**GSFC ESDIS CMO**

**September 2020**

**Released**



National Aeronautics and

Space Administration

**Goddard Space Flight Center**

**Greenbelt, Maryland**

**Unified Metadata Model - Variable (UMM-Var)**

Signature/Approval Page

|  |  |  |
| --- | --- | --- |
| **Prepared by:** |  |  |
|  |  |  |
| Name |  | Date |
| Title/Role |  |  |
| Organization |  |  |
|  |  |  |
| **Reviewed by:** |  |  |
|  |  |  |
| Name |  | Date |
| Title/Role |  |  |
| Organization |  |  |
|  |  |  |
| **Approved by:** |  |  |
|  |  |  |
| Name |  | Date |
| Title/Role |  |  |
| Organization |  |  |
|  |  |  |
| **Concurred by:** |  |  |
|  |  |  |
| Name |  | Date |
| Title/Role |  |  |
| Organization |  |  |

**[Electronic] Signatures available in B32 Room E148**

**online at: /** **https://ops1-cm.ems.eosdis.nasa.gov/cm2/**

Preface

This document is under ESDIS Project configuration control. Once this document is approved, ESDIS approved changes are handled in accordance with Class I and Class II change control requirements described in the ESDIS Configuration Management Procedures, and changes to this document shall be made by change bars or by complete revision.

Any questions should be addressed to: esdis-esmo-cmo@lists.nasa.gov

ESDIS Configuration Management Office (CMO)  
NASA/GSFC

Code 423

Greenbelt, Md. 20771

Abstract

This document describes the Unified Metadata Model for Variables (UMM-Var) to be used by the National Aeronautics and Space Administration (NASA) Earth Science community and addresses the need for describing the types of variables that exist within data products that are described by the Unified Metadata Model for Granules (UMM-G) metadata records. Developers, engineers and architects should reference this document and the Unified Metadata Model (UMM) as a guide while implementing Common Metadata Repository (CMR) components, CMR clients or services that make use of the CMR or CMR clients. Data providers should use this model as a guide during metadata generation.

This version of the variable model focuses on hierarchical variables and what is the minimum variable metadata needed to support the User Interface/User Experience (UI/UX) leading to an improved user experience. Since there will be many thousands of variables in the CMR, it also supports the notion of auto-population of variable metadata records. This aims to reduce the workload on the metadata curator in their task to manage variable metadata over time.

***Keywords:*** UMM-Var, UMM-S, UMM-G, Variables, NASA Earthdata Search, EOSDIS, ESDIS, CMR

Change History Log

|  |  |  |
| --- | --- | --- |
| **Revision** | **Effective Date** | **Description of Changes**  (Reference the CCR & CCB Approval Date) |
|  |  | CCR 423-ESDIS-XXX; CCB Approved  Pages: |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

Table of Contents

[1 Introduction 1](#_Toc63081313)

[1.1 Purpose 2](#_Toc63081314)

[1.2 Scope 3](#_Toc63081315)

[1.3 Related Documentation 3](#_Toc63081316)

[1.3.1 Applicable Documents 3](#_Toc63081317)

[1.3.2 Reference Documents 3](#_Toc63081318)

[1.4 Impact 4](#_Toc63081319)

[1.5 Copyright Notice 4](#_Toc63081320)

[1.6 Feedback 4](#_Toc63081321)

[1.7 Document Conventions 4](#_Toc63081322)

[2 Unified Metadata Model - Variables 5](#_Toc63081323)

[2.1 Use Cases 5](#_Toc63081324)

[2.1.1 Browse Variables of a Collection 5](#_Toc63081325)

[2.1.2 Cross-site Data Subsetting 6](#_Toc63081326)

[2.2 UMM-Var Metadata Model 8](#_Toc63081327)

[2.2.1 Name [R] 10](#_Toc63081328)

[2.2.2 StandardName 17](#_Toc63081329)

[2.2.3 LongName [R] 18](#_Toc63081330)

[2.2.4 Definition [R] 18](#_Toc63081331)

[2.2.5 Units 19](#_Toc63081332)

[2.2.6 DataType 19](#_Toc63081333)

[2.2.7 Scale and Offset 21](#_Toc63081334)

[2.2.8 VariableType 21](#_Toc63081335)

[2.2.9 VariableSubType 22](#_Toc63081336)

[2.2.10 IndexRanges 22](#_Toc63081337)

[2.2.11 ScienceKeywords 22](#_Toc63081338)

[2.2.12 FillValues 23](#_Toc63081339)

[2.2.13 Dimensions 24](#_Toc63081340)

[2.2.14 ValidRanges 24](#_Toc63081341)

[2.2.15 MeasurementIdentifiers 25](#_Toc63081342)

[2.2.15.1 MeasurementContextMedium [R] 26](#_Toc63081343)

[2.2.15.2 MeasurementContextMediumURI 27](#_Toc63081344)

[2.2.15.3 MeasurementObject [R] 27](#_Toc63081345)

[2.2.15.4 MeasurementObjectURI 27](#_Toc63081346)

[2.2.15.5 MeasurementQuantities 27](#_Toc63081347)

[2.2.16 Sets 29](#_Toc63081348)

[2.2.17 SamplingIdentifiers 34](#_Toc63081349)

[2.2.18 AdditionalIdentifiers 34](#_Toc63081350)

[Appendix A Tags Glossary 36](#_Toc63081351)

[Appendix B Keywords and Measurements Governance Structure 37](#_Toc63081352)

[Appendix C Analysis of CSDMS and CF Standard Names as a Source of Tagging 39](#_Toc63081353)

[Appendix D Definitions of Terms 41](#_Toc63081354)

[Appendix E Examples 43](#_Toc63081355)

[Appendix F Abbreviations and Acronyms 47](#_Toc63081356)

List of Figures

[Figure 1. Use Case: Browse Variables of a Collection 6](#_Toc63081359)

[Figure 2. GES DISC User interface view for a user to choose a subset of variables for the AIRX2RET collection 7](#_Toc63081360)

[Figure 3. GES DISC User interface view for a user to choose a subset of variables for the OMDOAO3 collection 7](#_Toc63081361)

[Figure 4. GES DISC User interface view for a user to choose a subset of variables for the MOD08 collection 8](#_Toc63081362)

[Figure 5. UMM-Var Metadata Model 9](#_Toc63081363)

[Figure 6. The sea\_surface\_temperature Variable Highlighted within the VIIRS\_SST\_NPP L3C-GHRSST-SST Data Set 10](#_Toc63081364)

[Figure 7. The sea\_surface\_temperature Variable Structure 11](#_Toc63081365)

[Figure 8. A sea\_surface\_temperature Variable Plot 11](#_Toc63081366)

[Figure 9. The quality\_level Variable Structure 11](#_Toc63081367)

[Figure 10. The LST\_Day\_1km Variable Highlighted within the MOD11A1 Data Set 12](#_Toc63081368)

[Figure 11. The LST\_Day\_1KM Variable Structure 12](#_Toc63081369)

[Figure 12. A LST\_Day\_1km Plot 13](#_Toc63081370)

[Figure 13. The QC\_Day Variable Structure 13](#_Toc63081371)

[Figure 14. The CERES\_SW\_Filtered\_Radiances\_Upwards Variable Highlighted within the CER\_BDS\_Aqua-FM3\_Edition1 Data Set Structure 14](#_Toc63081372)

[Figure 15. The CERES\_SW\_Filtered\_Radiances\_Upwards Variable Structure 14](#_Toc63081373)

[Figure 16. CERES\_SYN\_1km Data Set Structure 15](#_Toc63081374)

[Figure 17. The SW\_TOA\_Clear-Sky Variable Highlighted within the CERES\_SYN\_1km Data Set Structure 15](#_Toc63081375)

[Figure 18. The SW\_TOA\_Clear-Sky Variable Structure 15](#_Toc63081376)

[Figure 19. A SW\_TOA\_Clear-Sky Variable Plot 16](#_Toc63081377)

[Figure 20. The mole\_fraction\_of\_carbon\_dioxide\_in\_free\_troposphere Variable Highlighted within the AIRS.2012.02.09.L3.CO2Std008 Data Set 16](#_Toc63081378)

[Figure 21. The mole\_fraction\_of\_carbon\_dioxide\_in\_free\_troposphere Variable Structure 16](#_Toc63081379)

[Figure 22. A mole\_fraction\_of\_carbon\_dioxide\_in\_free\_troposphere Plot 17](#_Toc63081380)

[Figure 23. Subset variable choices for the MOD08 collection 30](#_Toc63081381)

[Figure 24. SMAP\_L3\_SM\_P Variables Representing the Morning (AM) Crossing 32](#_Toc63081382)

[Figure 25. SMAP\_L3\_SM\_P Variables Representing the Afternoon (PM) Crossing 33](#_Toc63081383)

[Figure 26. Suggested Governance Structure 37](#_Toc63081384)

List of Tables

[Table 1. Applicable Documents 3](#_Toc63081385)

[Table 2. Reference Documents 3](#_Toc63081386)

[Table 3. Cardinality 5](#_Toc63081387)

[Table 4. Example Values for the Variable's Unit 19](#_Toc63081388)

[Table 5. HDF4 User Guide as a Possible Source 20](#_Toc63081389)

[Table 6. HDF5 User Guide as a Possible Source 20](#_Toc63081390)

[Table 7. Tag's Glossary 36](#_Toc63081391)

|  |  |  |
| --- | --- | --- |
| V1.7.0 | August 2020 | * Name field now must be unique within a collection and should start with a '/'  and include a '/' delimited hierarchy if the variable is hierarchical. Added StandardName. Added AdditionalIdentifiers. Removed Alias, AcquisitionSourceName, and SizeEstimation from the schema. Removed Characteristics/GroupPath. This is now part of the top level Name field. Removed Characteristics. Removed CoordinatesType. IndexRanges has been pulled out of Characteristics and made a top level field. Dimensions, Sets, DataType, Scale, and Offset are now optional fields and no longer required. |

# Introduction

The NASA Earth Observing System Data and Information System (EOSDIS) generates, archives, and distributes massive amounts and a large variety of Earth Science data via twelve Distributed Active Archive Centers (DAACs). Reliable, consistent and high-quality metadata are essential to enable cataloging and proper use of these data. To improve the quality and consistency among its metadata holdings, EOSDIS has developed models for metadata that it archives and maintains. This model aims to document vital elements that may be represented across various data models and standards and unify them through mainstream fields useful for data discovery, data use, and service invocations. This unified model, aptly named the Unified Metadata Model (UMM) will be used by the CMR and will drive search metadata cataloged within that system and retrieve data discovered through such searches.

This document describes the Unified Metadata Model for Variables (UMM-Var). It includes the uses cases for UMM-Var model itself and its relationship with other UMM models, element descriptions, and examples.

Listed below are some definitions with examples that will help the reader understand this model.

