423-RQMT-003-E, Original, Revision -
Earth Science Data and Information Systems (ESDIS)
Project, Code 423

Appendix E. Metadata Requirements Base Reference for Unified Metadata Model Variables (UMM-Var)

Appendix E. Metadata Requirements Base Reference for United Metadata Model Variables (UMM-Var)

Signature/ Approval Page

Prepared by:

Signature obtained on file ***10/30/2018***
Valerie Dixon — Date
ESDIS CMR Manager
GSFC Code 586

Reviewed by:

Signature obtained on file ***11/05/2018***
Karen Michael — Date
EOSDIS System Manager
GSFC Code 423

Signature obtained on file ***10/30/2018***
Jeanne Behnke — Date
ESDIS Deputy Manager for Science Operations
GSFC Code 423

Approved by:

Signature obtained on file ***10/31/2018***
Andrew Mitchell — Date
ESDIS Project Manager
GSFC Code 423

[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. Changes to this document shall be verified by a document change notice (DCN) and implemented 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 NASA Earth Science community and addresses the need for describing the types of variables that exist within data products that are described by the UMM-G metadata records. Developers, engineers and architects should reference this document and the UMM as a guide while implementing 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 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-Common, UMM-C, UMM-S, UMM-G, Variables, NASA Earthdata Search, EOSDIS, ESDIS, CMR

Change History Log

Revision	Effective Date	Description of Changes
V0.8.0	July 2015	• Provisional NASA Internal Review Release.
Baseline, Original, Rev -	10/31/2018	CCR 423-ESDIS-193; approved 10/10/2018 • Baseline in COMET as stand-alone Appendix E to the Metadata Requirements Base Reference as Revision B. Pages: E-1 through E-79

Table of Contents

List of Figures

List of Tables

E.1 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 Common Metadata Repository (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>>, e.g. 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.

The description of the variable may include what was intended to be measured (i.e., the observable property, and how the variable was measured (e.g., 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.

- Sample Illustration:
  - Example 1:
    - Measurement:
      - Aerosol Optical Depth
    - Variables:
      - Aerosol Optical Depth 550nm (Dark Target) - this is a science variable example
      - Aerosol Optical Depth 550nm (Deep Blue, Land Only) - this is another science variable example
      - Deep_Blue_Aerosol_Optical_Depth_550_Land_QA - this is a quality variable example

        - Deep_Blue_Algorithm_Flag_Land - this is an ancillary variable example

    - Example 2:

        - Measurement:

            - Ozone Mixing Ratio reported in parts per billion by volume

        - Variable:

            - O3_ppbv - this is a science variable example

    - Example 3:

        - Measurement:

            - Integrated Column NO2 loading

        - Variables:

            - NO2_Column - this is a science variable example

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, e.g. Aerosol Optical Depth 550nm (Dark Target), Aerosol Optical Depth 550nm (Deep Blue, Land Only), along with its associated data quality variables, and ancillary variables (e.g. latitude, longitude information).

E.1.1 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, e.g. 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 User Interface (UI) for typical data transformations, e.g. 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 National Aeronautics and Space Administration (NASA) Earth Science community. Distribution is unlimited.

E.1.2 Scope

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

E.1.3 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 DAACs, non-DAAC data providers, instrument PIs, 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.

E.1.4 Copyright Notice

E.1.5 Feedback

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

E.1.6 Document Conventions

Each section of this document describes an element of the model. Elements that feature in multiple models are fully documented in the UMM-Common model. Any references to the UMM-Common model are included, as required. Elements that are unique to the Variable model are documented in the following way:

- Element Specification: Identifies the individual fields for 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.
- Cardinality: Indicates the specification of counts for this element, summarized in the following table:

Table E.1: 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
1..N	At least one of this element is required, up to and including N number of this element may be present

Interaction diagrams presented in this document are based on the UML Notation.

E.1.7 Related Documentation

There is a document that fully describes metadata elements that are used in multiple models. This document, the UMM-Common (https://wiki.earthdata.nasa.gov/display/CMR/CMR+Documents) is documented separately. The Variable model makes several references to the UMM-Common throughout.

The latest versions of all documents below should be used. The latest ESDIS Project documents can be obtained from 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.

E.1.7.1 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.

423-RQMT-003	Metadata Requirements Base Reference Document
CMR Life Cycle	https://wiki.earthdata.nasa.gov/display/CMR/CMR+Documents
423-RQMT-003-A UMM-Common	Metadata Requirements Base Reference Appendix A UMM-Common https://wiki.earthdata.nasa.gov/display/CMR/CMR+Documents
SERF	https://gcmd.nasa.gov/Aboutus/xml/serf/serf.xsd https://gcmd.nasa.gov/add/serfguide/index.html
CMR End-To-End Services Study (Task 25) EED2-TP-025	https://drive.google.com/open?id=0BzJ0Mge7A2GEQ0NlZ0tnQktlbnM

E.1.7.2 Reference Documents

The following documents are not binding on the content but referenced herein and amplify or clarify the information presented in this document.

Tags	http://en.wikipedia.org/wiki/Tag_%28metadata%29
Translators	Translators to ISO can be found at https://cdn.earthdata.nasa.gov/iso/resources/transforms/
XPath	XPath is a language for addressing parts of an XML document, designed for use with XSLT.
XLinks	http://en.wikipedia.org/wiki/XLink
MENDS	More information on the Base Metadata Requirements established by the MENDS group: https://wiki.earthdata.nasa.gov/display/NASAISO/NASA+Base+Metadata+Requirements http://www.w3schools.com/schema/el_choice.asp

E.2 UNIFIED METADATA MODEL - VARIABLES

E.2.1 Variable Context Diagram and Metadata Model Relationships

Figure E.1 shows the UMM-Var metadata model at a high level and its relationships with other key models: Collection (UMM-C), Granule (UMM-G), and Service (UMM-S). It is a high-level diagram, with abbreviated models, showing the key associations between the new UMM-Var and the other models in the UMM. Note that the figure specifically highlights the Variable model's role in the context of the Unified Metadata Model.

