NOAA Technical Information Series NESDIS
DSMR-00132 Version 1.0

doi: 10.25923/jnnt-6f09

Data Stewardship Maturity Report for NOAA JPSS Visible Infrared Imaging Radiometer Suite (VIIRS) Cloud Mask Environmental Data Record (EDR) from NDE

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 - 5	Usability - 4.5
Production Sustainability - 5	Data Quality Assurance - 4.5	Data Quality Control/Monitoring - 4.75
Data Quality Assessment - 3.5	Transparency/Traceability - 3.5	Data Integrity - 5

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 NOAA JPSS Visible Infrared Imaging Radiometer Suite (VIIRS) Cloud Mask Environmental Data Record (EDR) from NDE

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, NOAA JPSS Visible Infrared Imaging Radiometer Suite (VIIRS) Cloud Mask Environmental Data Record (EDR) from NDE, is assessed based on a reference stewardship maturity framework. The current maturity ratings of NOAA JPSS Visible Infrared Imaging Radiometer Suite (VIIRS) Cloud Mask Environmental Data Record (EDR) from NDE are at Level 1 or higher for all nine key components with zero Level 1, zero Level 2, two Level 3, four Level 4, and three Level 5 key components.

NOAA Technical Information Series NESDIS DSMR-00132 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-00132

Version 1.0

doi: 10.25923/jnnt-6f09

Data Stewardship Maturity Report for NOAA JPSS Visible Infrared Imaging Radiometer Suite (VIIRS) Cloud Mask Environmental Data Record (EDR) from NDE

Lori Hager, Paul Lemieux III
NOAA's National Centers of Environmental Information (NCEI)
151 Patton Avenue, Asheville, NC 28801, (828) 271-4800

Recommended Citation

Lori Hager, Paul Lemieux III. (2021), Data Stewardship Maturity Report for NOAA JPSS Visible Infrared Imaging Radiometer Suite (VIIRS) Cloud Mask Environmental Data Record (EDR) from NDE

NOAA Technical Information Series NESDIS DSMR-00132 Version 1.0, 21pp.,
doi: 10.25923/jnnt-6f09

Table of Contents

List of Tables

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-00132 Version 1.0

Data Stewardship Maturity Report for NOAA JPSS Visible Infrared Imaging Radiometer Suite (VIIRS) Cloud Mask Environmental Data Record (EDR) from NDE

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

This data set contains a high quality Environmental Data Record (EDR) of cloud masks from the Visible Infrared Imaging Radiometer Suite (VIIRS) instrument onboard the Suomi-NPP satellite and is produced by the NOAA Satellite and Information Service (NESDIS). The EDR contains an image in the form of a binary cloud mask that identifies pixels within a coverage region as clear or cloudy. The production of the clear sky mask is an important step in the processing of many other VIIRS products that use the information generated in the production of the clear sky mask to determine the presence of clouds. Additionally, the product includes data quality information for the binary cloud mask data values for on-earth pixels. The algorithm utilizes 13 VIIRS bands from 0.415μm to 12.02μm including the VIIRS Day Night Band (DNB). The EDR provides global coverage approximately every 12 hours at a spatial resolution of 750 meters at nadir. VIIRS Cloud Mask data distributed by the NESDIS Data Exploitation (NDE) system and obtained from the Comprehensive Large-Array Stewardship System (CLASS) are distributed as single 86-second granules in NetCDF-4 format with metadata attributes included.