* Measurement: The act or process of measuring an observable property, usually geophysical, geo-biophysical, physical, or chemical. In the case of air temperature, for instance, the object of the measurement is air and the property being measured is temperature. For models, it is a simulated observable property.
  + Using Scott Peckham's model as a basis for a measurement naming convention, the Measurement names can be expressed as: <<object, quantity>>, object = "Aerosol", quantity = "Optical Depth".
  + Examples: Aerosol Optical Depth, Air Temperature, Surface Albedo, Solar Irradiance, Surface Reflectance, Atmospheric Moisture, Methane Concentration, Sulphur Dioxide Concentration, Ozone Concentration.
* Variable: A named set of data that contains the recorded values of a measurement. In this context, the variable is described by its name and characteristics. For instance, a variable contained within the MYD08\_M3V5 dataset is called: Optical\_Depth\_Land\_Maximum. There are other variables in the set, including variables which contain information about geographic position and quality.

Variables may be classified as science variables, quality variables and ancillary variables (or other, when one of these classifications cannot be used). A variable can also be the output of a model.

* Examples: Aerosol Optical Depth 550nm (Dark Target), Aerosol Optical Depth 550nm (Deep Blue, Land Only), Air Temperature (Daytime/Ascending), Air Temperature at 2m, Air Temperature at Surface (Daytime/Ascending), Air Temperature at Surface (Nighttime/Descending), Relative Humidity (Daytime/Ascending), Relative Humidity (Nightime/Descending), Water Vapor Mass Mixing Ratio (Daytime/Ascending), Water Vapor Mass Mixing Ratio (Nightime/Descending), Methane Total Column (Nighttime/Descending), SO2 Column Mass Density, SO2 Column Mass Concentration, Ozone - reported in parts per billion by volume.

The term "Measurement" is the act or process of measuring an observable property, and is mostly likely to be used as a search term, as an alternate to the Science Keywords. The term "Variable" is an artifact that represents a Measurement. The UMM-Var model is not interested in the direct measurement that the instrument made. It is the "feature of interest" and the "observed property" represented by the data that are of interest. The Variable class will be used to store metadata about each variable. The Variable metadata will consist of its name and other characteristics. The CMR Variable class can be utilized to simplify search and retrieval of data products at the variable level.

In terms of the data product and its file structure, variables are stored within a data granule, such as Aerosol Optical Depth 550nm (Dark Target), Aerosol Optical Depth 550nm (Deep Blue, Land Only), along with its associated data quality variables, and ancillary variables, such as latitude and longitude information.

## Purpose

The purpose of UMM-Var is to express a variable model applicable to CMR that stores variable metadata. In addition, the UMM-Var model is related to the other CMR metadata models, such as UMM-S, which supports the specification of variables which have associated services.

Note: the previous variable design principally addressed the concept of parameters. The parameter version of this model, known as UMM-P, sought to bridge the divide between variables and collection-level additional attributes. However, this new model, UMM-Var, considers variables in their own right. Now variables can be stored and discovered in ways described by a new set of use cases. Granule data may be subsetted by variable, or transformed in other ways, as supported by services. The user experience guides what selections and choices a user makes at the UI for typical data transformations such as spatial subsetting, reprojection, reformatting, etc. The user is exclusively concerned about what choices are available for a specific data set and the back-end services take care of any needed processing.

This document provides information to the NASA Earth Science community. Distribution is unlimited.

## Scope

This document describes the Unified Metadata Model - Variables (UMM-Var) model version 1.7.

## Related Documentation

The latest versions of all documents below should be used. The latest ESDIS Project documents can be obtained from Uniform Resource Locator (URL): https://ops1-cm.ems.eosdis.nasa.gov. ESDIS documents have a document number starting with either 423 or 505. Other documents are available for reference in the ESDIS project library website at: http://esdisfmp01.gsfc.nasa.gov/esdis\_lib/default.php unless indicated otherwise.

### Applicable Documents

The following documents are referenced within, are directly applicable, or contain policies or other directive matters that are binding upon the content of this document.Note: Links require NASA Credentials

Table . Applicable Documents

|  |  |
| --- | --- |
| **Document Number** | **Document Title** |
| N/A | CMR Life Cycle  https://wiki.earthdata.nasa.gov/display/CMR/CMR+Documents |
| EED2-TP-025 | CMR End-To-End Services Study (Task 25) EED2-TP-025  https://wiki.earthdata.nasa.gov/download/attachments/83624411/EED2-TP-025\_CMR%20End-To-End%20Services%20Study.pdf?api=v2 |
| N/A | Scale Calibration Attributes  https://support.hdfgroup.org/release4/doc/UG\_PDF.pdf (Section 3.10.6 Calibration Attributes) |
| UG\_3.6.3 | Scale Attribute Conventions  https://cdn.earthdata.nasa.gov/conduit/upload/495/  netcdf\_UG\_3.6.3.pdf (See Appendix B Attribute Conventions) |

### Reference Documents

The following documents are not binding on the content but referenced herein and amplify or clarify the information presented in this document. Note: Links require NASA Credentials

Table . Reference Documents

|  |  |
| --- | --- |
| **Document Number** | **Document Title** |
| N/A | Tags  http://en.wikipedia.org/wiki/Tag\_%28metadata%29 |
| N/A | XPath  XPath is a language for addressing parts of an XML document, designed for use with XSLT. |

## Impact

This document outlines a model intended to be compatible with existing NASA Earth Science metadata implementations within the CMR. It will impact providers from NASA Distributed Active Archive Centers (DAACs), non-DAAC data providers, instrument Principal Investigators (PI), CMR client developers, metadata catalog developers, and users. Users will be impacted specifically in terms of data discovery and data use. This is very important for science research purposes.

## Copyright Notice

The contents of this document are not protected by copyright in the United States and may be used without obtaining permission from NASA.

## Feedback

Questions, comments and recommendations on the contents of this document should be directed to support@earthdata.nasa.gov

## Document Conventions

There are two main sections to the rest document: the use cases and the detailed description of the metadata model. The use case section describes the use cases used to create the metadata model. Each use case section contains the following information:

* Scenarios: One or more related scenarios are described in this section.
* Outcomes: A description of what the system provides the user as a result of the scenarios.
* Use Case Diagram: A diagram that highlights the actor's interaction with the system.
* Activity Diagram: A diagram that shows the flow of data in terms of the user experience.
* Sequence Diagram: A diagram which shows the key components of the system and the sequences of actions within the system.

The detailed description of the metadata model section of this document describes each element within the model. Variable model elements are documented in the following way:

* Element Name: Specifies the element name.
* Element Specification: Provides the sub-elements, cardinality of the sub-elements within (), any valid values within <>, applicable comments and notes within {}, and any other major factors that make up the element.
* Description: Provides background information on the purpose of the element and how it should be used. Any notes about the current usage of this element are documented here as well as any recommendations for usage or unresolved issues.
* Tags: Provides specific, related categorical values associated with this element, which are defined in Appendix A: Tags Glossary.

With the exception of Element Name each of the element's sections are that are included are listed in bold to make it easier for the reader to distinguish between the element's section headings and the descriptions.

Table . Cardinality

| **Value** | **Description** |
| --- | --- |
| 1 | Exactly one of this element is required |
| 0..N | This element is optional; up to and including N number of this element may be present |
| 0..\* | Optionally, many of this element may be present |
| 1..\* | At least one of this element is required, many may be present |

Interaction diagrams presented in this document are based on the Unified Modeling Language (UML) notation.

# Unified Metadata Model - Variables

## Use Cases

This section provides information about use cases identified for the UMM-Var.

### Browse Variables of a Collection

Scenario: A user starts with a collection and wants to know what variables it includes.

**Outcomes:** Enables a user without any knowledge of variable names to search for collections, select one, and be presented with a list of variables for that collection under the customized data access method.



Figure . Use Case: Browse Variables of a Collection

### Cross-site Data Subsetting

Scenario: As a subsetting GUI, I can present the variables for a given collection and further subset the data into more specific groups based on additional criteria.

**Outcomes:** Enables users of a subsetting GUI to perform cross-site subsetting variables based on the selection of a collection. Cross-site subsetting occurs when a variable (by its association with a granule) can exist in more than one collection and these collections may be sourced from multiple sites (i.e. Goddard Earth Sciences Data and Information Services Center (GES DISC), Level-1 and Atmosphere Archive and Distribution System (LAADS), etc.). The CMR can perform a cross-site search since it houses metadata from all sites. This use case enables a user to go on to perform subsetting via a GUI.

Note: In the example shown below using GES DISC, the measurement term used was "Ozone". This resulted in three collections being returned from the search: AIRX2RET v005, OMDOAO3 v003, and MOD08 v006. In the subsetting GUI, variables are shown grouped for each collection. The user will be able to subset the variable fields for specific granules of interest.

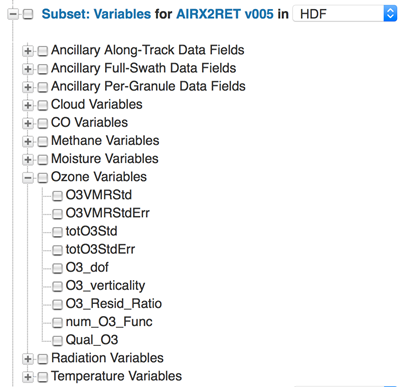


Figure . GES DISC User interface view for a user to choose a subset of variables for the AIRX2RET collection

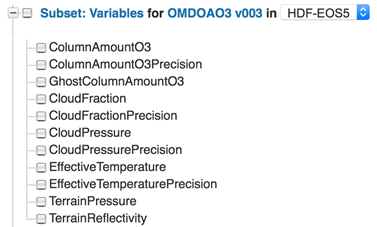


Figure . GES DISC User interface view for a user to choose a subset of variables for the OMDOAO3 collection



Figure . GES DISC User interface view for a user to choose a subset of variables for the MOD08 collection

## UMM-Var Metadata Model

As shown in Figure 5, the UMM-Var Metadata Model asserts that a Variable metadata instance is related to one collection. The remaining classes are discussed in more detail throughout the remainder of this document. Each class and relationship express a different type of information conveyed by the variable.