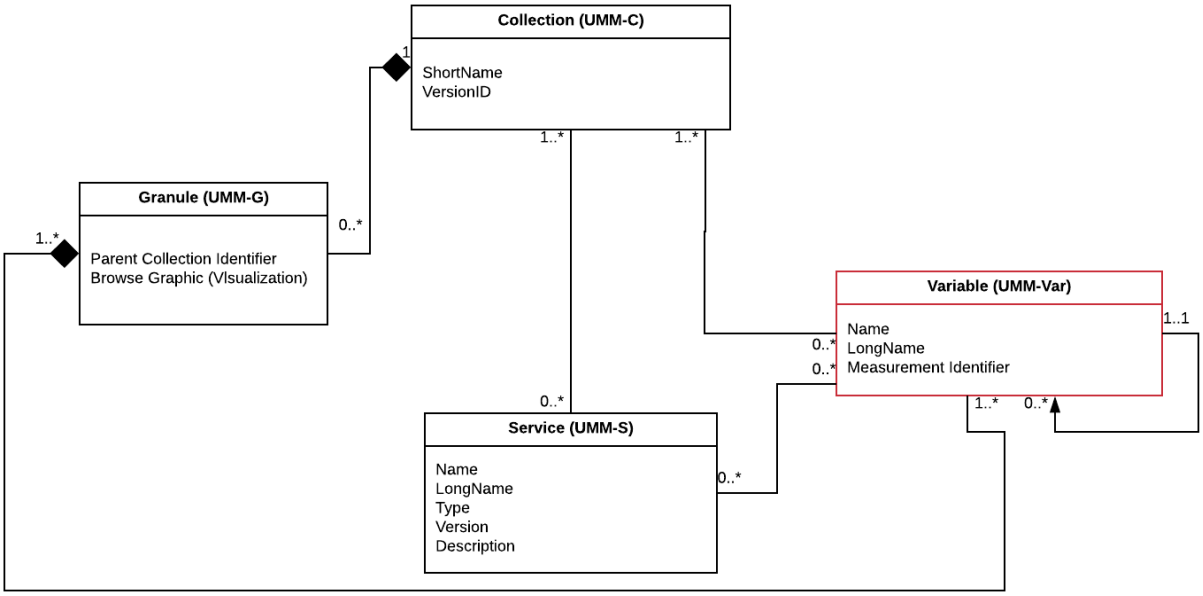

Figure E.1: Variable (UMM-Var) as part of the UMM context diagram

E.2.2 Use Cases

This section provides information about use cases identified for the UMM-Var. In keeping with the UML methodology, we provide a use case diagram showing the actor's interaction with the system. An activity diagram shows the flow of data, in terms of the user experience, and a sequence diagram which shows the sequences of actions within the system, and the key components of the system.

E.2.2.1 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, grouped by measurement.

See the use case diagram below.

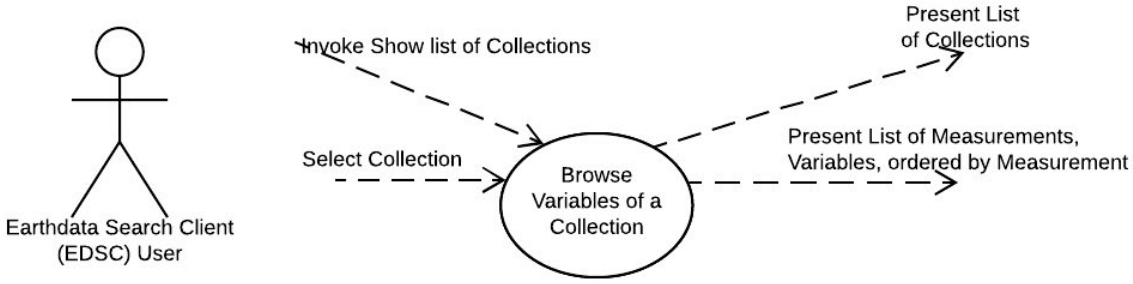

Figure E.2: Use Case: Browse Variables of a Collection

See the user experience activity diagram below.

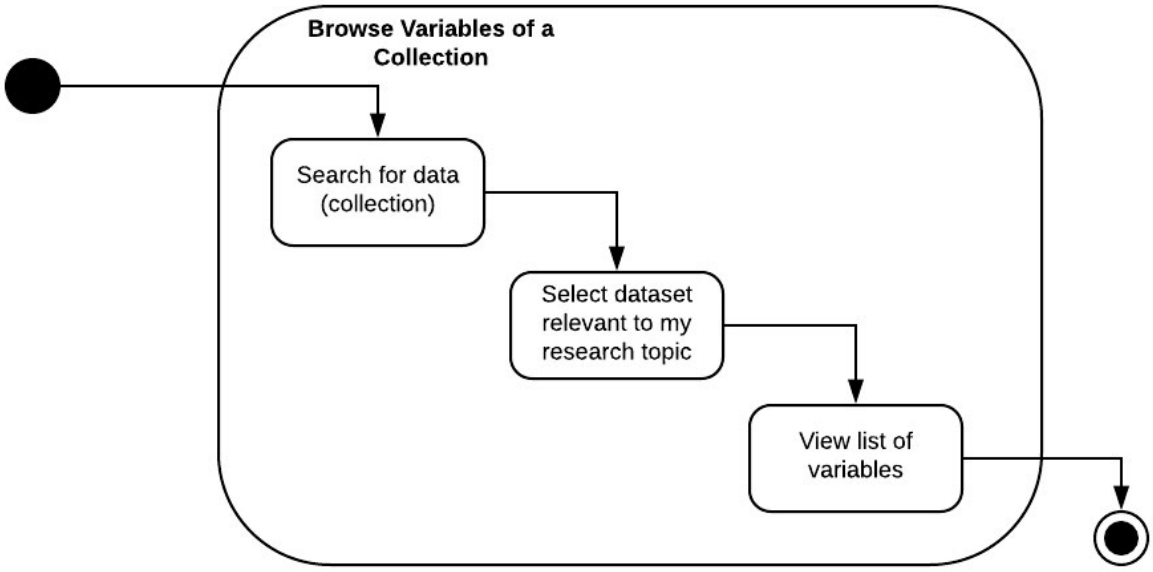

Figure E.3: Activity Diagram: Browse Variable of a Collection

See the system workflow sequence diagram below.

