSD/DAB
Stewardship Maturity Ratings (kc1/kc2/kc3/kc4/kc5/kc6/kc7/kc8/kc9)	4.0/2.0/2.5/4.5/3.5/2.5/3.0/2.5/3.5
SMM Original Assessment Date (MM/DD/YYYY)	06/08/2015
SMM Original Assessment POC (Name; E-mail; Affiliation)	Valerie Toner; valerie.toner@noaa.gov; Archive Specialist, Contractor with Team ERT/STG, an affiliate of NOAA's National Centers for Environmental Information (NCEI)
SMM Last Modified Date (MM/DD/YYYY)	09/04/2015
SMM Last Modification POC (Name; E-mail; Affiliation)	Rich Baldwin; Rich.Baldwin@noaa.gov; NCEI/DSD/DAB
SMM Modified Date (MM/DD/YYYY)	07/22/2015

SMM Modification POC (Name; E-mail; Affiliation)	Ge Peng; Ge.Peng@noaa.gov; NCEI/CWC/CSD/PRB & CICS-NC; Jay Lawrimore; Jay.Lawrimore@noaa.gov; NCEI/CWC/CSD/DSB; Christina Lief; Christina.Lief@noaa.gov; NCEI/DSD/AB; Valerie Toner; valerie.toner@noaa.gov; NCEI/DSD/AB
SMM Modified Date (MM/DD/YYYY)	07/06/2015
SMM Modification POC (Name; E-mail; Affiliation)	Jay Lawrimore; Jay.Lawrimore@noaa.gov; NCEI/CWC/CSD/DSB
SMM Modified Date (MM/DD/YYYY) [2]	06/25/2015
SMM Modification POC (Name; E-mail; Affiliation)	Christina Lief; Christina.Lief@noaa.gov; NCEI/DSD/AB

[1] NCEI includes the organizations previously referred to as National Climatic Data Center (NCDC), National Geophysical Data Center (NGDC), and National Oceanographic Data Center (NODC).

[2] Repeat these last two lines to capture the SMM modification history

Maturity Scale (across)	**Level 1** **Ad Hoc** **Not Managed**	**Level 2** **Minimal** **Managed Limited**	**Level 3** **Intermediate** **Managed Defined, Partially Implemented**	**Level 4** **Advanced** **Managed Well-Defined, Fully Implemented**	**Level 5** **Optimal** **Level 4 + Measured , Controlled , Audit**		
Key Component (below)						**Stewardship Maturity Rating And Justification or Evidence**	**Comments**
Preservability *(The state of being preservable)*	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	❖ ***Level 4*** - Archived at NCEI-NC (Designated NOAA data center that is compliant to NARA archive standards.) - Conforming to the NCEI-NC archiving process and guideline that are following OAIS RM - Compliant to NCEI-NC defined archive procedure and requirement set forth by Submission Agreement (SA) - Offsite backup copy available - Collection level Metadata conforming to ISO 19115 metadata standards: (http://gis.ncdc.noaa.gov/geoportal/catalog/search/resource/details.jsp?id=gov.noaa.ncdc:C00839) –	