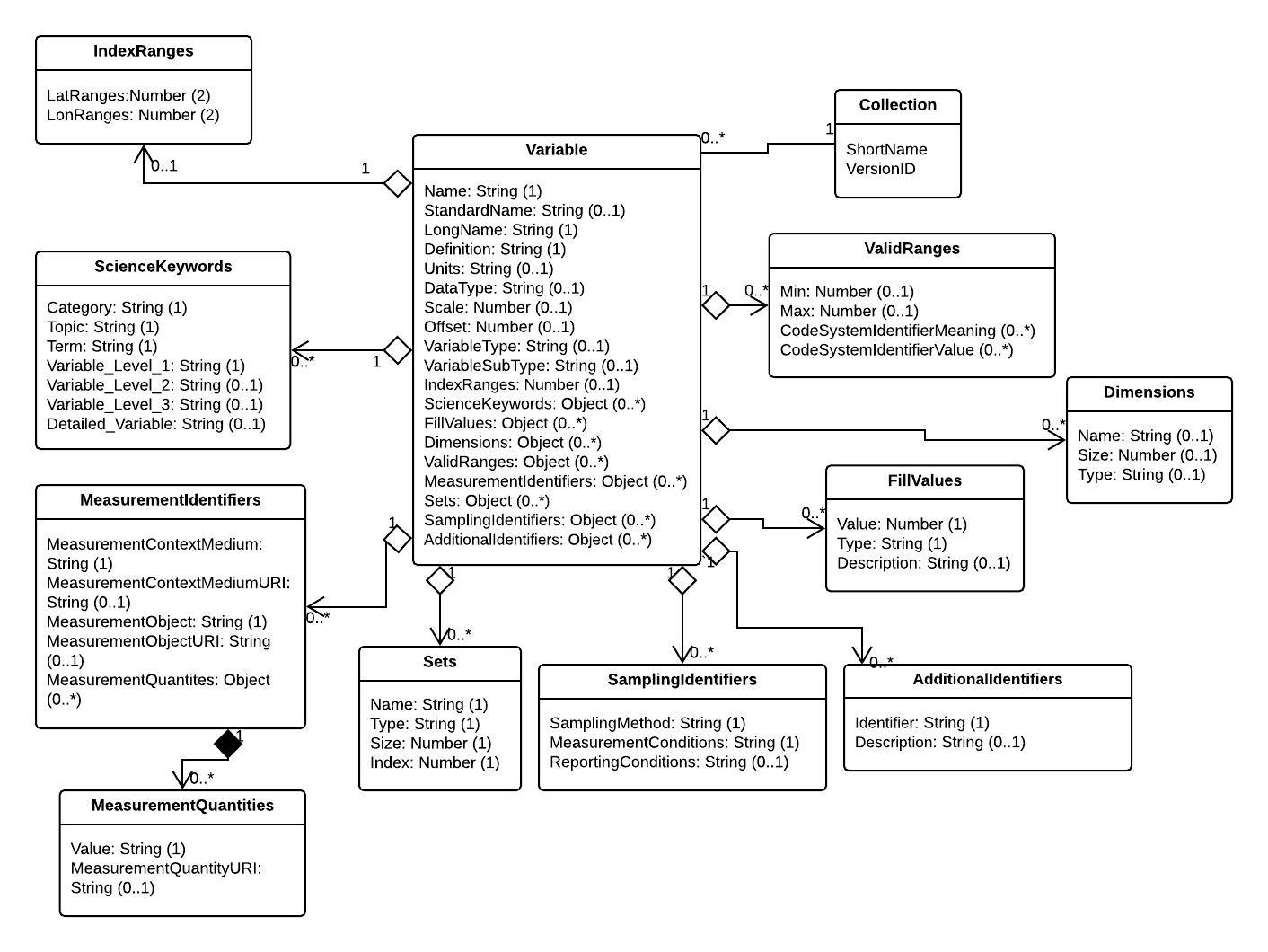


Figure . UMM-Var Metadata Model

The author of a Variable metadata record should be aware of the following:

1. A Collection can only be associated to one Variable.
2. A Variable must be created with an association to a Collection.
3. Names must be unique within a Collection.
4. Name must include GroupPath.
5. The elements of the Science Keywords section also apply to a Variable. The Science Keywords may be sourced from Global Change Master Directory (GCMD) Keywords. (See Appendix B).
6. The elements of the Measurement section apply to a Variable. The Measurement names may be sourced from the CSDMS standard names or the CF Convention standard names. This process will be dictated by a GCMD-style Governance process. (See Appendix C).
7. A Variable record may be created / updated via the MMT GUI or via a json file.
8. A Variable's record should answer all parts of the following question: What measurement type, collections, variables, granules, visualizations are associated with the Variable?

### Name [R]

**Element Specification:**

Name (1)

**Description:**

A variable short name given by the data provider. The name should include a '/' delimited hierarchy and must be unique within a collection.

Variables are available in a wide range of forms. These variables are named similarly across a family of collections, but these names differ considerably across collections. The variety of variables is illustrated using some examples across a sample of collections in Figure 6 below.

The VIIRS\_SST\_NPP L3C-GHRSST-SST Data Set structure is represented in Figure 6.

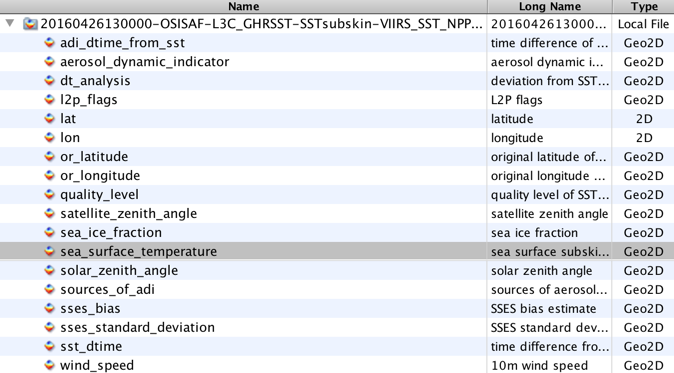


Figure . The sea\_surface\_temperature Variable Highlighted within the VIIRS\_SST\_NPP L3C-GHRSST-SST Data Set

The highlighted sea\_surface\_temperature variable structure is shown in Figure 7 with a plot shown in Figure 7. Note the dimensionality of the variable is: time=1, nj=3072 and ni=4096.

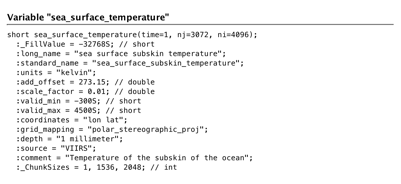


Figure . The sea\_surface\_temperature Variable Structure

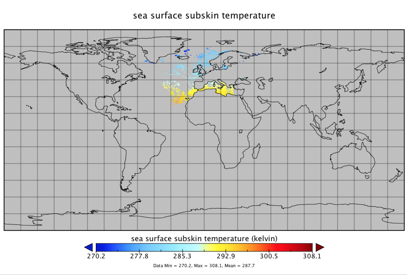


Figure . A sea\_surface\_temperature Variable Plot

The corresponding data quality variable is shown in Figure 9. Note the dimensionality of the variable is: time=1, nj=3072 and ni=4096.

​



Figure . The quality\_level Variable Structure

​

Conversely, the LST variable contained within the MOD11A1 Data Set Structure is shown in Figure 10.

​

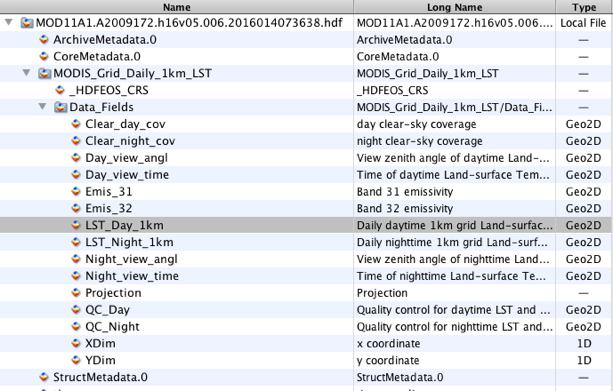


Figure . The LST\_Day\_1km Variable Highlighted within the MOD11A1 Data Set

​

t

The LST\_Day\_1km variable structure is represented as shown in Figure 11 with a plot in Figure 12. Note the dimensionality of the variable is: YDim=1200 and XDim=1200.

​

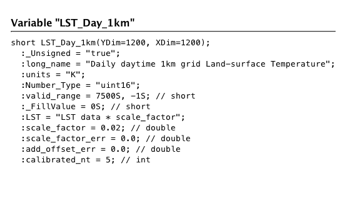


Figure . The LST\_Day\_1KM Variable Structure

​

​

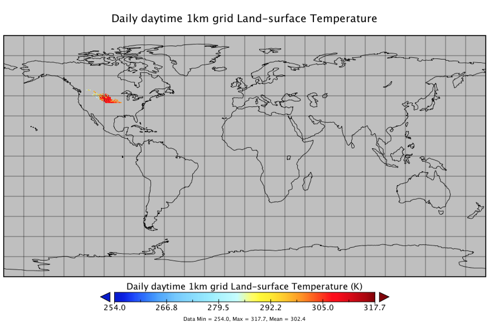


Figure . A LST\_Day\_1km Plot

The corresponding quality variable is represented as shown in Figure 13. Note the dimensionality of the variable is: YDim=1200 and XDim=1200.

​

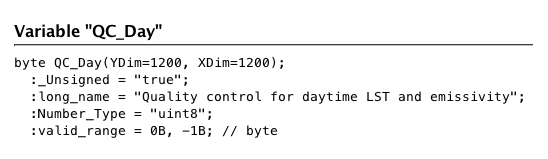


Figure . The QC\_Day Variable Structure

​

CER\_BDS\_Aqua-FM3\_Edition1 Data Set structure is represented in Figure 14.

​

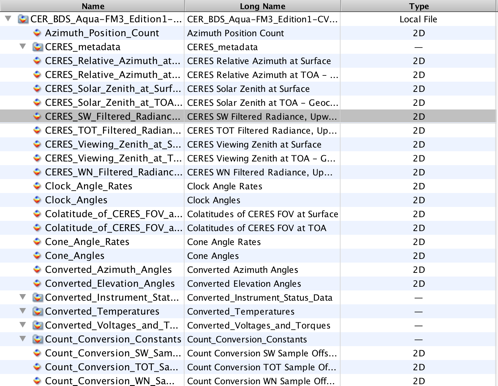


Figure . The CERES\_SW\_Filtered\_Radiances\_Upwards Variable Highlighted within the CER\_BDS\_Aqua-FM3\_Edition1 Data Set Structure

​

The selected CERES\_SW\_Filtered\_Radiances\_Upwards variable structure is represented in Figure 15. Note the dimensionality of the variable is: Records=13091 and Samples=660.

​

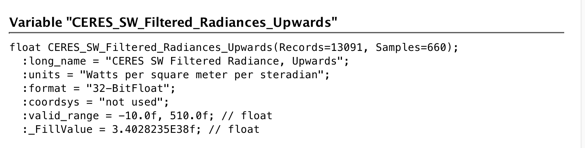


Figure . The CERES\_SW\_Filtered\_Radiances\_Upwards Variable Structure

​.

CERES\_Solar\_Zenith\_at\_Surface variable structure is represented in Figure 16.

