

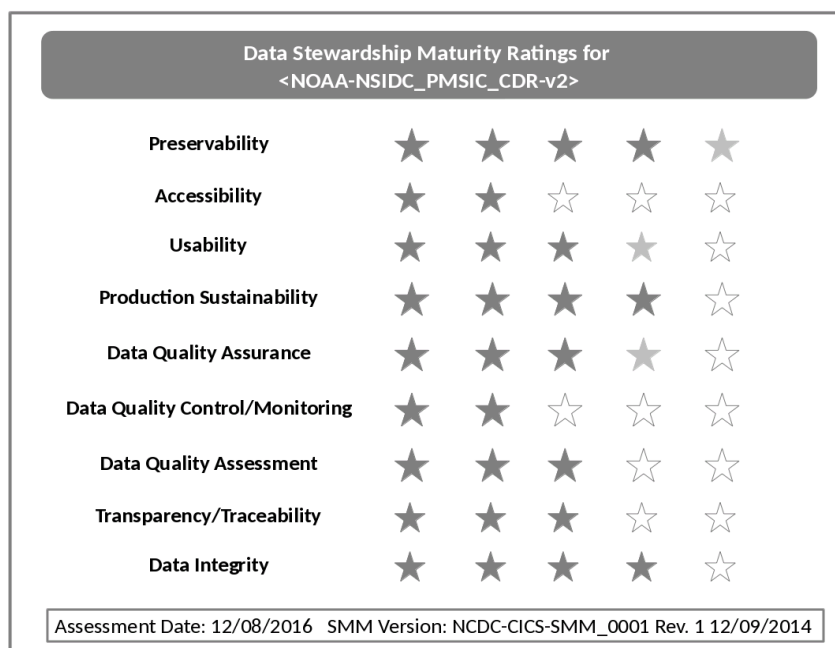
Data Stewardship Maturity Report for NOAA Climate Data Record (CDR) of Passive Microwave Sea Ice Concentration, Version 2

Paul Lemieux III¹, Ge Peng², and Donna J. Scott³

¹ ERT, Inc. and NOAA's National Centers of Environmental Information (NCEI)
151 Patton Avenue, Asheville, NC 28801, USA

² Cooperative Institute for Climate and Satellites - North Carolina (CICS-NC), North Carolina
State University and NCEI, Asheville, NC, USA

³ National Snow and Ice Data Center, University of Colorado, Boulder, CO, USA



Dark solid filled stars – completely satisfied
 Light solid filled stars – partially satisfied
 Non-filled stars – not satisfied

Cover Image: Data stewardship rating diagram for <NOAA-NSIDC_PMSIC_CDR-v2>. One to five stars are used to represent Level 1 to 5 ratings, denoting Ad Hoc, Minimal, Intermediate, Advance, and Optimal stages for each of the nine key components, respectively. The dark filled stars indicate that all the practices are completely satisfied. The light filled ones indicated that not all the practices are satisfied. And the non-filled ones indicated that the practices are not satisfied.

The stewardship maturity of NCEI data product, <NOAA-NSIDC_PMSIC_CDR-v2>, is assessed based on a reference stewardship maturity framework. The current maturity ratings of <NOAA-NSIDC_PMSIC_CDR-v2> are at Level 2 or higher for all nine key components with two Level 2, four Level 3, and three Level 4 key components.

ASSESSMENT REVISION HISTORY

Revision	Description	Date
V03r00	Data stewardship maturity report for NOAA Climate Data Record (CDR) of Passive Microwave Sea Ice Concentration, Version 2.	12/08/2016

Recommended Citation:

Lemieux III, P., Peng, G., and Scott, D. J. (2017), Data stewardship maturity report for NOAA Climate Data Record (CDR) of Passive Microwave Sea Ice Concentration, Version 2. *Figshare*. 20 pp. doi: [10.6084/m9.figshare.5279932](https://doi.org/10.6084/m9.figshare.5279932)

