NOAA Technical Information Series NESDIS
DSMR-00109 Version 1.0

doi: 10.25923/8ke1-0b74

Data Stewardship Maturity Report for GHRSST Level 4 REMO_OI_SST_5km Regional Foundation Sea Surface Temperature Analysis (GDS version 2)

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 - 5	Accessibility - 5	Usability - 4.5
Production Sustainability - 5	Data Quality Assurance - 1	Data Quality Control/Monitoring - 1
Data Quality Assessment - 0	Transparency/Traceability - 1	Data Integrity - 3

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 GHRSST Level 4 REMO_OI_SST_5km Regional Foundation Sea Surface Temperature Analysis (GDS version 2)

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, GHRSST Level 4 REMO_OI_SST_5km Regional Foundation Sea Surface Temperature Analysis (GDS version 2), is assessed based on a reference stewardship maturity framework. The current maturity ratings of GHRSST Level 4 REMO_OI_SST_5km Regional Foundation Sea Surface Temperature Analysis (GDS version 2) are at Level 1 or higher for all nine key components with three Level 1, zero Level 2, one Level 3, one Level 4, and three Level 5 key components.

NOAA Technical Information Series NESDIS DSMR-00109 Version 1.0

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	12/16/2021

NOAA Technical Information Series NESDIS DSMR-00109

Version 1.0

doi: 10.25923/8ke1-0b74

Data Stewardship Maturity Report for GHRSST Level 4 REMO_OI_SST_5km Regional Foundation Sea Surface Temperature Analysis (GDS version 2)

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 GHRSST Level 4 REMO_OI_SST_5km Regional Foundation Sea Surface Temperature Analysis (GDS version 2)

NOAA Technical Information Series NESDIS DSMR-00109 Version 1.0, 20pp.,
doi: 10.25923/8ke1-0b74

Table of Contents

List of Tables 7

Preface 8

1. Introduction 9

2. Results 10

3. Acknowledgment 15

4. References 16

Appendix I The Scientific Data Stewardship Maturity Matrix (DSMM) 17

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.

NOAA Technical Information Series NESDIS DSMR-00109 Version 1.0

Data Stewardship Maturity Report for GHRSST Level 4 REMO_OI_SST_5km Regional Foundation Sea Surface Temperature Analysis (GDS 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 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

A Group for High Resolution Sea Surface Temperature (GHRSST) Level 4 sea surface temperature (SST) analysis produced daily on an operational basis by the Oceanographic Modeling and Observation Network (REMO) at Applied Meteorology Laboratory/Federal University of Rio de Janeiro (LMA/UFRJ) using the Barnes sub optimal interpolation (OI) technique on a regional 0.05 degree grid. REMO uses Advanced Very High Resolution Radiometer (AVHRR) data from National Oceanic and Atmospheric Administration (NOAA) satellites series (NOAA 15, NOAA 16, NOAA 17, NOAA 18 and NOAA 19) and Microwave Imager (TMI) data from Tropical Rainfall Measuring Mission (TRMM) which is a joint mission between NASA and the Japan Aerospace Exploration Agency (JAXA) to generate 0.05 degree daily cloud free blended (infrared and microwave) SST products (approximately 5.5 km). The data lies between latitudes 45 S and 15 N and longitudes 70 W and 15 W region and are fully validated by in situ measurements from eleven buoys of Prediction and Research Moored Array in the Tropical Atlantic (PIRATA).

AVHRR is a scanning radiometer capable of detecting energy from land, ocean and atmosphere. It operates with six spectral bands arranged in the regions of visible and infrared region. TRMM was launched in December, 1997, having an orbital inclination of 53 degree and altitude 350 km, an equatorial orbit that ranges from 40 N to 40 S and a spatial resolution of 0.25 degree (~27.75 km). Although infrared AVHRR SST data have high spatial resolution, they are contaminated by cloud cover and aerosols, while lower resolution microwave TMI data are barely influenced by these.