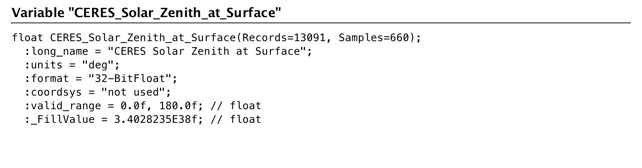


Figure . CERES\_SYN\_1km Data Set Structure

The SW\_TOA\_Clear-Sky variable is highlighted within the CERES\_SYN\_1km Data Set structure as shown in Figure 17.

​

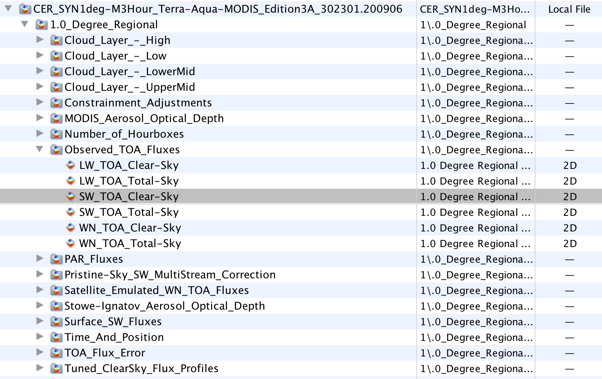


Figure . The SW\_TOA\_Clear-Sky Variable Highlighted within the CERES\_SYN\_1km Data Set Structure

Figure 18 is the SW\_TOA\_Clear-Sky variable structure is representation and in Figure 19 is a plot. Note the dimensionality of the variable is: Mean\_&\_Stdev=2, Synoptic\_Hours\_(1, 4, 7, 10, 13, 16, 19, 22)=8, 1.0\_deg.regional\_colat.zones=180 and 1.0\_deg.\_regional\_long.\_zones=360.

​

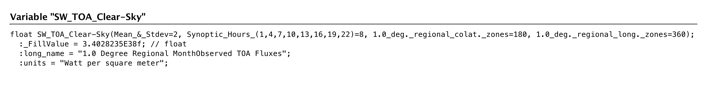


Figure . The SW\_TOA\_Clear-Sky Variable Structure

​

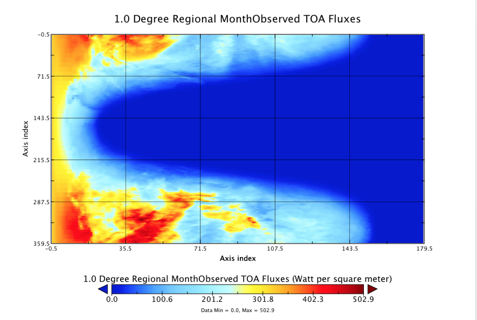


Figure . A SW\_TOA\_Clear-Sky Variable Plot

​

AIRS.2012.02.09.L3.CO2Std008 data set structure is represented in Figure 20.

​

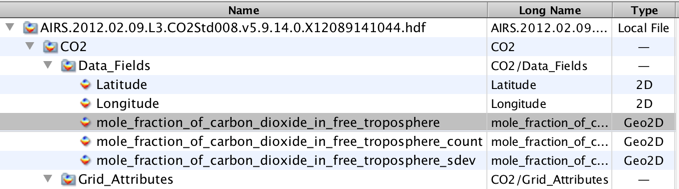


Figure . The mole\_fraction\_of\_carbon\_dioxide\_in\_free\_troposphere Variable Highlighted within the AIRS.2012.02.09.L3.CO2Std008 Data Set

The highlighted mole\_fraction\_of\_carbon\_dioxide\_in\_free\_troposphere variable structure is shown in Figure 21 and its plot in Figure 22. Note the dimensionality of the variable is: LatDim=91, LonDim=144.

​

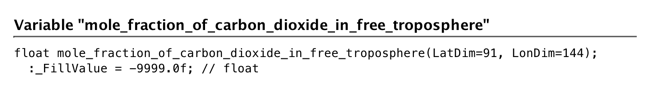


Figure . The mole\_fraction\_of\_carbon\_dioxide\_in\_free\_troposphere Variable Structure

​

​

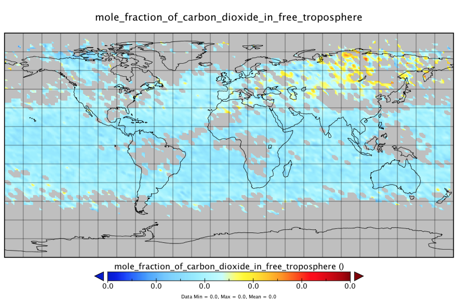


Figure . A mole\_fraction\_of\_carbon\_dioxide\_in\_free\_troposphere Plot

​

**Sample Values:**

/science/grids/data/amplitude

/gt1l/land\_segments/canopy/h\_canopy

sea\_surface\_temperature

**Tags:**

*Required, Free Text Search*

### StandardName

**Element Specification:**

StandardName (0..1)

**Description:**

Refers to its more formal/scientific name, e.g. the CF Standard Name.

**Sample Values:**

sea\_surface\_foundation\_temperature

air\_density

**Tags:**

*Optional, Free Text Search*

### LongName [R]

**Element Specification:**

LongName (1)

**Description:**

The expanded or long name given by the data provider.

**Sample Values:**

sea\_surface\_temperature (sea surface temperature)

quality\_level (quality level of the sea surface temperature)

LST\_1km\_Day (daily daytime 1km grid land surface temperature)

QC\_day (quality control for daytime LST and emissivity)

CERES\_SW\_Filtered\_Radiances\_Upwards (CERES SW filtered radiances, upwards)

CERES\_Solar\_Zenith (CERES solar zenith at surface)

SW\_TOA\_Clear-Sky (1 degree regional month observed TOA fluxes)

>mole\_fraction\_of\_carbon\_dioxide\_in\_free\_troposphere (mole fraction of carbon dioxide in free troposphere)

psl (mean sea level pressure)

O3\_ppbv (ozone mixing ratio reported in parts per billion by volume)

Scat\_550 (total dry aerosol scattering coefficient at 550 nm)

LST\_1KM\_Day (daily daytime 1km grid land surface temperature)

Sur\_Refl\_b01 (surface reflectance band 1)

WDB\_L3MCA10 v004 (Aerosol Optical Depth 550nm (Land Only)

**Tags:**

*Required, Free Text Search*

### Definition [R]

**Element Specification:**

Definition (1)

**Description:**

The meaning of the variable given by the data provider. This can typically be found in the Collection User Guide corresponding to the variable. Ideally, it should include the details of what is being measured, the scope of the measurement, and any other information to assist a scientist's understanding of the variable's particularities. See the SamplingIdentifiers class for details about the sampling method and the measurement and reporting conditions.

**Sample Value:**

Angstrom Exponent is an exponent that expresses the spectral dependence of aerosol optical thickness (τ) with the wavelength of incident light (λ). The spectral dependence of aerosol optical thickness can be approximated (depending on size distribution) by, τa = β λα where α is the Angstrom exponent (β = aerosol optical thickness at 1 μm).

**Tags:**

*Required*

### Units

**Element Specification:**

Units (0..1)

**Description:**

The unit used to report the variable. The list of units will be sourced from the Dataset Interoperability Working Group (https://wiki.earthdata.nasa.gov/display/ESDSWG/Dataset+Interoperability+Working+Group). The list will be managed as a KMS-managed list.

**Sample Values:**

Table . Example Values for the Variable's Unit

| **Coordinate Variable** | **Unit Value** | **Examples** |
| --- | --- | --- |
| latitude | degrees\_north | 89.9 degrees\_north |
| longitude | degrees\_east | -179.9 degrees\_east |
| pressure | Pa or hPa | 50 Pa |
| height (depth) | meter (m) or kilometer (km) | 10,000 m |
| time | Seconds, minutes, hours, days, etc., since a specific starting point in time, often (but not always) representing a canonical time (1 Jan 1970, TAI93, start of mission, etc.). | Time is in International Organization for Standardization (ISO)-8601 format. seconds since 1992-10-08T15:15:42.5-6:00 days since 1970-01-01T00:00:0 |

**Tags:**

*Optional, Controlled Vocabulary*

### DataType

**Element Specification:**

DataType  (0..1) <"byte", "float", "float32", "float64", "double", "ubyte", "ushort", "uint", "uchar", "string", "char8", "uchar8", "short", "long", "int", "int8", "int16", "int32", "int64", "uint8", "uint16", "uint32", "uint64", "OTHER">

**Description:**

Specifies the basic computer science data type of a variable. These types can be either short, long, character, binary, etc. Table 3 and Table 4 list out some data types from the Hierarchical Data Format (HDF) version 4 and 5 specifications.

Table . HDF4 User Guide as a Possible Source

| **HDF Data Type** | **Data Type Flag and Value** | **Description** |
| --- | --- | --- |
| char8 | DFNT\_CHAR8 (4) | 8-bit character type |
| uchar8 | DFNT\_UCHAR8 (3) | 8-bit unsigned character type |
| int8 | DFNT\_INT8 (20) | 8-bit integer type |
| uint8 | DFNT\_UINT8 (21) | 8-bit unsigned integer type |
| int16 | DFNT\_INT16 (22) | 16-bit integer type |
| uint16 | DFNT\_UINT16 (23) | 16-bit unsigned integer type |
| int32 | DFNT\_INT32 (24) | 32-bit integer type |
| uint32 | DFNT\_UINT32 (25) | 32-bit unsigned integer type |
| float32 | DFNT\_FLOAT32 (5) | 32-bit floating-point type |
| float64 | DFNT\_FLOAT64 (6) | 64-bit floating-point type |

Table . HDF5 User Guide as a Possible Source

| **HDF5 Data Type** | **Data Type Flag and Value** | **Description** |
| --- | --- | --- |
| string | NC\_STRING | string type |
| char | NC\_CHAR | character type |
| ubyte | NC\_UBYTE | unsigned byte type |
| ushort | NC\_USHORT | unsigned short type |
| uint | NC\_UINT | unsigned integer type |
| uint64 | NC\_UINT64 | 64-bit unsigned integer type |
| byte | NC\_BYTE | byte type |
| short | NC\_SHORT | short type |
| int | NC\_INT | integer type |
| int64 | NC\_INT64 | 64-bit integer type |
| double | NC\_DOUBLE | double type |

**Sample Value:**

 float

**Tags:**

*Recommended, Controlled Vocabulary*

### Scale and Offset

**Element Specification:**

Scale (0..1)

Offset (0..1)

**Description:**

The Scale is the numerical factor by which all values in the stored data field are multiplied in order to obtain the original values. The Offset is the value which is either added to or subtracted from all values in the stored data field in order to obtain the original values. Scale and Offset may be used together. The formula by which the Scale and Offset are applied is usually one of the following:

1. Additive Offset formula: actual data value = (scale factor \* scaled value) + offset
2. Subtractive Offset formula: actual data value = scale factor \* (scaled value - offset

Note: the additive offset formula is the standard one, with the subtractive being non-standard, and rarely used. Exceptions include science variables from: Moderate Resolution Imaging Spectroradiometer (MODIS), MOD08\_M3 (MODIS/Terra Aerosol Cloud Water Vapor Ozone Monthly L3 Global 1Deg CMG)and MCD43A4 (MODIS/Terra+Aqua BRDF/Albedo Nadir BRDF-Adjusted Ref Daily L3 Global - 500m) which use the subtractive offset formula.