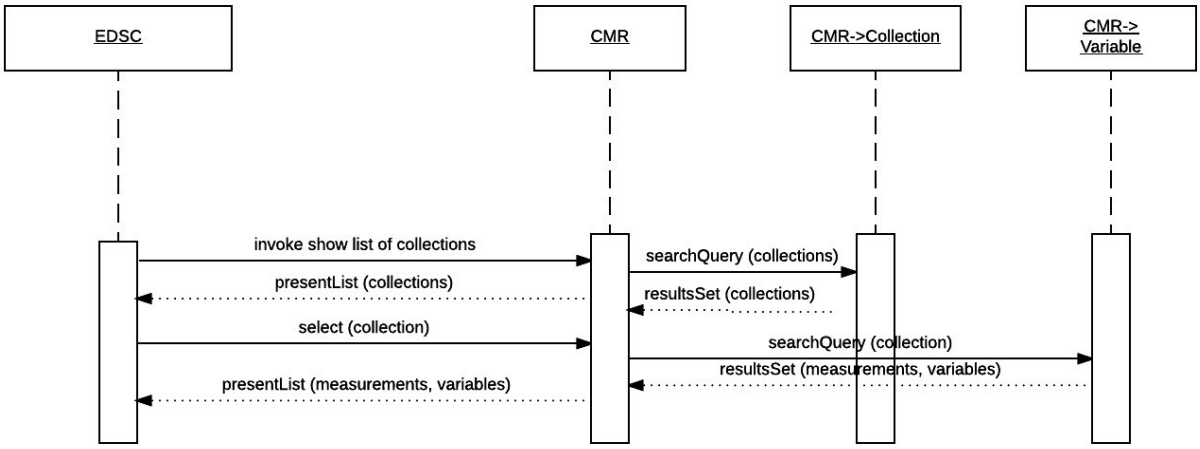

Figure E.4: Sequence Diagram: Browse Variables of a Collection

E.2.2.2 Faceted Browse

Scenario [a]: A user of a search tool, i.e. Earthdata Search Client (EDSC - https://search.earthdata.nasa.gov), can get a list of Measurement facets from the CMR.
Scenario [b]: A user of the EDSC can click on a "Measurement" facet value and constrain the lists to the collections that match the selected Measurement type and any other constraints (i.e. spatial, temporal) that have been selected.

Outcomes: A user of the EDSC, with no knowledge of the types of Measurements available within the CMR, can get a list of types of Measurements, and can further constrain the lists to Collections which match by selecting that type of Measurement, and any other constraints.
Note: The user can select a Measurement type to see the list of Variables available for that Measurement type, and an associated list of quality and ancillary variables. The association between Measurement types and Variables may be made previously by the Metadata Curator using the Metadata Management Tool (MMT), or a client specific to that provider.

Note: Scenarios [a] and [b] are related, in that the first represents a starting point for faceted browse, whereby the user can see a list of facets. Typically, the user wants to narrow the search criteria as much as possible, so starting with the list of facets (scenario[a]), the user selects one or more facets (scenario[b]) and via the EDSC UI, this yields a search result which displays those variables available in association with a specific Measurement type.