Accessibility *(The state of being searchable and accessible publicly)*	Not publicly available Person-to-person	Publicly available Direct file download (e.g., via anonymous FTP server) Collection/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	❖ **Level 2** - direct file download via ftp server (http://www.ncdc.noaa.gov/ ghcnm/v3.php) - data and metadata are ASCII files, space delimited, in gzip'd tar files. (1 file containing all station data and 1 file containing all station metadata) - Collection-level searchable (Google, NCEI, NOAA Catalog, Geoportal) but not searchable at file-level - **Info needed on if dissemination report available internally**	Next version will be at level 3 (metadata will be in the Historical Observing Metadata Repository (HOMR) and data will be provided via Climate Data Online (CDO) portal)
Usability *(The state of being easy to use)*	Extensive product-specific knowledge required No documentation online	Non-standard data format Limited documentation (e.g., user's guide) online	Community standard-based interoperable format & metadata Documentation (e.g., source code, product algorithm document, processing or/and data flow diagram) online	Level 3 + Basic capability (e.g., subsetting, aggregating) & data characterization (overall/global, e.g., climatology, error estimates) available online	Level 4 + Enhanced online capability (e.g., visualization, multiple data formats) Community metrics of data characterization (regional/cell) online External ranking	❖ **Level 2.5** - README file online - Product algorithm reference list is online - Some source code online - All source code/data flow/process flow diagram not online - Data and file-level metadata are in ASCII format which is in situ community supported format - The current ASCII files are not self-describing	Next version will be at level 3 (with all source code, data flow and process flow diagrams) with self-describing data format.
Production Sustainability *(The state of data production being sustainable and extendable)*	Ad Hoc or Not applicable No obligation or deliverable requirement	Short-term Individual PI's commitment (grant obligations)	Medium-term Institutional commitment (contractual deliverables with specs and schedule defined)	Long-term Institutional commitment Product improvement process in place	Level 4 + National or international commitment Changes for technology planned	❖ **Level 4.5** - Long-term institutional and international commitment - Data are being updated regularly - Product improvement in place - Product under version control	
Data Quality Assurance *(The state of data quality being assured)*	Data quality assurance (DQA) procedure unknown or none	Ad Hoc and random DQA procedure not defined and documented	DQA procedure defined and documented and partially implemented	DQA procedure well documented, fully implemented and available online with master reference data Limited data quality assurance metadata	Level 4 + DQA procedure monitored and reported Conforming to community quality metadata & standards External review	❖ **Level 3.5** - DQA procedure defined in JGR-Atmospheres journal article (http://onlinelibrary.wiley.com/doi/10.1029/2011JD016187/abstract) and also described online (https://www.ncdc.noaa.gov/ghcnm/v3.php under "Data Assurance" tab) - Quality assurance procedures fully implemented - Community metrics are produced and made available online (ftp://ftp.ncdc.noaa.gov/pub/data/ghcn/v3/products/) - No data quality assurance metadata	
Data Quality Control/Monitoring	None or Sampling	Sampling and analysis are regular	Level 2+ Sampling and analysis are	Level 3 + Anomaly detection	Level 4 + Cross-validation of temporal	❖ **Level 2.5** - Quality flagged and statistics metrics are online at: http://www.ncdc.noaa.gov/ghcnm/v3.php	For next version, documentation on the procedure(s) will be online – level 3

(The state of data quality being controlled and monitored)	unknown or spotty Analysis unknown or random in time	in time and space Limited product-specific metrics defined & implemented	frequent and systematic but not automatic Community metrics defined and partially implemented Procedure documented and available online	procedure well-documented and fully implemented using community metrics, automatic, tracked and reported Limited quality monitoring metadata	& spatial characteristics Physical consistency check Conforming to community quality metadata & standards Dynamic providers/users feedback in place	- Regular monthly manual reviews of automatically generated plots or statistics are conducted. - Quality monitoring metrics are consistent with in situ community - Procedure is not documented and available online - No data quality control/monitoring metadata	
Data Quality Assessment *(The state of data quality being assessed)*	Algorithm/method/model theoretical basis assessed (methods and results online)	Level 1 + Research product assessed (methods and results online)	Level 2 + Operational product assessed (methods and results online)	Level 3 + Quality metadata assessed Limited quality assessment metadata	Level 4 + Assessment performed on a recurring basis Conforming to community quality metadata & standards External ranking	❖ **Level 3** - Information on the product algorithm and data quality assessment procedures is available in the JGR-Atmospheres article (http://onlinelibrary.wiley.com/doi/10.1029/2011JD016187/abstract) and online at http://www.ncdc.noaa.gov/ghcnm/v3.php. - Assessment of the operational product, i.e., GHCN-Monthly version 3.x, was done comparing with other datasets and included in the latest IPCC report (http://www.ipcc.ch/pdf/assessment-report/ar5/wg1/WG1AR5_Chapter02_FINAL.pdf) - No data quality assessment metadata	
Transparency /Traceability *(The state of being transparent, trackable, and traceable)*	Limited product information available Person-to-person	Product information available in literature	Algorithm Theoretical Basis Document (ATBD) & source code online Dataset configuration managed (CM) Unique Object Identifier (OID) assigned (dataset, documentation, source code) Data citation tracked (e.g., utilizing Digital Object Identifier (DOI) system)	Level 3 + Operational Algorithm Description (OAD) online, OID assigned, and under CM	Level 4 + System information online Complete data provenance online	❖ **Level 2.5** - Product information in the literature (http://onlinelibrary.wiley.com/doi/10.1029/2011JD016187/abstract) - Dataset ID is assigned (NCDC DSI 9100_03) and under CM - dataset doi is assigned and tracked (http://dx.doi.org/10.7289/V5X34VDR) - Detailed summary of each software modification and the resulting impacts to global temperatures is available at: ftp://ftp.ncdc.noaa.gov/pub/data/ghcn/v3/techreports/Technical Report NCDC No112_02_3.2.0-29Aug12.pdf; ftp://ftp.ncdc.noaa.gov/pub/data/ghcn/v3/techreports/Technical Report GHCNM No15-01.pdf - The Pairwise Homogeneity Adjustment algorithm software is available online at ftp://ftp.ncdc.noaa.gov/pub/data/ghcn/v3/software/	Descriptive Product Information Document will be generated and available for next version – Level 3
Data Integrity *(The state of data integrity being verifiable)*	Unknown or no data ingest integrity check	Data ingest integrity verifiable (e.g., checksum technology)	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	❖ **Level 3.5** - Data integrity is checked at ingest, archive, and dissemination using the check-sum technology - No check-sum available online for users to verify data files at access	Recommend including checksum/MANIFEST on ftp when staging the data files