Contents

List of Figures	4
List of Tables	5
Preface	6
1. Introduction	7
2. Results	8
3. Acknowledgment	15
4. References	15
Appendix I: The Scientific Data Stewardship Maturity Matrix (DSMM)	17

List of Figures

Figure [1](#). Data stewardship maturity scoreboard for <NOAA-NSIDC_PMSIC_CDR-v2>, highlighted with 5-level progressive green scales for each of the nine key components (across), representing Ad Hoc, Minimal, Intermediate, Advance, and Optimal stages (vertical). If more than two cells are highlighted, it denotes that the dataset has completely satisfied the criterion for the lower level but not yet so at the current level.

Figure [2](#). Data stewardship rating diagram for <NOAA-NSIDC_PMSIC_CDR-v2>. One to five stars are used to represent Level 1 to 5 ratings, denoting Ad Hoc, Minimal, Intermediate, Advance, and Optimal stages for each of the nine key components, respectively. The dark filled stars indicate that all the practices are completely satisfied. The light filled ones indicated that not all the practices are satisfied. And the non-filled ones indicated that the practices are not satisfied. The current maturity ratings of <NOAA-NSIDC_PMSIC_CDR-v2> are at Level 2 or higher for all nine key components with two Level 2, four Level 3, and three Level 4 key components.

List of Tables

Table [1](#). Dataset and Data Stewardship Maturity Assessment Metadata.

Table [2](#). Stewardship Maturity Levels and Detailed Justifications for Each of Nine DSMM Key Components for the <NOAA-NSIDC_PMSIC_CDR-v2> Dataset.

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.*, 2015]. 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 NOAA Climate Data Record (CDR) of Passive Microwave Sea Ice Concentration, Version 2

1. Introduction

1.1 Purpose

The purpose of this document is to describe the results of stewardship maturity assessment for NOAA Climate Data Record (CDR) of Passive Microwave Sea Ice Concentration, Version 2, utilizing the Scientific Data Stewardship Maturity Matrix or *DSMM* [Peng, *et al*, 2016]. DSMM defines 5 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, 2015].

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 the consistent and transparent stewardship maturity information to data users and decision-makers.

1.3 Dataset Outline

This data set provides a Climate Data Record (CDR) of sea ice concentration from passive microwave data. It provides a consistent, daily and monthly time series of sea ice concentrations from 09 July 1987 through the most recent processing for both the north and south polar regions on a 25 km x 25 km polar stereographic grid.

The NOAA/NSIDC CDR is based on the recommendations from the National Research Council (NRC) (2004). It is produced from gridded brightness temperatures from the Defense Meteorological Satellite Program (DMSP) F8, F11, and F13 Special Sensor Microwave Imager (SSM/I) passive microwave radiometers and the DMSP F17 Special Sensor Microwave Imager/Sounder (SSMIS) passive microwave radiometer.

Variables containing standard deviation, quality flags, and projection information are also included in the netCDF files. Data are available via FTP.

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 information about dataset and stewardship maturity assessment is summarized in Table 1. The data stewardship maturity ratings are displayed as the scoreboard (Figure 1) and rating diagram (Figure 2) with the detailed justifications in Table 2.

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

**NOAA Climate Data Record (CDR) of Passive Microwave Sea Ice Concentration,
Version 2**

Maturity Level as of
12/08/2016

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 Granular/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: <https://dx.doi.org/10.7265/N55M63M1>
Dataset POC: NOAA Climate Data Record Program Office, sea_ice_concentration_contacts@noaa.gov

SMM POC: Ge Peng: Ge.Peng@noaa.gov
SMM Assessment POC: Paul Lemieux III, paul.lemieux@noaa.gov

Figure 1. Data stewardship maturity scoreboard for <NOAA-NSIDC_PMSIC_CDR-v2>, highlighted with 5-level progressive green scales for each of the nine key components (across), representing Ad Hoc, Minimal, Intermediate, Advance, and Optimal stages (vertical). If more than two cells are highlighted, it denotes that the dataset has completely satisfied the criterion for the lower level but not yet so at the current level.