See the use case diagram below

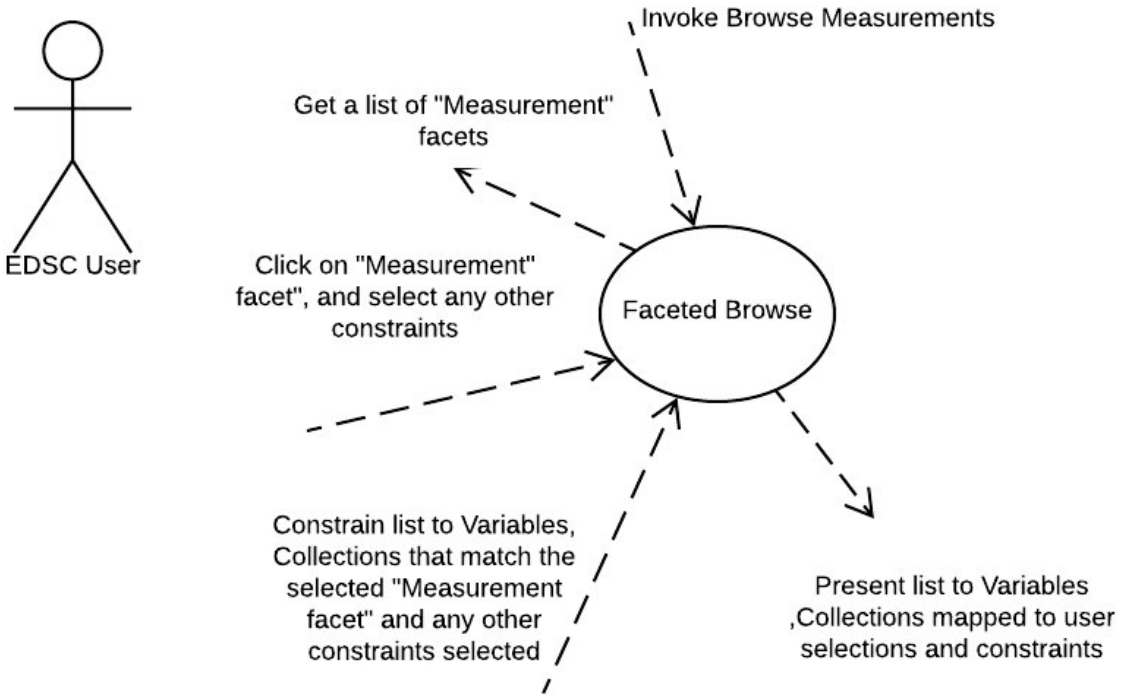

Figure E.5: Use Case: Faceted Browse

See the user experience activity diagram below

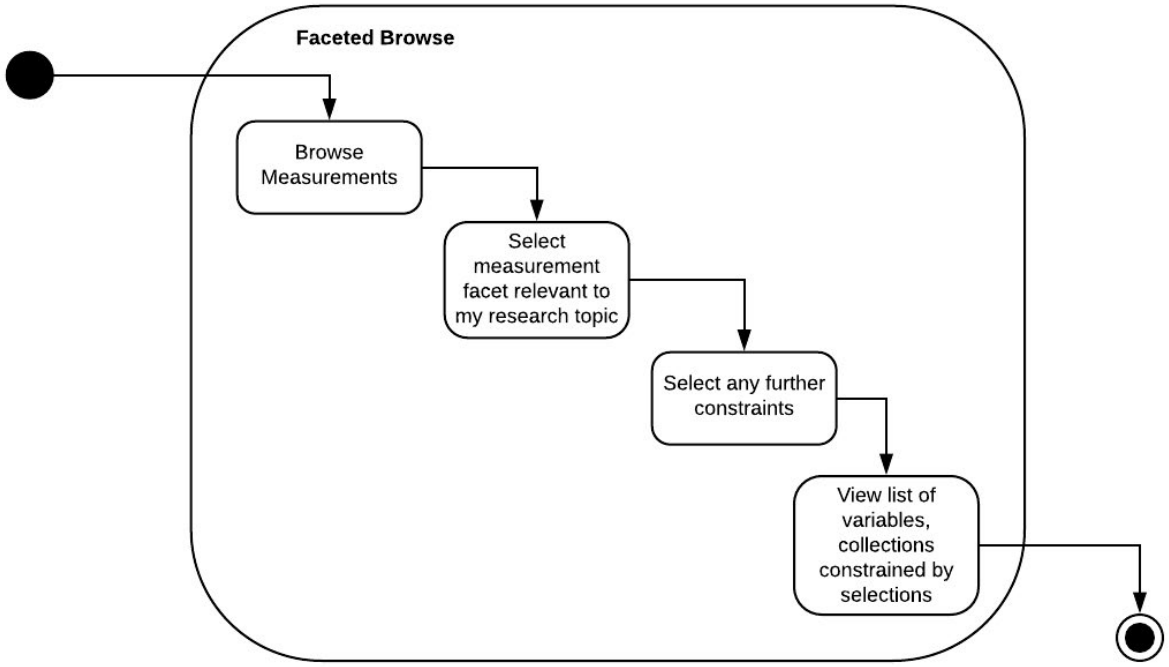

Figure E.6: Activity Diagram: Faceted Browse

See the workflow sequence diagram below

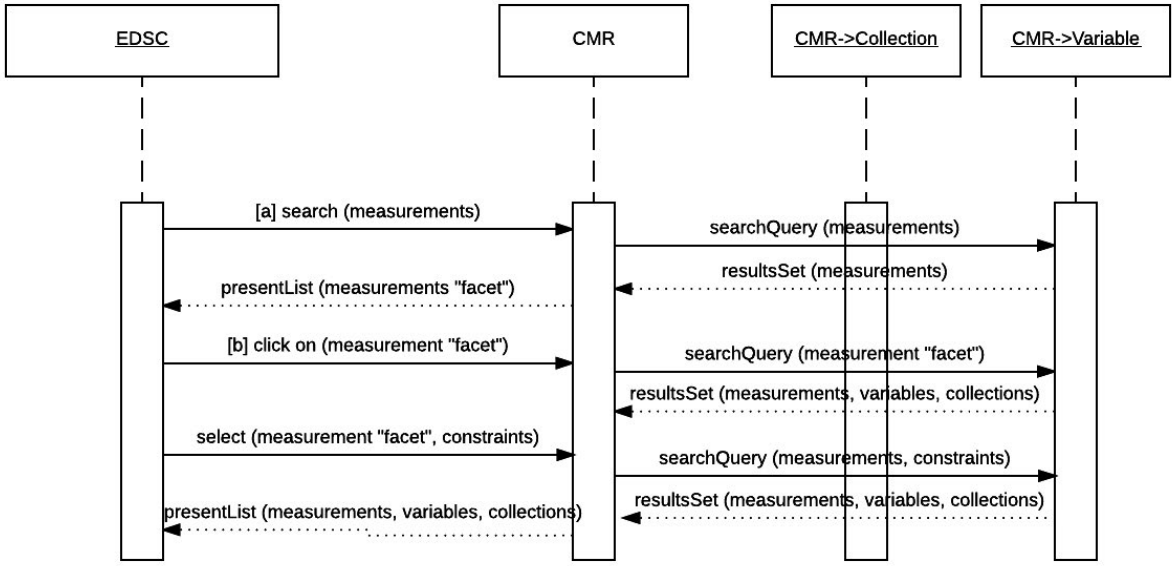

Figure E.7: Sequence Diagram: Faceted Browse

E.2.2.3 Update Variable Associations

Scenario [a]: A user of the CMR client, such as the MMT or other suitable metadata curation tool, can associate multiple variables with a collection.
Scenario [b]: A user of the CMR client, can submit multiple collections and all of the variables listed for each collection.
Scenario [c]: A metadata curator can populate the list of valid measurements associated with variable with selections from the Science Keywords hierarchy, Community Surface Dynamics Modeling System (CSDMS) or NetCDF Climate and Forecast (CF) metadata convention.

Outcomes: The curator seeded the CMR with new valid measurements chosen from the Science Keywords hierarchy and allowed editors to maintain variable and collection associations. A metadata curation tool, e.g. MMT, or another suitable metadata curation tool, may be used to maintain/update the variable and collection associations.

See the use case diagram below.