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	GHRSST Level 4 REMO_OI_SST_5km Regional Foundation Sea Surface Temperature Analysis (GDS version 2)
Dataset Information URL	https://www.ncei.noaa.gov/metadata/geoportal/rest/metadata/item/gov.noaa.nodc%3AGHRSST-REMO_OI_SST_5km-UFRJ-L4-SAMERICA/html
Data Provider POC (Name; Email; Affiliation)	National Centers for Environmental Information, NESDIS, NOAA, U.S. Department of Commerce301-713-3277NCEI.Info@noaa.gov
Dataset POC (Name; Email; Affiliation)	Rosa Paes; rosa@lma.ufrj.br;LMA/UFRJ, Applied Meteorology Laboratory/Federal University of Rio de Janeiro
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)	v01r05
SMM Assessment Date (MM/DD/YYYY)	04/17/2019
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)	5/5/4.5/5/1/1/0/1/3
SMM Original Assessment Date (MM/DD/YYYY)	06/21/2016
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)	11/08/2021
SMM Last Modification POC (Name; E-mail; Affiliation)	Katy Luquire, catherine.luquire@noaa.gov , CASE Consultants International
SMM Modified Date (MM/DD/YYYY)	04/17/2019
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
Preservability	Level 5 ▪ Archived by NCEI, which is NOAA designated repository. NOAA is compliant to NARA standards ▪ Metadata following ISO 19115-2 standards. ▪ Compliant to OIAS RM ▪ Plans to update metadata to ISO 19115-1 at a later date ▪ Using NCEI Silver Spring Archive Management System, AMS. Comments:
Accessibility	Level 5 ▪ Collection level searchable online ▪ Granule level is searchable online ▪ Additional search options available from collection level site ▪ Direct file download available from ▪ THREDDS: ▪ https://www.ncei.noaa.gov/thredds-ocean/catalog/ghrsst/L4/SAMERICA/UFRJ/REMO_OI_SST_5km/catalog.html ▪ HTTP: https://www.ncei.noaa.gov/data/oceans/ghrsst/L4/SAMERICA/UFRJ/REMO_OI_SST_5km/ ▪ FTP: ftp://ftp-oceans.ncei.noaa.gov/pub/data.nodc/ghrsst/L4/SAMERICA/UFRJ/REMO_OI_SST_5km/ ▪ Data citation is also available from NASA PODAAC site: https://podaac.jpl.nasa.gov/dataset/REMO_OI_SST_5km-UFRJ-L4-SAMERICA-v1.0 ▪ Dissemination reports are available to the public https://www.ncei.noaa.gov/access/ghrsst-long-term-stewardship-and-reanalysis-facility/ ▪ Future technology changes are planned Comments:
Usability	Level 4.5 ▪ The format is interoperable: nc. for granules ▪ User Guide [GHRSST, 2011] is available online https://www.nodc.noaa.gov/archive/arc0072/0123222/1.1/data/0-data/GHRSSTUserGuidev91.pdf ▪ All GHRSST collections have enhanced online capability (e.g., visualization, multiple data formats): TDS, DAP (*data servers maintained at NCEI); access from metadata main landing page. ▪ A GHRSST User Guide, Quick Start Guide, GHRSST Data Specification (GDS) manual, and other relevant documents describing GHRSST data sets can be found in the archive accession, Documentation for The Group for High Resolution Sea Surface Temperature (GHRSST) data archived at NODC (NODC Accession 0123222), https://www.ncei.noaa.gov/access/metadata/landing-page/bin/iso?id=gov.noaa.nodc:0123222 ▪ No external ranking ▪ No Algorithm or ATBD documents exist for GHRSST collections Comments: No algorithm documents available. Data providers have them, but these are not available at NCEI. Will be planned for the future.

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
Production Sustainability	Level 5 ▪ The dataset is currently supported, according to LTSRF Table: https://www.ncei.noaa.gov/access/ghrsst-long-term-stewardship-and-reanalysis-facility/ ▪ Long-term institutional commitment through UFRJ (Federal University of Rio de Janeiro) ▪ Long-term international commitment (GHRSST is an international group) ▪ Changes for technology are available from individual dataset producers. Comments: Changes for technology are available from individual dataset producers. NOAA does not have them documented. From LTSRF page, the product is listed under UFRJ (Federal University of Rio de Janeiro), then REMO_OI_SST_5km SAMERICA
Data Quality Assurance	Level 1 ▪ No documentation exists Comments:
Data Quality Control/ Monitoring	Level 1 ▪ No Quality Control metrics are available: https://www.star.nesdis.noaa.gov/sod/sst/squam/ Comments:
Data Quality Assessment	Level 0 ▪ No Algorithm Based Documentation (ATBD) available Comments:
Transparency / Traceability	Level 1 ▪ Limited product information available, metadata only on the GHRSST_L4_REMO_OI_SST_5km_RFSSTA landing page: https://www.ncei.noaa.gov/metadata/geoportal/rest/metadata/item/gov.noaa.nodc%3AGHRSST-REMO_OI_SST_5km-UFRJ-L4-SAMERICA/html Comments:

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	Level 3 ▪ Data archive integrity verifiable - Checksum technology is available, each GHRSST_L4_REMO_OI_SST_5km_RFSSTA package is accompanied by a manifest in XML format containing hash digests generated using various algorithms, including MD5, SHA-1, SHA-384, etc. That includes checksums (.md5) for every file package. https://www.nodc.noaa.gov/archive/arc0073/0125971/0125971.1.1.xml ▪ Data authenticity is verifiable (since data can be downloaded via HTTPS and HTTPS uses certificates to prove site authenticity) ▪ NCEI-MD does not provide digital signatures for data dissemination Comments: Example of a checksum file (.md5 file) can be also seen at ftp://podaac-ftp.jpl.nasa.gov/allData/ghrsst/data/GDS2/L4/SAMERICA/UFRJ/REMO_OI_SST_5km/v1/2010/025/ PODAAC ftp site: ftp://podaac-ftp.jpl.nasa.gov/allData/ghrsst/data/GDS2/L4/SAMERICA/UFRJ/REMO_OI_SST_5km/v1

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 naturity 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.

GHRSST User Guide version 9.1, 2011, retrieved online: https://www.nodc.noaa.gov/archive/arc0072/0123222/1.1/data/0-data/GHRSSTUserGuidev91.pdf (Accessed December 22, 2016)

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** ***Ad hoc*** Little or no management	**Level 2** ***Minimal*** Limited management	**Level 3** ***Intermediate*** Defined management, partially implemented	**Level 4** ***Advanced*** Well-defined management, fully implemented	**Level 5** ***Optimal*** Full management, audited, measured, controlled
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
Accessibility *(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

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
Production Sustainability *(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
Data Quality Assurance *(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

Data Quality Control/ Monitoring *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 implemented Procedure documented and available online	Level 3 + Anomaly detection procedure well-documented and fully implemented using community metrics, automatic, tracked and reported Limited quality monitoring metadata	Level 4 + Cross-validation of temporal & spatial characteristics Physical consistency check Conforming to community quality metadata & standards
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
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	Level 3 + Operational Algorithm Description (OAD) online, OID assigned, and under CM	Level 4 + System information online Complete data provenance online

Data Integrity *(The state of data integrity being verifiable)*	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