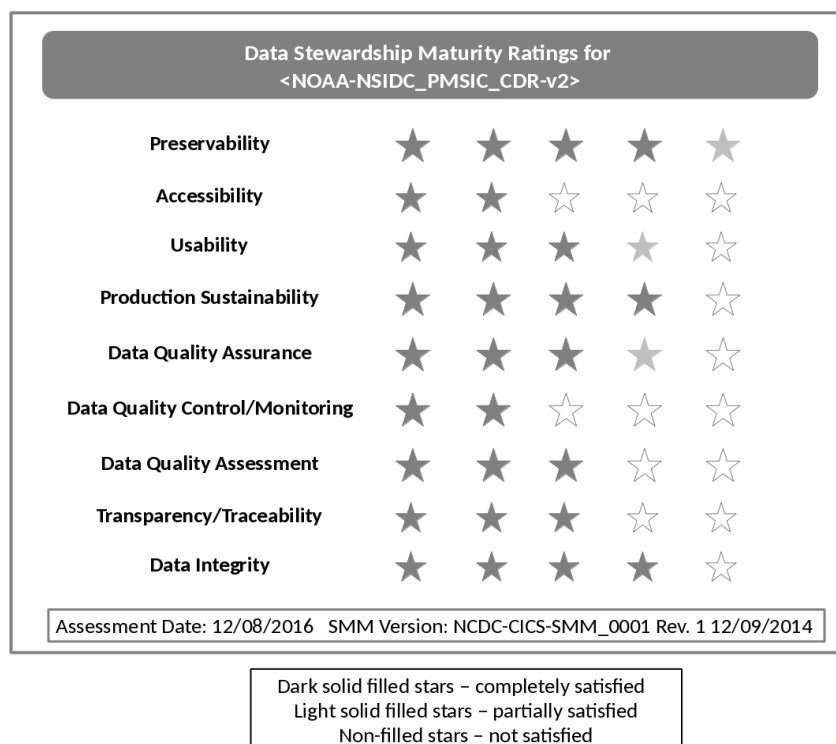


Figure 2. Data stewardship rating diagram for <NOAA-NSIDC_PMSIC_CDR-v2>. One to five stars are used to represent Level 1 to 5 ratings, denoting Ad Hoc, Minimal, Intermediate, Advance, and Optimal stages for each of the nine key components, respectively. The dark filled stars indicate that all the practices are completely satisfied. The light filled ones indicated that not all the practices are satisfied. And the non-filled ones indicated that the practices are not satisfied. The current maturity ratings of <NOAA-NSIDC_PMSIC_CDR-v2> are at Level 2 or higher for all nine key components with two Level 2, four Level 3, and three Level 4 key components.

Table 1. Dataset and Data Stewardship Maturity Assessment Metadata.

Dataset Title	NOAA Climate Data Record (CDR) of Passive Microwave Sea Ice Concentration, Version 2
Dataset Information URL	https://dx.doi.org/10.7265/N55M63M1 ; http://nsidc.org/data/G02202/
Data Provider POC (Name; E-mail; Affiliation)	NOAA National Centers for Environmental Information, ncei.orders@noaa.gov ,
Dataset POC (Name; E-mail; Affiliation)	NOAA Climate Data Record Program Office, sea_ice_concentration_contacts@noaa.gov ,
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	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)	V03r00
SMM Assessment Date (MM/DD/YYYY)	12/08/2016
SMM Assessment POC (Name; E-mail; Affiliation)	Paul Lemieux III, paul.lemieux@noaa.gov , NOAA National Centers for Environmental Information (NCEI)
Stewardship Maturity Ratings (each key component) (kc1/kc2/kc3/kc4/kc5/kc6/kc7/kc8/kc9)	4.5/2.0/3.5/4.0/3.5/2.0/3.0/3.0/4.0
SMM Original Assessment Date (MM/DD/YYYY)	10/16/2014
SMM Original Assessment POC (Name; E-mail; Affiliation)	Donna Scott, dscott@nsidc.org , NSIDC; Ge Peng, Ge.Peng@noaa.gov , Cooperative Institute for Climate and Satellites North Carolina (CICS-NC) North Carolina State University (NCSU) & NOAA National Centers for Environmental Information (NCEI)
SMM Last Modified Date (MM/DD/YYYY)	12/08/2016
SMM Last Modification POC	Paul Lemieux III, Paul.Lemieux@noaa.gov , NOAA