**Sample Values:**

Scale: 0.00100000004749745

Offset: 0.0

**Tags:**

*Recommended*

### VariableType

**Element Specification:**

VariableType (0..1) <"SCIENCE\_VARIABLE", "QUALITY\_VARIABLE", "ANCILLARY\_VARIABLE", "OTHER">

**Description:**

This element is controlled and specifies the basic type of a variable. These types can be either: "SCIENCE\_VARIABLE", "QUALITY\_VARIABLE", "ANCILLARY\_VARIABLE", or "OTHER".

**Sample Value:**

SCIENCE\_VARIABLE

**Tags:**

*Recommended, Controlled Vocabulary*

### VariableSubType

**Element Specification:**

VariableSubType (0..1) <"SCIENCE\_SCALAR", "SCIENCE\_VECTOR", "SCIENCE\_ARRAY", "SCIENCE\_EVENTFLAG", "OTHER">

**Description:**

This element is controlled and specifies the sub type of a variable. There are different types of science variables and this information is variable specific and important for data use. The sub-types can be used in the following way: science\_scalar (e.g., O3, NO, NO2, CH2O, CN, etc.); science\_vector (e.g., wind direction); science\_array (e.g., radiation spectrum, aerosol number size distribution); science\_eventflag (e.g., cloud flag, pollution plume). There are other types of variables not included here.

**Sample Value:**

SCIENCE\_SCALAR

**Tags:**

*Recommended, Controlled Vocabulary*

### IndexRanges

**Element Specification:**

IndexRanges (0..1)

IndexRanges/LatRange (2)

IndexRanges/LonRange (2)

**Description:**

Describes the spatial index ranges of a variable, which consist of a LatRange as a pair of values and a LonRange as a pair of values. If the IndexRanges element is used the LatRange and LonRange sub elements are required. In the example shown below, the index ranges represent the ranges of the latitude and longitude respectively of the variable. Each range is described by a pair of values.

**Sample Values:**

IndexRanges/LatRange: [89.5, -89.5]

IndexRanges/LonRange: [-179.5, 179.5]

**Tags:**

*Recommended*

### ScienceKeywords

**Elements:**

ScienceKeywords (0..\*)

ScienceKeywords/Category [R]

ScienceKeywords/Topic [R]

ScienceKeywords//Term [R]

ScienceKeywords/Variable\_Level1

ScienceKeywords/Variable\_Level2

ScienceKeywords/Variable\_Level3

ScienceKeywords/Detailed\_Variable

**Description:**

Science Keywords are derived from the ESDIS keyword management system. The keywords are provided to enable better searches by the use of human-readable measurement terms. Note that the keywords have a more complex structure than the Measurement class. ScienceKeywords are hierarchical with the higher level keywords. Category, Topic, Term are required and the lower level keywords, VariableLevel1, VariableLevel2 and VariableLevel3 and DetailedVariable are optional. It is important to recognize that the measurement terms are sometimes used in any one of the lower level keywords. So for example the measurement term "Methane" may be entered into the "DetailedVariable" element for a collection which possesses Methane variables, such as AIRX3STD.006.

Science Keywords search is offered as the primary way to discover variables. ScienceKeywords and Measurements could be used interchangeably for faceted browse in search clients. Elements in this category are used for search and faceting purposes.

**Sample Values:**

Category: EARTH SCIENCE, Topic: ATMOSPHERE, Term: ATMOSPHERIC CHEMISTRY, Variable\_Level1: NITROGEN COMPOUNDS, Variable\_Level2: Peroxyacyl Nitrate

**Tags:**

*Recommended, Controlled Vocabulary*

### FillValues

**Element Specification:**

FillValues (0..\*)

FillValues/Value (1)

FillValues/Type (1) <"SCIENCE\_FILLVALUE", "QUALITY\_FILLVALUE", "ANCILLARY\_FILLVALUE", "OTHER">

FillValues/Description (0..1)

**Description:**

The fill value of the variable in the data file. It is generally a value which falls outside the valid range. For example, if the valid range is '0, 360', the fill value may be '-1'. The fill value type is data provider-defined. It is typically a value out of valid range, although some cases have been reported of exceptions to this rule.

**Sample Values:**

Value: -1

Type: SCIENCE\_FILLVALUE

Description: Valid Science Fill Value

Value: -9999

Type: QUALITY\_FILLVALUE

Description: Valid Quality Fill Value

**Tags:**

*Recommended, Controlled Vocabulary*

### Dimensions

**Element Specification:**

Dimensions (0..\*)

Dimensions/Name (0..1)

Dimensions/Size (0..1)

Dimensions/Type  (0..1) <"LATITUDE\_DIMENSION", "LONGITUDE\_DIMENSION", "PRESSURE\_DIMENSION", "HEIGHT\_DIMENSION", "DEPTH\_DIMENSION", "TIME\_DIMENSION", "OTHER">

**Description:**

A variable consists of one or more dimensions. An example of a dimension name is 'XDim'. An example of a dimension size is '1200'. For the example where time=1; Name = time, Size = 1, and Type = TIME\_DIMENSION. Variables are rarely one dimensional. More commonly, they are two or three dimensional.

**Sample Values:**

Name: time, Size: 1, Type:  TIME\_DIMENSION

Name: nj, Size: 3072, Type:  ALONG\_TRACK\_DIMENSION

Name: ni, Size: 4096, Type:  CROSS\_TRACK\_DIMENSION

**Tags:**

*Recommended, Controlled Vocabulary*

### ValidRanges

**Element Specification:**

ValidRanges (0..1)

ValidRanges/Max (0..1)

ValidRanges/Min (0..1)

ValidRanges/CodeSystemIdentifierMeaning (0..\*)

ValidRanges/CodeSystemIdentifierValue (0..\*)

**Description:**

ValidRanges specifies the minimum and maximum valid values of the variable represented in the data field. Optionally, if the valid range is not continuous, a code system can be defined. If there is a discrete number system used for the data values, then there needs to be a code system identifier. The CodeSystemIdentifierMeaning element can be used to specify a code system identifier meaning. For example, 'Open Shrubland' corresponds to the value of '7'. The CodeSystemIdentifierValue element describes the textual or numerical value assigned to each meaning. The number of code system identifier meanings must match the number of values. Other examples include cloud masks, land surface classification variables, etc.

**Sample Values:**

ValidRanges/Max: 5000

ValidRanges/Min: -100

ValidRangeCodeSystemIdentifierMeaning: no\_data, bad\_data, worst\_quality, low\_quality, acceptable\_quality, best\_quality

ValidRange/CodeSystemIdentifierValue: 0B, 1B, 2B, 3B, 4B, 5B

**Tags:**

*Recommended*

### MeasurementIdentifiers

**Elements:**

MeasurementIdentifiers (0..\*)

MeasurementIdentifiers/MeasurementContextMedium (1)

MeasurementIdentifiers/MeasurementContextMediumURI (0..1)

MeasurementIdentifiers/MeasurementObject (1)

MeasurementIdentifiers/MeasurementObjectURI (0..1)

MeasurementIdentifiers/MeasurementQuantities (0..1)

MeasurementIdentifiers/MeasurementQuantities/Value (0..1)

MeasurementIdentifiers/MeasurementQuantitiesMeasurementQuantityURI (0..1)

**Description:**

Elements in this category are used for search purposes. The measurement identifier is structured according to the form defined by : <<contextmedium, object, quantity>>, and it is for this reason that the measurement identifiers class contains the ContextMedium, Object and Quantity elements. Every measurement identifier has a contextmedium part that describes where the measurement was taken, an object part that describes the object of the measurement, and a quantity part that describes a particular attribute of the object that can be quantified. There can be multiple quantity terms for a given object. These terms are sorted alphabetically and other sorting methods can be added later.

In addition, the definition of the meaning of each of the terms has a URI, which points to a location of the specific term where it resides on the semantic web.

In consultation with the GCMD team, it is recommended that MeasurementIdentifier's valid values should be enumerations in KMS and not keywords. It is also recommended that the MeasurementIdentifier's valid values are managed via the current ESDIS Standards Office (ESO) process but not until the valid values have matured.

In order to complete the selection of values for the ContextMedium, Object, Quantity (or Quantities) and Units, we need to follow a repeatable set of rules for creating the structured measurements identifiers.  
   
The proposed syntactic rules are: (1) in the first position, put the medium/context within which the measurement occurred; (2) in the second position, put the object that was measured; (3) in the third position, put the quantity (4) use \_ to separate the words in multi-word terms (5) use \_\_ to separate the three terms  
   
A term refers to either a context, an object or a quantity term.  
Terms may be made up of more than one word.  
   
The reason for having a set of syntactic rules is to clarify what order to place these terms and to ensure that different people, with different backgrounds in science or data engineering will arrive at the same set or measurement names, given the same set of choices. It also provides the basis for machine automation of the parsing of community-sourced lists when there are large numbers of terms to handle at once.

**Sample Values:**

MeasurementContextMedium: atmosphere, MeasurementContextMediumURI: http://www.ontobee.org/ontology/ENVO?iri=http://purl.obolibrary.org/obo/ENVO\_01000267, MeasurementObject: cloud, MeasurementObjectURI: http://www.ontobee.org/ontology/ENVO?iri=http://purl.obolibrary.org/obo/ENVO\_01000760, MeasurementQuantities/Value: albedo, MeasurementQuantitiesMeasurementQuantityURI: http://www.ontobee.org/ontology/ENVO?iri=http://purl.obolibrary.org/obo/PO\_0009087.

**Tags:**

*Recommended, Controlled Vocabulary*

#### MeasurementContextMedium [R]

**Description:**

This element describes the context/medium within which the measurement was made.

MeasurementContextMedium: The related GCMD science keyword (usually a topic or term) of the Measurement. e.g. atmosphere. Other terms, e.g. "at\_cloud\_top" can be added as needed. Any additional terms will be separated using a "-" character.

**Sample Value:**

MeasurementIdentifiers/MeasurementContextMedium: atmosphere

MeasurementIdentifiers/MeasurementContextMedium: atmosphere-at\_cloud\_top

MeasurementIdentifiers/MeasurementContextMedium: land

**Tags:**

*Recommended, Controlled Vocabulary*

#### MeasurementContextMediumURI

**Description:**

This element contains the URI for the context/medium.