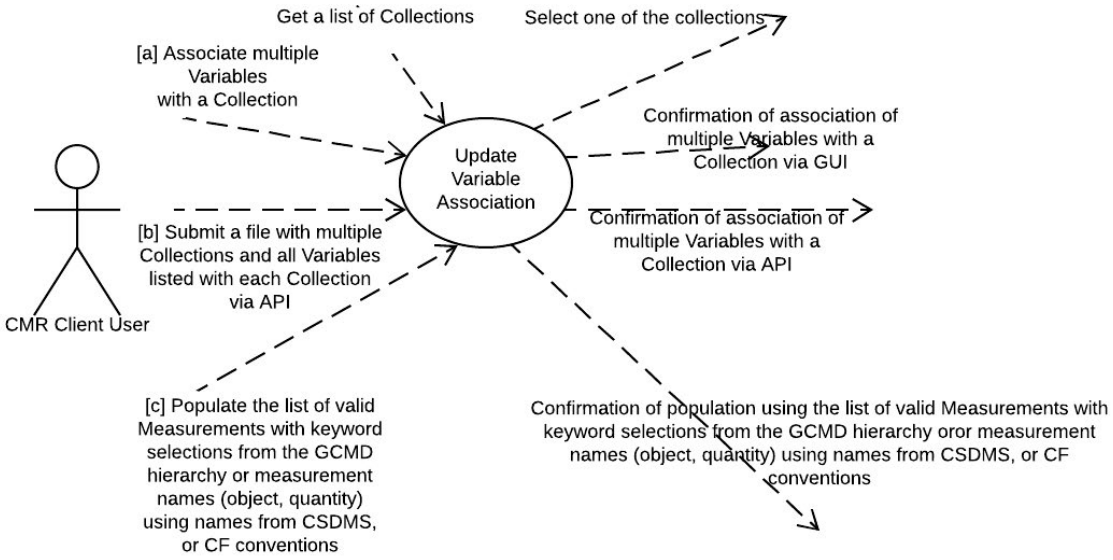

Figure E.8: Use Case: Update Variable Associations

See the user experience activity diagram below.

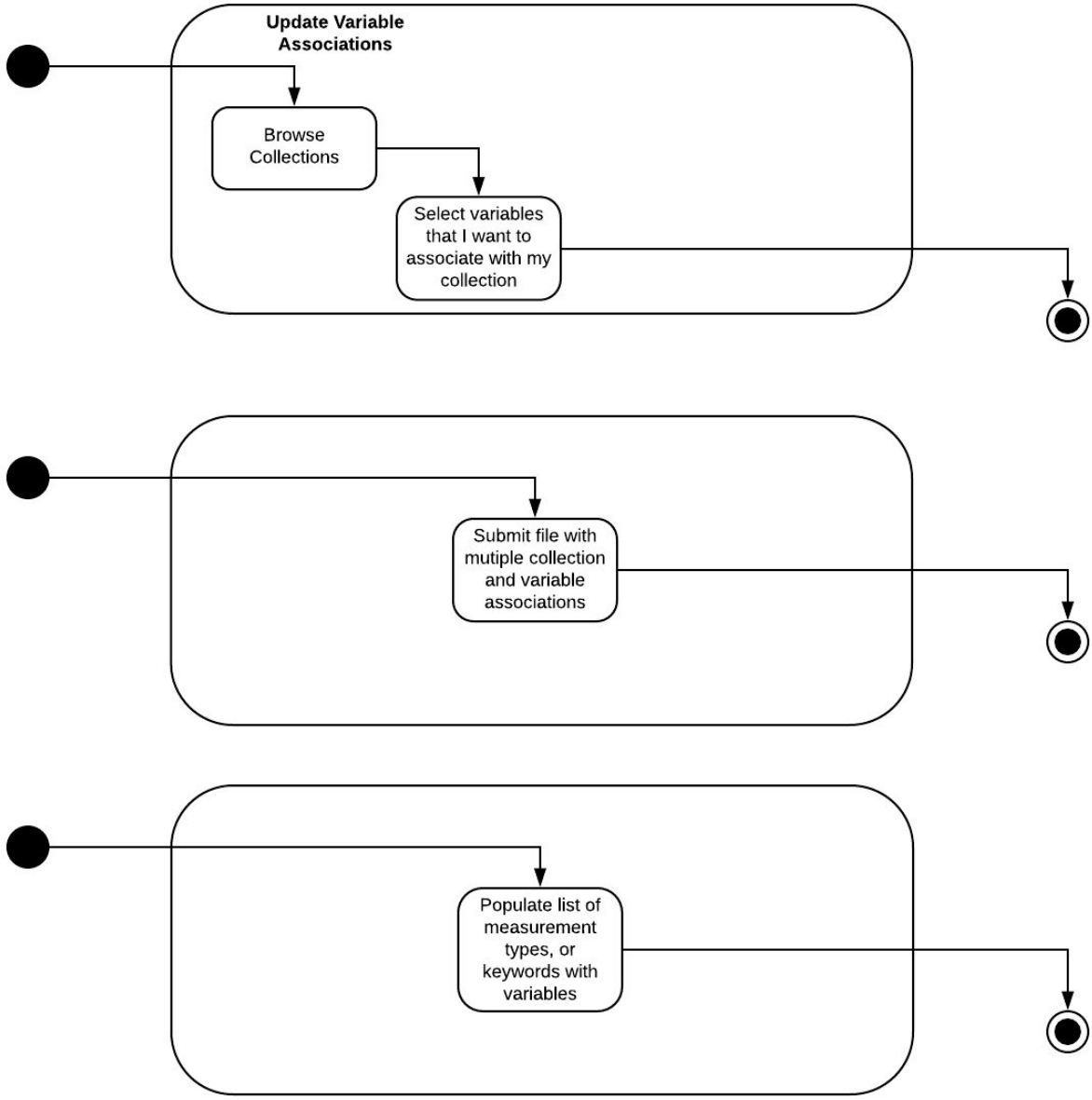

Figure E.9: Activity Diagram: Update Variable Associations

See the workflow sequence diagram below.