(Name; E-mail; Affiliation)	National Centers for Environmental Information
SMM Modified Date (MM/DD/YYYY)	12/08/2016
SMM Modification 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 National Centers for Environmental Information (NCEI)

Table 2. Stewardship Maturity Levels and Detailed Justifications for Each of Nine DSMM Key Components for the <NOAA-NSIDC_PMSIC_CDR-v2> Dataset.

DSMM Key Component	Stewardship Maturity Rating, Justification, and Comments
<i>Preservability</i>	<p style="text-align: center;">❖ Level 4.5</p> <p>Justification:</p> <ul style="list-style-type: none"> • Archived at NOAA NCEI • Following NOAA Climate Data Record (CDR) Research-2-Operation (R2O) transition process with the Initial Operation Capability (IOC) • Following ISO OAIS RM • Conforming to ISO 19115-2 metadata standards • Conforming to NetCDF CF metadata conventions • Conforming to CDR Program (CDRP) guidelines on coding and NCEI Archive Branch (AB) guidance on file and variable naming conventions per Submission Agreement (SA) • Plans to transition ISO metadata to newer 19115-1 standard <p>Comment:</p> <ul style="list-style-type: none"> • No known audits of the archiving processes
<i>Accessibility</i>	<p style="text-align: center;">❖ Level 2.0</p> <p>Justification:</p> <ul style="list-style-type: none"> • Public ftp site: ftp://sidads.colorado.edu/pub/DATASETS/NOAA/G02202_v2 • Collection searchable online • Reports available internally for the FTP/HTTP servers • New technology for OneStop search and discovery planned (i.e. ElasticSearch, Hyrax Servers, etc.) This is part of the CDR data group that will be OneStop ready. <p>Comment:</p> <ul style="list-style-type: none"> • Dissemination reports are available internally, but are not available online
<i>Usability</i>	<p style="text-align: center;">❖ Level 3.5</p> <p>Justification:</p> <ul style="list-style-type: none"> • NetCDF-4 data format (CF compliant) • Data Flow Diagram [<i>Fetterer and NOAA CDR Program</i>, 2015] is available online here: https://www.ncdc.noaa.gov/cdr/oceanic/sea-ice-concentration • C-ATBD [<i>Meier and Windnagel</i>, 2015] is available online here: https://www.ncdc.noaa.gov/cdr/oceanic/sea-ice-concentration • Error assessment and estimates available in the C-ATBD [<i>Meier and Windnagel</i>, 2015] available online here: https://www.ncdc.noaa.gov/cdr/oceanic/sea-ice-concentration • Some data characterization on the global scale available in literature [<i>DeRepentigny, Tremblay, Newton, et al</i>, 2016] which is available online here: https://dx.doi.org/10.3402/polar.v33.21004