**Sample Value:**

http://www.ontobee.org/ontology/ENVO?iri=http://purl.obolibrary.org/obo/ENVO\_00002042 (URI reference for surface water)

**Tags:**

*Recommended, Controlled Vocabulary*

#### MeasurementObject [R]

**Description:**

This element describes the object which was measured.

MeasurementObject: The name of the object of measurement. The object part describes a particular object or phenomenon which is being measured. For example, the object of a measurement of aerosol optical depth would be aerosol, and the object of a measurement of longwave radiative flux would be longwave radiation

**Sample Value:**

MeasurementIdentifiers/MeasurementObject: air

MeasurementIdentifiers/MeasurementObject: snow

**Tags:**

*Recommended, Controlled Vocabulary*

#### MeasurementObjectURI

**Description:**

This element contains the URI for the object which was measured.

**Sample Value:**

http://www.ontobee.org/ontology/ENVO?iri=http://purl.obolibrary.org/obo/ENVO\_01000704 (URI referemce for hydrological\_evaporation)

**Tags:**

*Recommended, Controlled Vocabulary*

#### MeasurementQuantities

**Element Specification:**

MeasurementIdentifiers/MeasurementQuantities (0..\*)

MeasurementIdentifiers/MeasurementQuantities/Value (0..1)

MeasurementIdentifiers/MeasurementQuantities/MeasurementQuantityURI (0..1)

**Description:**

Value: The name of the quantity of the measurement. The quantity part describes a particular attribute of that object that can be quantified with a number. For example, the quantity of a measurement of aerosol optical depth would be optical depth, and the quantity of a measurement of longwave radiative flux would be radiative flux.

**Sample Values:**

MeasurementIdentifiers/MeasurementQuantities/Value: albedo

MeasurementIdentifiers/MeasurementQuantities/MeasurementQuantityURI: http://www.ontobee.org/ontology/ENVO?iri=http://purl.obolibrary.org/obo/PO\_0009087

MeasurementIdentifiers/MeasurementQuantities/Value: temperature

MeasurementIdentifiers/MeasurementQuantities/MeasurementQuantityURI: http://www.ontobee.org/ontology/ENVO?iri=http://purl.obolibrary.org/obo/PATO\_0000146

**Tags:**

*Recommended, Controlled Vocabulary*

##### Value

**Element Specification:**

MeasurementIdentifiers/MeasurementQuantities/Value (0..1)

**Description:**

The value of the quantity of measurement. The quantity part describes a quantifiable value for the object which is being measured.

**Sample Values:**

mass\_flux

geopotential\_height

**Tags:**

*Required*

##### MeasurementQuantityURI

**Element Specification:**

MeasurementIdentifiers/MeasurementQuantities/MeasurementQuantityURI (0..1)

**Description:**

The URI of the quantity of the measurement. The quantity part describes a particular attribute of that object that can be quantified with a number. The URI points to a location on the semantic web which defines the quantity.

**Sample Values:**

http://www.ontobee.org/ontology/ENVO?iri=http://purl.obolibrary.org/obo/PATO\_0001687 (URI reference for elevation)

**Tags:**

*Recommended*

### Sets

**Elements:**

Sets (0..\*)

Sets/Name (1)

Sets/Type (1)

Sets/Size (1)

Sets/Index (1)

**Description:**

Typically, science variables have quality variables associated with them and can also include other types. This element allows for variables to be grouped together as a set. The set is defined by the name, type, size, and index. The Set class is flexible enough to also include compound variables (a variable that groups related variables together to describe a phenomenon). The data provider will provide the set name, the set type - which is usually the theme of the group or just use the default string of General, the set size - which is the total number of variables in the set, and the index - which is just the numbering scheme for each variable in the set.

**Sample Values:**

This example shows what a variable set would look like for variables common to the 'Data\_Fields' group, within the MOD11A1 collection. The set class would be populated in the following way for the variable named 'LST\_Day\_1km'.

Name: Data\_Fields

Type: MODIS 1km gridded

Size: 15

Index: 7

This example shows what a variable set would look like for variables common to the AIRX3STD gridded data field group', within the AIRX3STD collection. The set class would be populated in the following way for the variable named 'EmisIR\_A\_ct'.

Name: AIRX3STD

Type: AIRS+AMSU Level 3 Gridded

Size: 867

Index: 13

Each variable in the set is numbered by Index, and the size of the set. So this is the 14th variable in a set of 867 variables if the numbering starts at 0.

For a phenomenon example, take the MOD08 v006 collection shown in Figure 23.

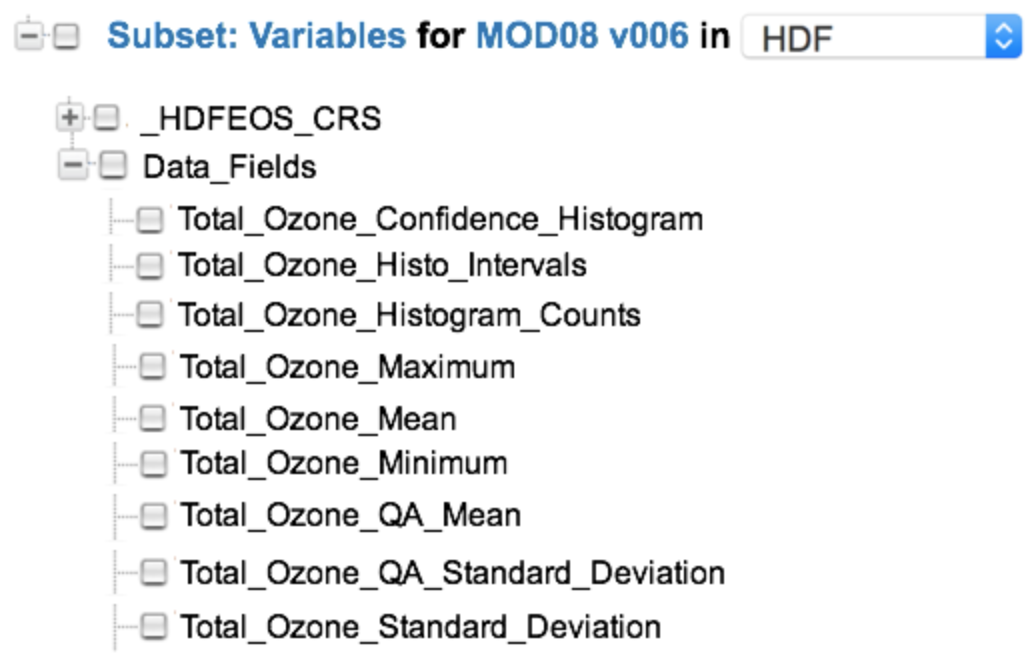


Figure . Subset variable choices for the MOD08 collection

The variables can be grouped which pertain to 'Total Ozone' in the following way:

Variable: {"Name": "Total\_Ozone\_Confidence\_Histogram",

...

"Sets": [{"Name": "Total Ozone","Type": "Data\_Field","Size": 9 ,"Index": 0}],

...

}

Variable: ("Name": "Total\_Ozone\_Histo\_Intervals",

...

"Sets": [{"Name": "Total Ozone","Type": "Data\_Field","Size": 9 ,"Index": 1}],

...

}

Variable: ("Name": "Total\_Ozone\_Histogram\_Counts",

...

"Sets": [{"Name": "Total Ozone","Type": "Data\_Field","Size": 9 ,"Index": 2}],

...

}

Variable: ("Name": "Total\_Ozone\_Maximum",

...

"Sets": [{"Name": "Total Ozone","Type": "Data\_Field","Size": 9 ,"Index": 3}],

...

}

Variable: ("Name": "Total\_Ozone\_Mean",

...

"Sets": [{"Name": "Total Ozone","Type": "Data\_Field","Size": 9 ,"Index": 4}],

...

}

Variable: ("Name": "Total\_Ozone\_Minimum",

...

"Sets": [{"Name": "Total Ozone","Type": "Data\_Field","Size": 9 ,"Index": 5}],

...

}

Variable: ("Name": "Total\_Ozone\_QA\_Mean",

...

"Sets": [{"Name": "Total Ozone","Type": "Data\_Field","Size": 9 ,"Index": 6}],

...

}

Variable: ("Name": "Total\_Ozone\_QA\_Standard\_Deviation",

...

"Sets": [{"Name": "Total Ozone","Type": "Data\_Field","Size": 9 ,"Index": 7}],

...

}

Variable: ("Name": "Total\_Ozone\_Standard\_Deviation",

...

"Sets": [{"Name": "Total Ozone","Type": "Data\_Field","Size": 9 ,"Index": 8}],

...

}

In general, variables are organized in a specific way within the structure of a data set. The examples shown above are for HDF4 structures. The arrangement of these structures varies considerably between HDF4, HDF5 and NetCDF-4, and NetCDF-CF.

In the following HDF5 examples, the variables for this SMAP\_L3\_SM\_P data set are organized into two sets. The first set, shown in Figure 24, contains the variables representing the morning (AM) crossing and the second set, shown in Figure 25, contains variables representing the afternoon (PM) crossing.