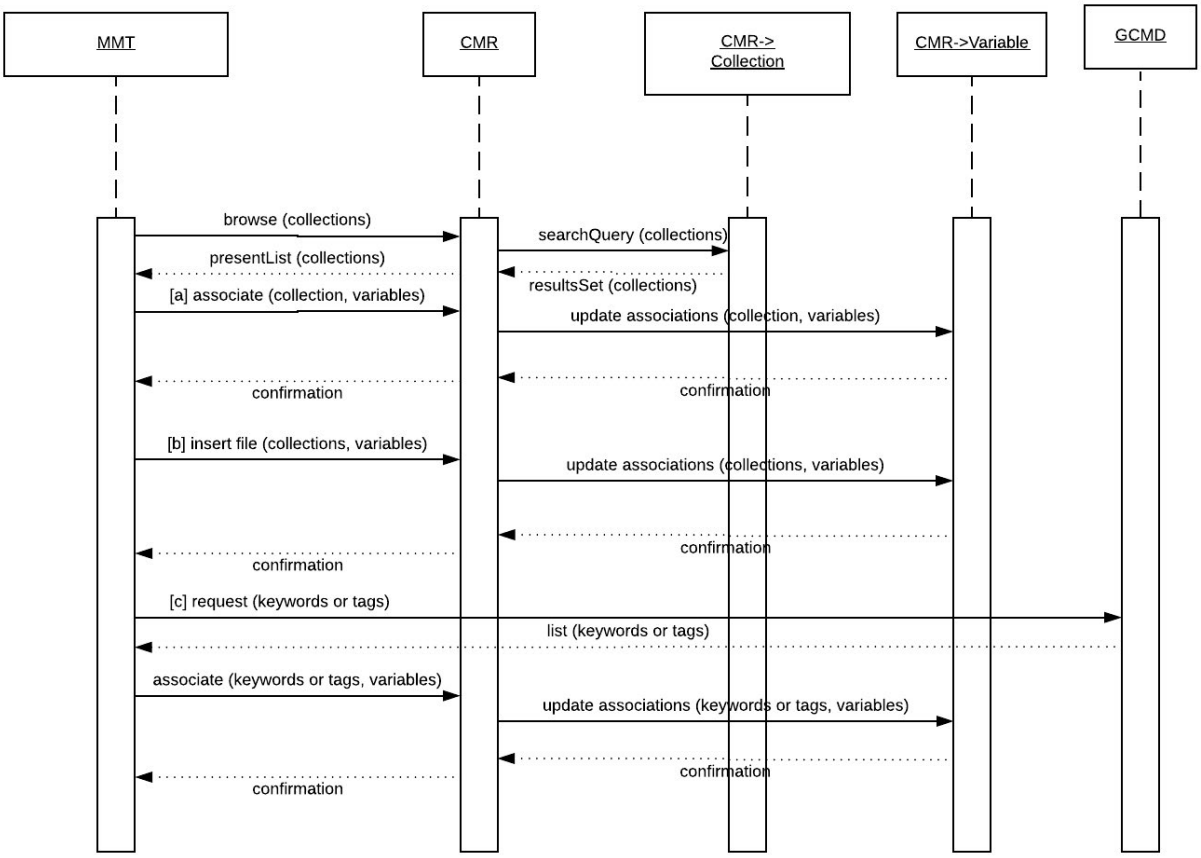

Figure E.10: Sequence Diagram: Update Variable Associations

E.2.2.4 Search Relevancy Ranking

Scenario: As a search engine (CMR), I can rank collections with a high relevance ranking when one or more of the search words appear in the measurement terms associated with the variables in the collection, as opposed to more generic fields such as the summary or references.

Outcomes: Returns to the user a list of Collections ranked by relevancy to the words used in the search term matching a measurement term.

See the use case diagram below.

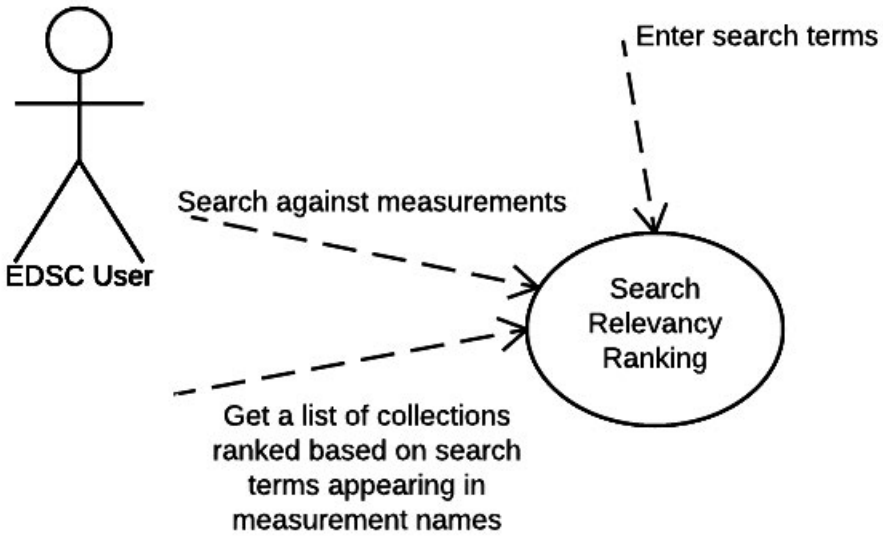

Figure E.11: Use Case: Search Relevancy Ranking

See the user experience activity diagram below.

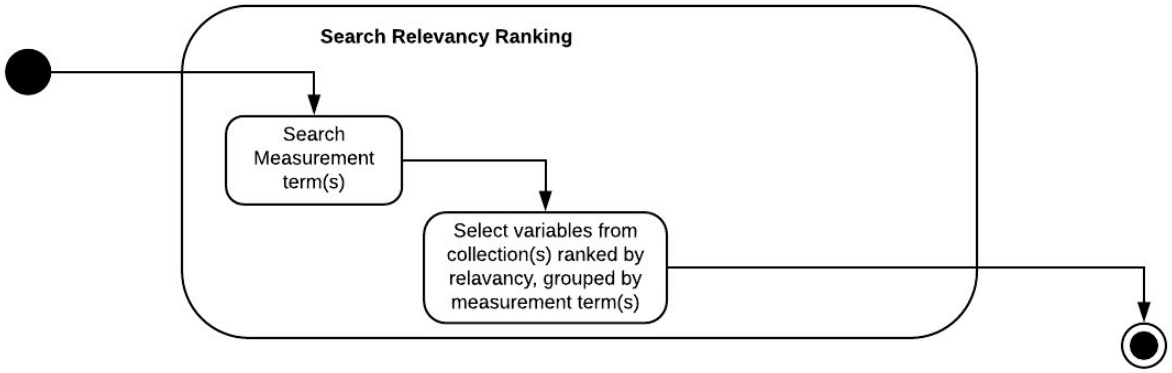

Figure E.12: Activity Diagram: Search Relevancy Ranking

See the workflow sequence diagram below.