DSMM Key Component	Stewardship Maturity Rating, Justification, and Comments
	<p>Comment:</p> <ul style="list-style-type: none"> • No subsetting or aggregating options in place • No known external ranking
Production Sustainability	<p style="text-align: center;">❖ Level 4.0</p> <p>Justification:</p> <ul style="list-style-type: none"> • Under NOAA CDR Operation & Maintenance (O&M) • Updated annually • Funding is allocated yearly to the CDR program • Product improvement (versioning) process in place <p>Comment:</p> <ul style="list-style-type: none"> • No comments
Data Quality Assurance	<p style="text-align: center;">❖ Level 3.5</p> <p>Justification:</p> <ul style="list-style-type: none"> • Agile development procedure in place with a defined/fixed set of analysis metrics • Master reference data are included in the source code package available online: https://www.ncdc.noaa.gov/cdr/oceanic/sea-ice-concentration • Weather filters • SST Mask • Cell data quality flag in each data file • No data quality assurance metadata implemented <p>Comment:</p> <ul style="list-style-type: none"> • No known external review
Data Quality Control/Monitoring	<p style="text-align: center;">❖ Level 2.0</p> <p>Justification:</p> <ul style="list-style-type: none"> • DQC is done after each data processing • Sampling is regular in space but no automatic • Procedure not documented or available online <p>Comment:</p> <ul style="list-style-type: none"> • No data quality information in the metadata record
Data Quality Assessment	<p style="text-align: center;">❖ Level 3.0</p> <p>Justification:</p> <ul style="list-style-type: none"> • Sea ice concentration retrieval algorithms (NASA Team and Bootstrap) have been

DSMM Key Component	Stewardship Maturity Rating, Justification, and Comments
	<p>validated extensively with many published peer-review papers</p> <ul style="list-style-type: none"> • So are research products (Goddard products) for this dataset • Verification of the dataset is done and described in literature [<i>DeRepentigny, Tremblay, Newton, et al</i>, 2016] available online here: https://dx.doi.org/10.3402/polar.v33.21004 <p>Comment:</p> <ul style="list-style-type: none"> • No data quality assessment information I the metadata record
Transparency/ Traceability	<p style="text-align: center;">❖ Level 3.0</p> <p>Justification:</p> <ul style="list-style-type: none"> • Product information available in literature [<i>Peng, Meier, Scott, et al</i>, 2013] which is available online here: https://dx.doi.org/10.5194/essd-5-311-2013 • C-ATBD [<i>Meier and Windnagel</i>, 2015] available online here: https://www.ncdc.noaa.gov/cdr/oceanic/sea-ice-concentration • DOI assigned: https://dx.doi.org/10.7265/N55M63M1 • NCEI OID assigned: DSI-3628-02 • Dataset Configuration Management is EIA-649-B standard compliant and diagramed in this presentation document [<i>Hutchins</i>, 2015] which is available online here: http://www1.ncdc.noaa.gov/pub/data/sds/cdr/conferences/2015%20PI%20Annual%20Meeting%20-%20Presentations/Day_1/(A-2)%20Operations%20and%20Maintenance%20(O_M)%20of%20NOAA%20IOC%20CDRs%20-%20(Hutchins).pdf <p>Comment:</p> <ul style="list-style-type: none"> • No comments
Data Integrity	<p style="text-align: center;">❖ Level 4.0</p> <p>Justification:</p> <ul style="list-style-type: none"> • Checksum is created by NSIDC for each month tar file staged for NCEI • NCEI ingest validates each file base on checksum before archive • NSIDC generates the checksum for each data file and online for user to verify data integrity <p>Comment:</p> <ul style="list-style-type: none"> • No comments

3. Acknowledgment

This work is supported by NOAA OneStop Project. We thank beneficial input from dataset POC(s) and collaborative effort by OneStop Teams, especially the Metadata Team. Guidance from Ge Peng on DSMM was beneficial.

The draft of this data stewardship maturity report is systematically generated by a tool created by Sonny Zinn, 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, Donald Collins, and Sonny Zinn with beneficial input from NOAA Central Library and NCEI Communication Team.

4. References

Casey, K. (2016), The NOAA OneStop data discover and access framework project, <https://ioos.noaa.gov/wp-content/uploads/2016/06/OneStop-IOOS-DMAC-03-June-2016.pdf>.

DeRepentigny, P., Tremblay, L., Newton, R., and Pfirman, S., (2016), Patterns of sea ice retreat in the transition to a seasonally ice-free arctic, *Journal of Climate*, 29(9), 6993—7008, doi:10.3402/polar.v33.21004.