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	NOAA JPSS Visible Infrared Imaging Radiometer Suite (VIIRS) Cloud Mask Environmental Data Record (EDR) from NDE
Dataset Information URL	https://www.ncei.noaa.gov/metadata/geoportal/rest/metadata/item/gov.noaa.ncdc%3AC01434/html
Data Provider POC (Name; Email; Affiliation)	Customer Engagement Branch, ncei.sat.info@noaa.gov, DOC/NOAA/NESDIS/NCEI > National Centers for Environmental Information, NESDIS, NOAA, U.S. Department of Commerce
Dataset POC (Name; Email; Affiliation)	Andy Heidinger, andrew.heidinger@noaa.gov, DOC/NOAA/NESDIS/STAR > Center for Satellite Applications and Research, 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)	v01r03
SMM Assessment Date (MM/DD/YYYY)	09/22/2017
SMM Assessment POC (Name; E-mail; Affiliation)	Paul Lemieux III, paul.lemieux@noaa.gov, Earth Resources Technology, Inc.
Stewardship Maturity Ratings (each key component) (kc1/kc2/kc3/kc4/kc5/kc6/kc7/kc8/kc9)	4.5 / 5 / 4.5 / 5 / 4.5 / 4.75 / 3.5 / 3.5 / 5
SMM Original Assessment Date (MM/DD/YYYY)	06/09/2016
SMM Original Assessment POC (Name; E-mail; Affiliation)	Paul Lemieux III, paul.lemieux@noaa.gov, Earth Resources Technology, Inc.
SMM Last Modified Date (MM/DD/YYYY)	10/07/2021
SMM Last Modification POC (Name; E-mail; Affiliation)	Lori Hager, lori.hager@noaa.gov, CASE Consultants International
SMM Modified Date (MM/DD/YYYY)	04/24/2019
SMM Modification POC (Name; E-mail; Affiliation)	Paul Lemieux III, paul.lemieux@noaa.gov, Riverside 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 4.5 ▪ Archived by NCEI which is a NOAA designated archive compliant to NARA standards. ▪ Metadata following ISO 19115-2. ▪ Compliant to OIAS RM ▪ Plans to update metadata to ISO 19115-1 at a later date and may be a pilot dataset for the OneStop initiative. ▪ Using CLASS Comments: CLASS is CMMI-Level 3.
Accessibility	Level 5 ▪ Collection level searchable online ▪ CLASS ordering: https://www.class.noaa.gov/saa/products/search?datatype_family=JPSS_GRAN ▪ Granules searchable and orderable via CLASS ▪ Direct file download available CLASS FTP: ftp://ftp-jpss.avl.noaa.gov ▪ CLASS has dissemination reports available internally and externally. Users have to e-mail the CLASS Help Desk to request access to the metrics tools. ▪ New technology for OneStop search and discovery planned (i.e. ElasticSearch, Hyrax Servers, etc.) This is part of the JPSS data group that will be OneStop ready. Comments: A CLASS improvement (GEARS) is in the works.
Usability	Level 4.5 ▪ Satellite community standard NetCDF data format following ACDD and CF conventions. ▪ Data paper [Kopp, Thomas, et al, 2014] available online here: https://doi.org/10.1002/2013JD020458 ▪ CLASS has aggregating/subsetting options for downloads. ▪ Climatology and Error Estimates available in the ATBD [Heidinger, Botambekov, et al, 2016] available online here: https://www.star.nesdis.noaa.gov/jpss/Docs.php ▪ External User's Guide [Lindeman, 2017] is available online here: https://www.star.nesdis.noaa.gov/jpss/Docs.php#S739241 Comments: No known external rankings.
Production Sustainability	Level 5 ▪ Long-term institutional commitment from NESDIS ▪ International commitment from EUMETSAT and Taiwan Weather Service. ▪ Product will initially be processed in NDE, however NDE hardware will transition to PDA. Comments: No 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 Quality Assurance	Level 4.5 ▪ Quality assurance procedure documented in the ATBD [Heidinger, 2016] and User's Guide [Lindeman, 2017] available online here: https://www.star.nesdis.noaa.gov/jpss/Docs.php ▪ Data quality assurance plan [Mikles and Liu, 2016] is reviewed externally by stakeholders and is available here: https://www.star.nesdis.noaa.gov/jpss/Docs.php. ▪ Quality flags exist in the dataset which can be considered limited data quality assurance metadata. Comments: No comments.
Data Quality Control/ Monitoring	Level 4.75 ▪ OSPO PAL will perform product quality monitoring as part of the Product Monitoring project (per the User's Guide). ▪ Cross-validation with CALIOP LIDAR, MODIS, and the EUMETSAT cloud mask algorithm. ▪ Product quality is monitored by ESPC Ops and e-mail alerts are automatically generated when anomalies occur. ▪ Additional metadata file in XML format is generated with each file that contains statistical information that is used to monitor product data quality and processing status. It is used internally at OSPO by monitoring team. ▪ Users can contact the ESPC help desk 24/7 for information about the data product and they can resolve issues through coordination with the PAL. Comments: Cannot verify if there are any physical consistency checks.
Data Quality Assessment	Level 3.5 ▪ Beta & Provisional (Research) & Validated (Operational) product assessments are available in the algorithm maturity review documents available here: https://www.star.nesdis.noaa.gov/jpss/AlgorithmMaturity.php ▪ ATBD [Heidinger, 2016] available online here: https://www.star.nesdis.noaa.gov/jpss/Docs.php ▪ Operational product assessment [Heidinger, Foster, et al, 2016] available in literature and online here: https://doi.org/10.3390/rs8060511 ▪ Data quality assessment information in the auxiliary metadata file that is generated by the system. Comments: No known external rankings.