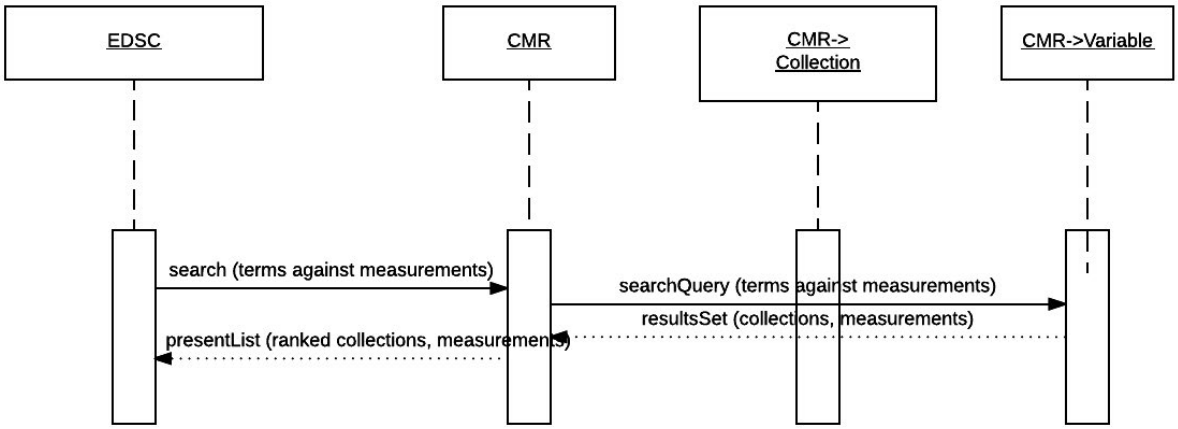

Figure E.13: Sequence Diagram: Search Relevancy Ranking

E.2.2.5 Cross-site Data Subsetting

Scenario: As a subsetting GUI, I can present the variables for a given collection in a logically categorized way, such as by measurement, 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 and categorized by measurement. 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. GES DISC, 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, 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.

See the use case diagram below.