SMM Document ID: NCDC-CICS-SMM_0001
Version: Rev. 1. 12/09/2014

Global Historical Climatology Network-Monthly

Data Stewardship Maturity Scoreboard

Maturity Level as of 09/04/2015

Maturity Scale	Preservability	Accessibility	Usability	Production Sustainability	Data Quality Assurance	Data Quality Control/Monitoring	Data Quality Assessment	Transparency /Traceability	Data Integrity
Level 1 – Ad Hoc **Not Managed**	Any storage location Data only	Not publicly available Person-to-person	Extensive product-specific knowledge required No documentation online	Ad Hoc or Not applicable No obligation or deliverable requirement	Data quality assurance (DQA) procedure unknown or none	None or Sampling unknown or spotty Analysis unknown or random in time	Algorithm/method/model theoretical basis assessed (method and results online)	Limited product information available Person-to-person	Unknown or no data ingest integrity check
Level 2 - Minimal **Managed Limited**	Non-designated repository Redundancy Limited archiving metadata	Publicly available Direct file download (e.g., via anonymous FTP server) Collection/dataset level searchable	Non-standard data format Limited documentation (e.g., user's guide) online	Short-term Individual PI's commitment (grant obligations)	Ad Hoc and random DQA procedure not defined and documented	Sampling and analysis are regular in time and space Limited product-specific metrics defined & implemented	Level 1 + Research product assessed (method and results online)	Product information available in literature	Data ingest integrity verifiable (e.g., checksum technology)
Level 3 - Intermediate **Managed Defined, Partially Implemented**	Designated archive Redundancy Community-standard archiving metadata Conforming to limited archiving process standards	Level 2 + Non-standard data service Limited data server performance Granule/file level searchable Limited search metrics	Community Standard-based interoperable format & metadata Documentation (e.g., source code, product algorithm document, processing or/and data flow diagram) online	Medium-term Institutional commitment (contractual deliverables with specs and schedule defined)	DQA procedure defined and documented and partially implemented	Level 2 + Sampling and analysis are frequent and systematic but not automatic Community metrics defined and partially implemented Procedure documented and available online	Level 2 + Operational product assessed (method and results online)	Algorithm Theoretical Basis Document (ATBD) & source code online Dataset configuration managed (CM) Unique Object Identifier (OID) assigned (dataset, documentation, source code) Data citation tracked (e.g., utilizing Digital Object Identifier (DOI) system)	Level 2 + Data archive integrity verifiable
Level 4 - Advanced **Managed Well-Defined, Fully Implemented**	Level 3 + Conforming to community archiving standards	Level 3 + Community-standard data services Enhanced data server performance Conforming to community search metrics Dissemination report metrics defined and implemented internally	Level 3 + Basic capability (e.g., subsetting, aggregating) & data characterization (overall/global, e.g., climatology, error estimates) available online	Long-term Institutional commitment Product improvement process in place	DQA procedure well documented, fully implemented and available online with master reference data Limited data quality assurance metadata	Level 3 + Anomaly detection procedure well-documented and fully implemented using community metrics, automatic, tracked and reported Limited quality monitoring metadata	Level 3 + Quality metadata assessed (method and results online) Limited quality assessment metadata	Level 3 + Operational Algorithm Description (OAD) online, OID assigned, and under CM	Level 3 + Data access integrity verifiable Conforming to community data integrity technology standard
Level 5 - Optimal **Level 4 + Measured, Controlled, Audit**	Level 4 + Archiving process performance controlled, measured, and audited Future archiving standard changes planned	Level 4 + Dissemination reports available online Future technology and standard changes planned	Level 4 + Enhanced online capability (e.g., visualization, multiple data formats) Community metrics of data characterization (regional/cell) online External ranking	Level 4 + National or international commitment Changes for technology planned	Level 4 + DQA procedure monitored and reported Conforming to community quality metadata & standards External review	Level 4 + Cross-validation of temporal & spatial characteristics Physical consistency check Conforming to community quality metadata & standards Dynamic providers/users feedback in place	Level 4 + Assessment performed on a recurring basis Conforming to community quality metadata & standards External ranking	Level 4 + System information online Complete data provenance available online	Level 4 + Data authenticity verifiable (e.g., data signature technology) Performance of data integrity check monitored and reported