Fetterer and NOAA CDR Program, (2015), Sea Ice Concentration – annual CDR (01B-11) data flow diagram, *Rep. CDRP-DIA-0183 Rev 3*, NOAA National Centers for Environmental Information, Asheville, NC., retrieved online: <https://www.ncdc.noaa.gov/cdr/oceanic/sea-ice-concentration> (Accessed 19 December 2016).

Hutchins, C. (2015), Operations and Maintenance (O&M) of NOAA IOC CDRs, [http://www1.ncdc.noaa.gov/pub/data/sds/cdr/conferences/2015%20PI%20Annual%20Meeting%20-%20Presentations/Day_1/\(A-2\)%20Operations%20and%20Maintenance%20\(O_M\)%20of%20NOAA%20IOC%20CDRs%20-%20\(Hutchins\).pdf](http://www1.ncdc.noaa.gov/pub/data/sds/cdr/conferences/2015%20PI%20Annual%20Meeting%20-%20Presentations/Day_1/(A-2)%20Operations%20and%20Maintenance%20(O_M)%20of%20NOAA%20IOC%20CDRs%20-%20(Hutchins).pdf) (Accessed 19 December 2016).

Meier, W., and Windnagel, A., (2015), Climate Algorithm Theoretical Basis Document (C-ATBD) Sea Ice Concentration, *Rep. CDRP-ATBD-0107*, NOAA National Centers for Environmental Information, Asheville, NC., retrieved online: <https://www.ncdc.noaa.gov/cdr/oceanic/sea-ice-concentration> (Accessed 19 December 2016).

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: <http://dx.doi.org/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.

Peng, G., Meier, W., Scott, D., and Savoie, M., (2013), A long-term and reproducible passive microwave sea ice concentration data record for climate studies and monitoring, *Earth System Science Data*, 5, 311—318, doi:10.5194/essd-5-311-2013.

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> (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
<i>Accessibility</i> (The state of being searchable and accessible publicly)	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
<i>Usability</i> (The state of being easy to use)	Extensive product-specific knowledge required	Non-standard data format Limited	Community standard-based interoperable format &	Level 3 + Basic capability (e.g., subsetting,	Level 4 + Enhanced online capability (e.g.,

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
	No documentation online	documentation (e.g., user's guide online)	metadata Documentation (e.g. source code, product algorithm document, processing or/and data flow diagram) online	aggregating) & data characterization overall/global, e.g., climatology, error estimates) available online	visualization, multiple data formats) Community metrics of data characterization (regional/cell) online External ranking
<i>Production Sustainability</i> (The state of data production being sustainable and extendable)	Ad Hoc or Not applicable To 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 echnology planned
<i>Data Quality Assurance</i> (The state of data quality being assured)	Data quality assurance (DQA) procedure unknown or none	Ad Hoc and random QA 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
<i>Data Quality Control/ Monitoring</i> The state of data quality being controlled and monitored	None or Sampling unknown or spotty Analysis unknown or random in time	Sampling and analysis are regular in time and space Limited product-specific metrics defined & implemented	Level 2 + Sampling and analysis are frequent and systematic but not automatic Community metrics defined and partially	Level 3 + Anomaly detection procedure well-documented and fully implemented using community metrics,	Level 4 + Cross-validation of temporal & spatial characteristics Physical consistency check

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
			implemented Procedure documented and available online	automatic, tracked and reported Limited quality monitoring metadata	Conforming to community quality metadata & standards
<i>Data Quality Assessment</i> <i>(The state of data quality being assessed)</i>	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
<i>Transparency/Traceability</i> <i>(The state of being transparent, trackable, and traceable)</i>	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
<i>Data Integrity</i> <i>(The state of data)</i>	Unknown or no data ingest integrity check	Data ingest integrity verifiable (e.g.,	Level 2 + Data archive	Level 3 + Data access	Level 4 + Data

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>integrity being verifiable)</i>		checksum technology)	integrity verifiable	integrity verifiable Conforming to community data integrity technology standard	authenticity verifiable (e.g., data signature technology) Performance of data integrity check monitored and reported