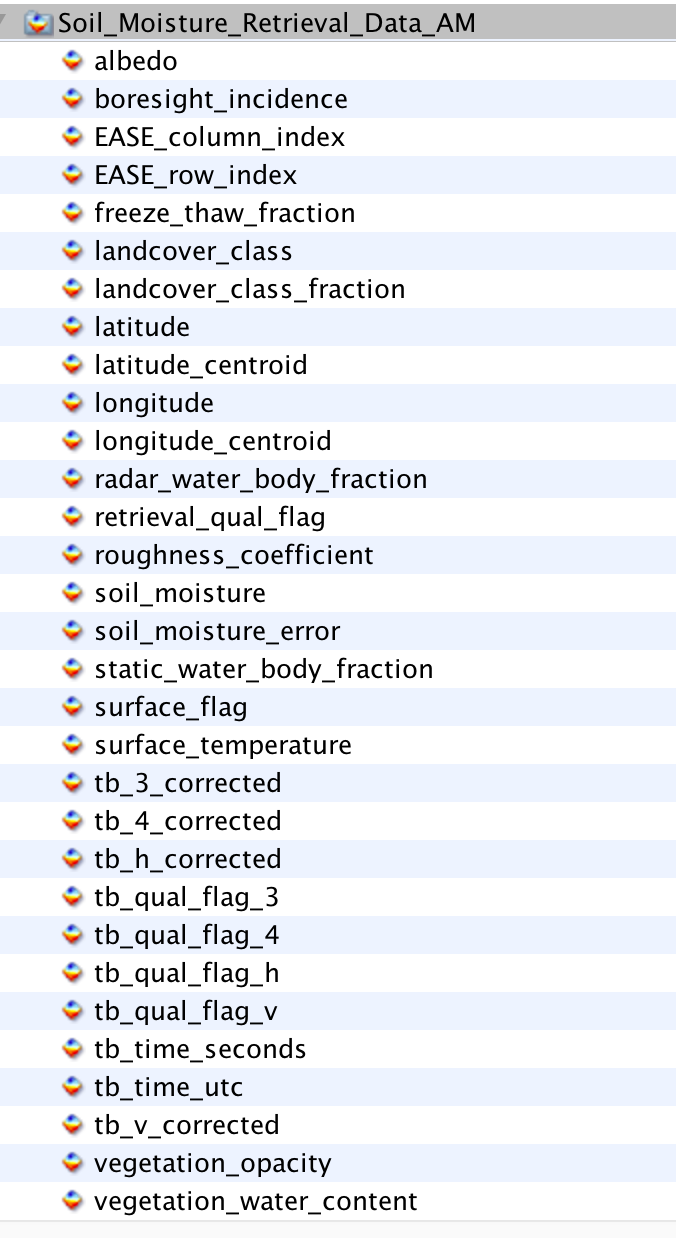


Figure . SMAP\_L3\_SM\_P Variables Representing the Morning (AM) Crossing

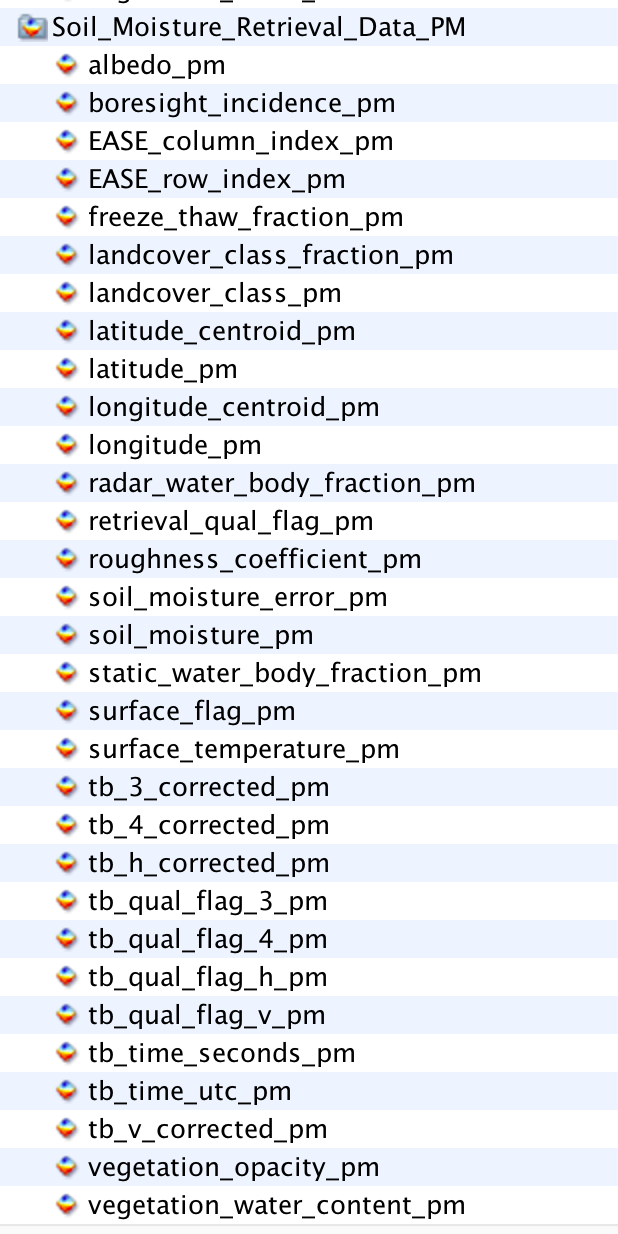


Figure . SMAP\_L3\_SM\_P Variables Representing the Afternoon (PM) Crossing

The benefit of using the Set class is to enable the CMR to preserve the order of the variables within the structure of the granule file.

**Tags**

*Recommended*

### SamplingIdentifiers

**Elements:**

SamplingIdentifiers (0..\*)

SamplingIdentifiers/SamplingMethod (1)

SamplingIdentifiers/MeasurementConditions (1)

SamplingIdentifiers/ReportingConditions (0..1)

**Description:**

Elements in this category are used for capturing information associated with sampling, including the method of sampling and the conditions at the time of measurement and reporting. SamplingMethod describes the name of the sampling method used for the measurement. An example of the SampleMethod includes 'radiometric detection within the visible and infra-red ranges of the electromagnetic spectrum'. MeasurementConditions and ReportingConditions are useful metadata for field campaign data sets. MeasurementConditions describes the conditions at the time the observation or measurement was recorded and the ReportingConditions describes the conditions over which the observation or measurement are valid. For example, MeasurementConditions could be 'Sampled Particle Size Range: 90 - 600 nm' and the ReportingConditions could be 'STP: 1013 mb and 273 K'.

**Sample Values:**

SamplingMethod: Test sampling method on Ship observations

MeasurementConditions: Test measurement conditions

ReportingConditions: Test reporting conditions

**Tags:**

*Recommended*

### AdditionalIdentifiers

**Element Specification:**

AdditionalIdentifiers (0..1)

AdditionalIdentifiers/Identifier (1)

AdditionalIdentifiers/Description (0..1)

**Description:**

An area to story any additional identifiers of a variable and its type such as a standard name or DOI.

**Sample Values:**

Identifier: CF\_Standard\_Description

Description: The sea surface subskin temperature is the temperature at the base of the conductive laminar sub-layer of the ocean surface, that is, at a depth of approximately 1 - 1.5 millimeters below the air-sea interface. For practical purposes, this quantity can be well approximated to the measurement of surface temperature by a microwave radiometer operating in the 6 - 11 gigahertz frequency range, but the relationship is neither direct nor invariant to changing physical conditions or to the specific geometry of the microwave measurements. Measurements of this quantity are subject to a large potential diurnal cycle due to thermal stratification of the upper ocean layer in low wind speed high solar irradiance conditions.

**Tags:**

*Optional, Free Text Search*

Appendix A Tags Glossary

Table 5 lists all tags used in this model and provides a description of the tags' usage.

Table . Tag's Glossary

|  |  |
| --- | --- |
| **Tag Name** | **Description** |
| Required | This element is required. |
| Free Text Search | This element will be indexed by the CMR as part of the Free Text Search. |
| Controlled Vocabulary | This element will have a vocabulary that will be used to validate the value. This will most likely be done via a vocabulary management service. |
| Recommended | This element is recommended. |

Appendix B Keywords and Measurements Governance Structure

The Governance Structure shown in Figure 26 is recommended for the selection of keywords and measurements. ESDIS chairs each of the measurement or keyword selection councils and provides overall science guidance, and the DAAC/Data Providers serve as the decision authority for the metadata associated with data sets sourced from their DAAC/Project.

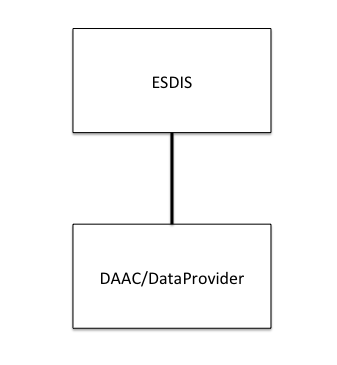


Figure . Suggested Governance Structure

Adding Keywords or Measurements are expected to be done by the Metadata Curator, via a GUI. Keywords are to be sourced from the GCMD Keywords and are controlled. What is being proposed here is not too different from the existing method used in the EDSC UI, with the exception being that the Keyword will be used for discovery at the Variable level, as opposed to the Collection or Granule level, which is currently the case. The challenge with Measurements is that they are uncontrolled. The concept is to start with a pre-seeded list of suggested measurements and, over time an alphabetically ordered list can be collected, by certain users, and by their use of a metadata management tool. The guidelines for adding keywords or measurements can be achieved by following the suggested steps below.

**Keywords**

Keywords may be selected from the GCMD Keywords set.

The GCMD Keywords are already subject to a strict governance process:

1. Review the controlled keyword/guidelines located at: https://cdn.earthdata.nasa.gov/conduit/upload/5182/KeywordsCommunityGuide\_Baseline\_v1\_SIGNED\_FINAL.pdf
2. Verify that the keyword does not already exist.
3. Map these to the appropriate variables.
4. Include a definition of the controlled keyword.

**Measurements**

Measurements may be selected from an array of standard sources, such as CSDMS, CF Conventions, etc.

The process by which measurements may be selected is simple.

* Determine level: i.e. Atmosphere, Oceans, Land (highest) or Atmosphere Air Temperature (mid), or Atmosphere Air Temperature Saturated Adiabatic Lapse Rate (lowest), etc.
* Determine whether the measurement is missing, and a new one is needed. For example, if we have Atmosphere Air Column Water Vapor and the next tag is Atmosphere Air Flow Azimuth Angle of Bolus Velocity, then Atmosphere Air Carbon Dioxide (and its derivatives) are missing.
* Select the most appropriate measurement to suit the need. If the measurement does not exist, apply crosswalk to another standard such as CSDMS to CF convention Standard Names.
* Add to the measurements list stored in the CMR so that all future users can use this measurement.
* Map these to the appropriate variables.
* Include a definition of the uncontrolled measurement.

Appendix C Analysis of CSDMS and CF Standard Names as a Source of Tagging

**Analysis of CSDMS Standard Names as a source of tagging**

The Community Surface Dynamics Modeling System (CSDMS) modeling framework provides mechanisms that allow models and data sets from different contributors (i.e. from different geoscience domains: hydrology, oceanography, meteorology, seismology) to coexist. Each geoscience domain has its own descriptive vocabulary to describe specific variables.  This presents a problem when trying to use the data sets together or when trying to discover the same variables that have different names. How do you find a variable by name when one knows it by a different name? The framework defines a unique and holistic approach to the semantic mediation problem by offering a set of standardized and precise descriptions for each variable and by giving a number of options to resolving which names and abbreviations are to be used for a variable.

The naming conventions of the CSDMS Standard Names are based on object-oriented principles and group the variable names by the following categories: Atmosphere, Atoms, Automobiles, Basins, Bedrock, Channel, Chocolate, Compounds and Mixtures, Earthquakes, Glaciers, Materials, Models, Molecules, Oceans, Planets, Projectiles, Radiation, River Deltas, Sea Ice, Snow, Soil, Sea Floor Debris, Topography and Water Tank. Only a subset of these categories are suitable for the Earth Observing System (EOS) mission and they include the following: Atmosphere, Basins, Channel, Earthquakes, Glaciers, Oceans, Radiation, River Deltas, Sea Floor Debris, Sea Ice, Soil, Snow, and Topography.  These Standard Names can be chosen as the primary source of tagging, since each group is highly relevant to the science domains which are covered by the EOS data sets and those likely to be covered in the future. The CSDMS Standard Names exhibit the Object Name plus Model Name Pattern structure to the name. An example of an Object Name is atmosphere\_water.  Examples of the corresponding Model Name Patterns are: domain time integral of precipitation leq volume flux, icefall mass per volume density, precipitation duration, precipitation leq volume flux, and precipitation mass flux. The first example of combining Object Name plus Model Name yields the resultant standard name: atmosphere water domain time integral of precipitation leq volume flux. An example of a missing name would be: atmosphere water precipitation, or the more common term, precipitation.