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
Transparency / Traceability	Level 3.5 ▪ ATBD [Heidinger, 2016] available online here: https://www.star.nesdis.noaa.gov/jpss/Docs.php ▪ DOI: 10.7289/TBD ▪ CLASS Assigned OID: VIIRS_CLM ▪ Product information available in literature [Heidinger, Foster, et al, 2016] and available online here: https://doi.org/10.3390/rs8060511 ▪ Configuration Management Plan [Zhao, 2014] available online here: https://www.star.nesdis.noaa.gov/jpss/Docs.php Comments: No OAD available.
Data Integrity	Level 5 ▪ CLASS offers data signatures option for downloads and checksums available. ▪ File level checksums are verified at ingest and at archive. ▪ CLASS maintains a copy of the checksum for validation during dissemination and that checksum is used for validation during the staging process when an order is fulfilled. ▪ Discrepancies identified in checksum validation are automatically reported to CLASS team for investigative purposes. Comments: No comments.

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.

Kopp, T., Thomas, W., et al, (2014), The VIIRS cloud mask: progress in the first year of S-NPP toward a common cloud detection scheme, _J. Geo. Res. Atm._, 119(5), 2441—2456, doi:10.1002/2013JD020458.

Heidinger, A., Botambekov, D., et al, (2016), Algorithm theoretical basis document a naïve Bayesian cloud mask delivered to NOAA enterprise version 1.1, NOAA Center for Satellite Applications and Research, College Park, MD, retrieved online: https://www.star.nesdis.noaa.gov/jpss/Docs.php (accessed 30 May 2017).

Lindeman, J., (2017), The NOAA JPSS Risk Reduction System External Users Manual Version 1.0, NOAA Center for Satellite and Application Research, College Park, MD, retrieved online: https://www.star.nesdis.noaa.gov/jpss/Docs.php# S739241, (accessed 2017 Sept 22).

Mikles, V., and Liu, X., (2016), STAR JPSS Quality Assurance Plan v1.0, NOAA Center for Satellite Applications and Research, College Park, MD, retrieved online: https://www.star.nesdis.noaa.gov/jpss/Docs.php (accessed 2017 May 24.

Heidinger, A., Foster, M., et al, (2016), Using the NASA EOS A-train to probe the performance of the NOAA PATMOS-x cloud fraction CDR, _Rem. Sen._, 8(6), 511—529, doi:10.3390/rs8060511.

Zhao, Y., (2014) STAR JPSS algorithms integration team configuration management plan, NOAA Center for Satellite Applications and Research (STAR), College Park, MD., retrieved online: https://www.star.nesdis.noaa.gov/jpss/Docs.php (Accessed 30 May 2017).

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