Dataset Information: **http://dx.doi.org/10.7289/V5X34VDR**
Dataset POC: **Jay Lawrimore; Jay.Lawrimore@noaa.gov**

SMM POC: **Ge Peng; Ge.Peng@noaa.gov**
SMM Assessment POC: **Jay Lawrimore; Jay.Lawrimore@noaa.gov**

Figure 1: Data stewardship maturity scoreboard of GHCN-M-v3.
If two cells are filled, it is an indication that only a partial maturity rating at the higher level is satisfied.

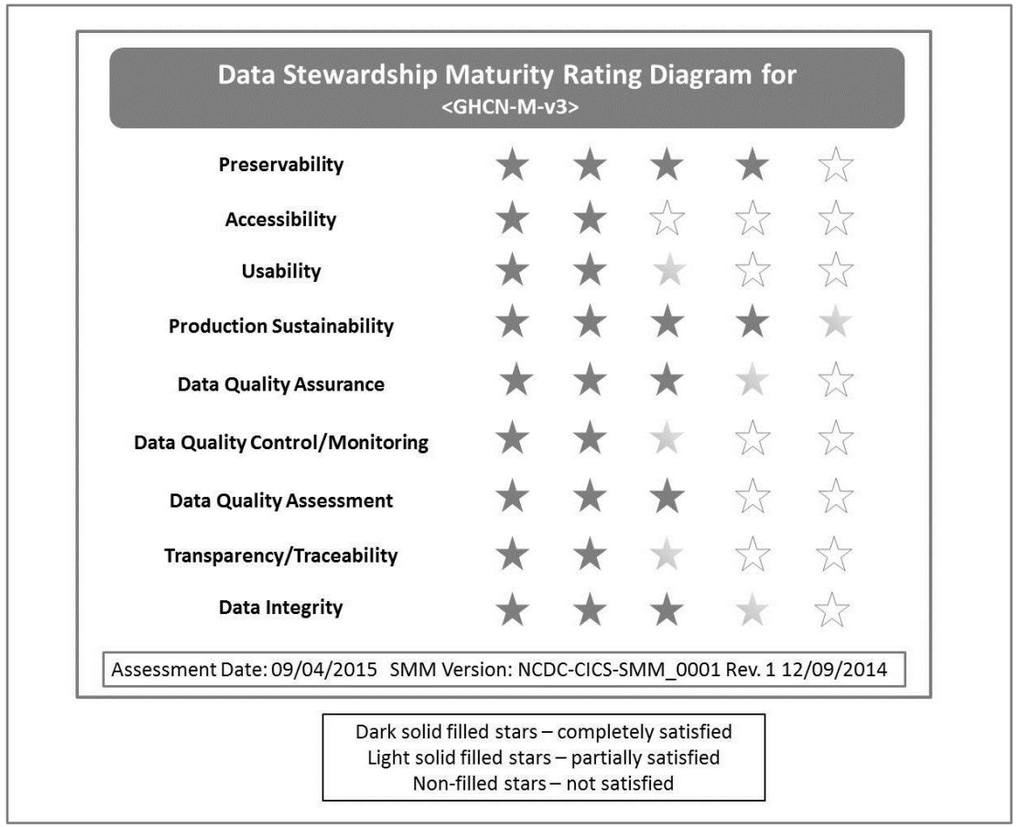

Figure 2: Data stewardship maturity rating diagram of GHCN-M-v3.