**Analysis of CF Standard Names as a source of tagging**

CF conventions for climate and forecast metadata are designed to promote the processing and sharing of files created with the netCDF Application Programmer Interface. The CF conventions generalize and extend the Cooperative Ocean/Atmosphere Research Data (COARDS) conventions. Most of the CF standard names have been derived from guidelines which have drawn on the European Center for Medium-Range Weather Forecasts (ECMWF), and National Centers for Environmental Prediction GRIdded Binary file format (NCEP GRIB) tables, the Program for Climate Model Diagnosis and Intercomparison (PCMDI) and GCMD. CF standard names consist of lower-letters, digits and underscores, and begin with a letter. Upper case is not used and US spelling is used, e.g. vapor, sulfur. The CF Standard Names can be chosen as a source of tagging supplementary to CSDMS. Examples of CF Standard Names are: precipitation amount, and precipitation flux, and precipitation flux onto canopy are included in the CF Standard Names and not in the CSDMS Standard Names. In this simple example, both CSDMS and CF Standard Names may be used as a source of tagging for search terms to locate all variables associated with the measurement: precipitation.

Appendix D Definitions of Terms

**CONCEPT**:

Measurement

**DEFINITION**:

The act or process of measuring an observable property, usually geophysical, geo-biophysical, physical, or chemical. In the case of air temperature, for instance, the object of the measurement is air and the property being measured is temperature. For models, it is a simulated observable property.  
Using Scott Peckham's model as a basis for a measurement naming convention, the Measurement names can be expressed as: <<object, quantity>>, object = "Aerosol", quantity = "Optical Depth".

Examples: Aerosol Optical Depth, Air Temperature, Surface Albedo, Solar Irradiance, Surface Reflectance, Atmospheric Moisture, Methane Concentration, Sulphur Dioxide Concentration, Ozone Concentration.

**CONCEPT**:

Variable

**DEFINITION**:

A named set of data that contains the recorded values of a measurement. In this context, the variable is described by its name and characteristics. For instance, a variable contained within the MYD08\_M3V5 dataset is called: Optical\_Depth\_Land\_Maximum. There are other variables in the set, including variables which contain information about geographic position and quality.

The description of the variable may include what was intended to be measured, i.e., the observable property, and how the variable was measured, such as measurement technique and the instrument used.

Variables may be classified as science variables, quality variables and ancillary variables (or other, when one of these classifications cannot be used). A variable can also be the output of a model.

Examples: Aerosol Optical Depth 550nm (Dark Target), Aerosol Optical Depth 550nm (Deep Blue, Land Only), Air Temperature (Daytime/Ascending), Air Temperature at 2m, Air Temperature at Surface (Daytime/Ascending), Air Temperature at Surface (Nightime/Descending), Relative Humidity (Daytime/Ascending), Relative Humidity (Nightime/Descending), Water Vapor Mass Mixing Ratio (Daytime/Ascending), Water Vapor Mass Mixing Ratio (Nightime/Descending), Methane Total Column (Nightime/Descending), SO2 Column Mass Density, SO2 Column Mass Concentration, Ozone - reported in parts per billion by volume.

Appendix E Examples

**Description:**

This section contains some variable metadata worked examples.

{

"Name": "sea\_surface\_temperature",

"StandardName": "sea\_surface\_foundation\_temperature",

"AdditionalIdentifiers": [

{

"Identifier": "CF\_Standard\_Description",

"Description": "The sea surface subskin temperature is the temperature at the base of the conductive laminar sub-layer of the ocean surface, that is, at a depth of approximately 1 - 1.5 millimeters below the air-sea interface. For practical purposes, this quantity can be well approximated to the measurement of surface temperature by a microwave radiometer operating in the 6 - 11 gigahertz frequency range, but the relationship is neither direct nor invariant to changing physical conditions or to the specific geometry of the microwave measurements. Measurements of this quantity are subject to a large potential diurnal cycle due to thermal stratification of the upper ocean layer in low wind speed high solar irradiance conditions."

}

],

"LongName": "sea surface subskin temperature",

"Definition": "sea surface subskin temperature in units of kelvin",

"VariableType": "SCIENCE\_VARIABLE",

"VariableSubType": "SCIENCE\_ARRAY",

"Units": "kelvin",

"DataType": "short",

"Scale": 0.01,

"Offset": 273.15,

"Dimensions": [

{ "Name": "time", "Size": 1, "Type": "TIME\_DIMENSION"},

{ "Name": "nj", "Size": 3072, "Type": "ALONG\_TRACK\_DIMENSION"},

{ "Name": "ni", "Size": 4096, "Type": "CROSS\_TRACK\_DIMENSION"}

],

"IndexRanges" : {

"LatRange" : [ -59.875, 89.875 ],

"LonRange" : [ -179.875, 179.875 ]

},

"SamplingIdentifiers": [

{ "SamplingMethod": "Test sampling method on Ship observations",

"MeasurementConditions": "Test measurement conditions",

"ReportingConditions": "Test reporting conditions"

}

],

"ScienceKeywords": [

{ "Category": "EARTH SCIENCE", "Topic": "SPECTRAL/ENGINEERING", "Term": "MICROWAVE", "VariableLevel1": "SEA SURFACE TEMPERATURE", "VariableLevel2" : "MAXIMUM/MINIMUM TEMPERATURE", "VariableLevel3" : "24 HOUR MAXIMUM TEMPERATURE", "DetailedVariable" : "details\_4385"},

{ "Category": "EARTH SCIENCE", "Topic": "SPECTRAL/ENGINEERING", "Term": "MICROWAVE", "VariableLevel1": "MICROWAVE IMAGERY"}

],

"ValidRanges": [

{ "Min": -2005, "Max": 45005, "CodeSystemIdentifierMeaning": ["Test Code System 2."], "CodeSystemIdentifierValue": ["CSIV 2."]}

],

"FillValues": [

{ "Value": -327685, "Type": "SCIENCE\_FILLVALUE" }

],

"Sets": [

{ "Name": "Data\_Fields", "Type": "General", "Size": 18 , "Index": 11 }

],

"MeasurementIdentifiers": [

{

"MeasurementContextMedium": "ocean",

"MeasurementContextMediumURI": "http://www.ontobee.org/ontology/ENVO?iri=http://purl.obolibrary.org/obo/ENVO\_01000324",

"MeasurementObject": "sea\_surface\_subskin",

"MeasurementObjectURI": "https://gcmd.earthdata.nasa.gov/kms/concept/68a09c56-be36-4100-8757-3a6eec7dc251",

"MeasurementQuantities": [

{"Value": "temperature", "MeasurementQuantityURI": "http://www.ontobee.org/ontology/PATO?iri=http://purl.obolibrary.org/obo/PATO\_0000146"}

]

}

]

}

{

"Name": "/gt1l/land\_segments/canopy/h\_canopy",

"StandardName": "canopy\_height",

"LongName": "height canopy",

"Definition": "98% height of all the individual canopy relative heights for the segment above the estimated terrain surface. Relative canopy heights have been computed by differencing the canopy photon height from the estimated terrain surface.",

"VariableType": "SCIENCE\_VARIABLE",

"Units": "meters",

"DataType": "float32",

"Scale": 1.0,

"Offset": 0.0,

"Dimensions": [

{ "Name": "/gt1l/land\_segments/delta\_time", "Size": 15201, "Type": "TIME\_DIMENSION"}

],

"FillValues": [

{ "Value": 3.402823466E38, "Type": "SCIENCE\_FILLVALUE" }

],

"Sets": [

{ "Name": "Data\_Fields", "Type": "General", "Size": 7, "Index": 0 }

]

}

{

"Name": "/science/grids/data/amplitude",

"LongName": "Amplitude",

"Definition": "2D Amplitude of IFG",

"VariableType": "SCIENCE\_VARIABLE",

"Units": "watt",

"DataType": "float32",

"Scale": 1,

"Offset": 0,

"Dimensions": [

{

"Name" : "latitude",

"Size" : 2166,

"Type" : "LATITUDE\_DIMENSION"

}, {

"Name" : "longitude",

"Size" : 4061,

"Type" : "LONGITUDE\_DIMENSION"

} ],

"ValidRanges" : [ {

"Min" : 0

} ],

"Sets": [

{ "Name": "data", "Type": "data", "Size": 4, "Index": 4 }

]

}

Appendix F Abbreviations and Acronyms

|  |  |
| --- | --- |
| AESIR | Application friendly EOSDIS Science Information Retriever |
| BODC | British Oceangraphic Data Centre |
| CF | Climate and Forecast metadata |
| CMR | Common Metadata Repository |
| COARDS | Cooperative Ocean/Atmosphere Research Data Service |
| CSDMS | Community Surface Dynamics Modeling System |
| DAAC | Distributed Active Archive Center |
| ECMWF | The European Center for Medium-Range Weather Forecasts |
| ECS | EOSDIS Core System |
| EDSC | Earthdata Search Client |
| EED | EOSDIS Evolution and Development |
| EOS | Earth Observing System |
| EOSDIS | Earth Observing System Data and Information System |
| ESDIS | Earth Science Data and Information System |
| ESO | Earth Science Office |
| GCMD | Global Change Master Directory |
| GES DISC | Goddard Earth Sciences Data and Information Services Center |
| GIBS | Global Imagery Browse Services |
| GRIB | GRIdded Binary file format |
| HDF | Hierarchical Data Format |
| ISO | International Organization for Standardization |
| KMS | Keyword Management System |
| LAADS | Level-1 and Atmosphere Archive and Distribution System |
| MMT | Metadata Management Tool |
| MODIS | Moderate Resolution Imaging Spectroradiometer |
| NASA | National Aeronautics and Space Administration |
| NCEP | National Centers for Environmental Prediction |
| PCMDI | Program for Climate Model Diagnosis and Intercomparison |
| PI | Principal Investigator |
| UI/UX | User Interface/User Experience |
| UML | Unified Modeling Language |
| UMM | Unified Metadata Model |
| UMM-C | Unified Metadata Model - Collections |
| UMM-Common | Unified Metadata Model - Common Elements |
| UMM-G | Unified Metadata Model - Granules |
| UMM-S | Unified Metadata Model - Services |
| UMM-Var | Unified Metadata Model - Variables |
| URL | Uniform Resource Locator |
| XML | Extensible Markup Language |
| XPath | XML Path Language |
| XSLT | Extensible Stylesheet Language Transformations |