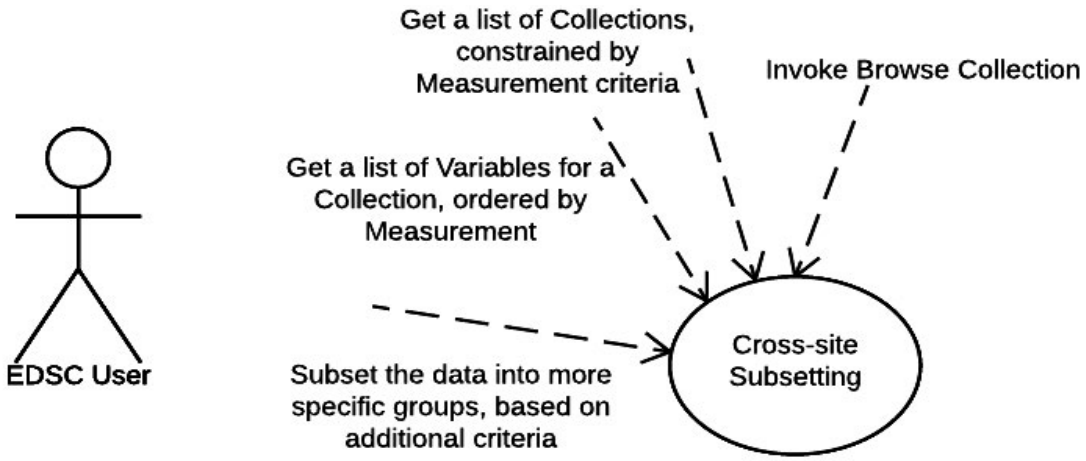

Figure E.14: Use Case: Cross-site Data Subsetting

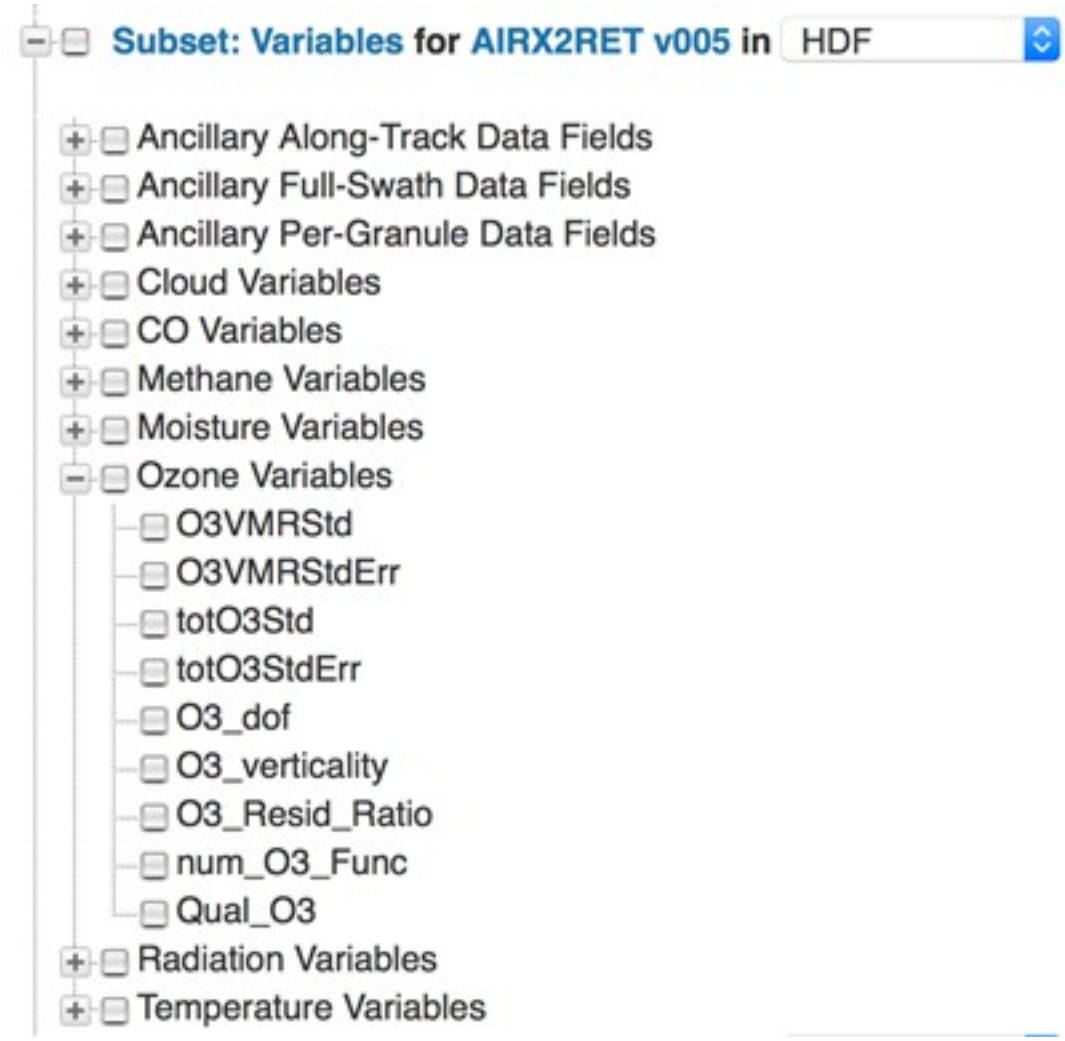

Figure E.15: User interface view for a user to choose a subset of variables for the AIRX2RET collection

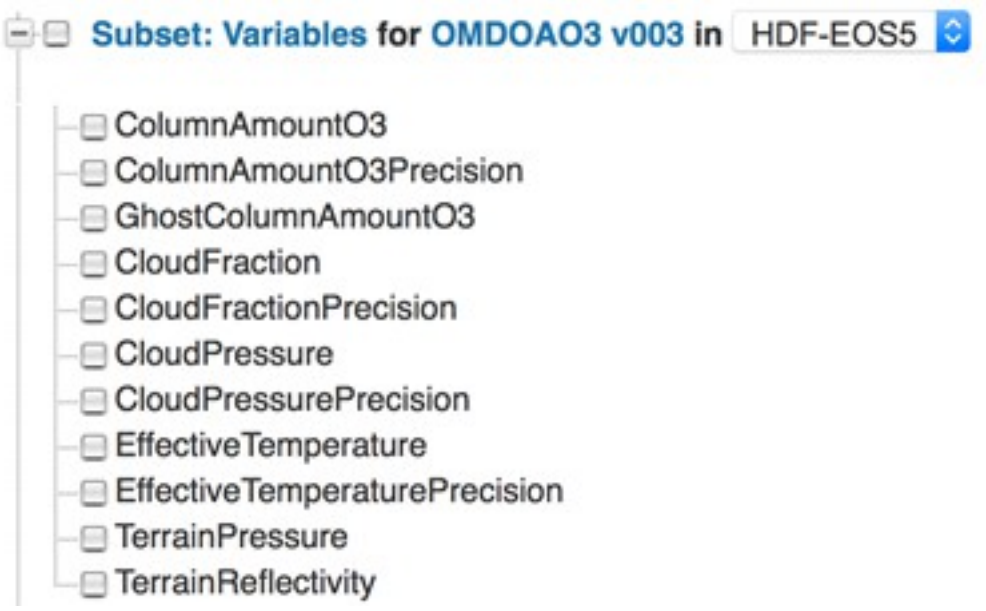

Figure E.16: User interface view for a user to choose a subset of variables for the OMDOAO3 collection

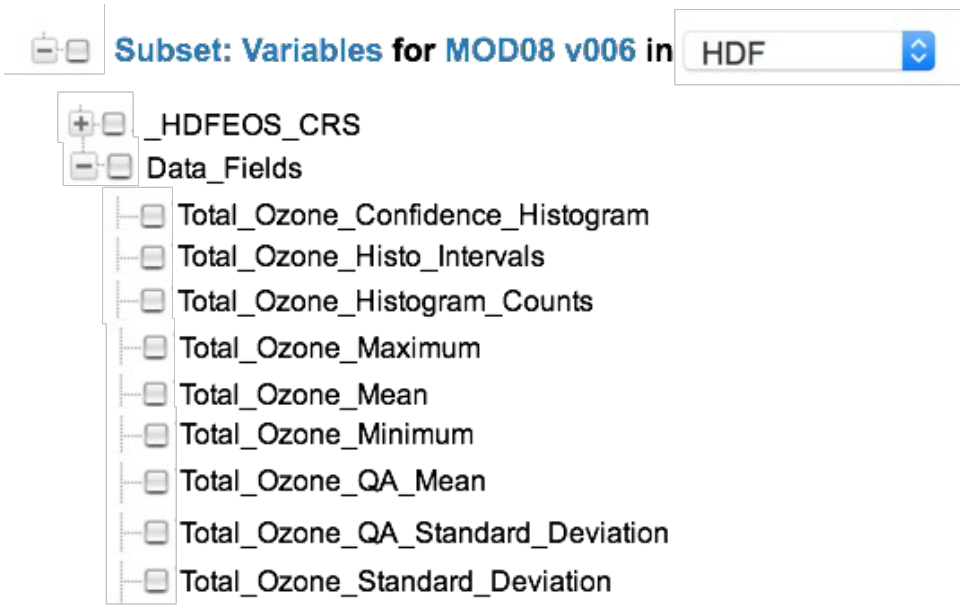

Figure E.17: User interface view for a user to choose a subset of variables for the MOD08 collection

See user experience activity diagram below.

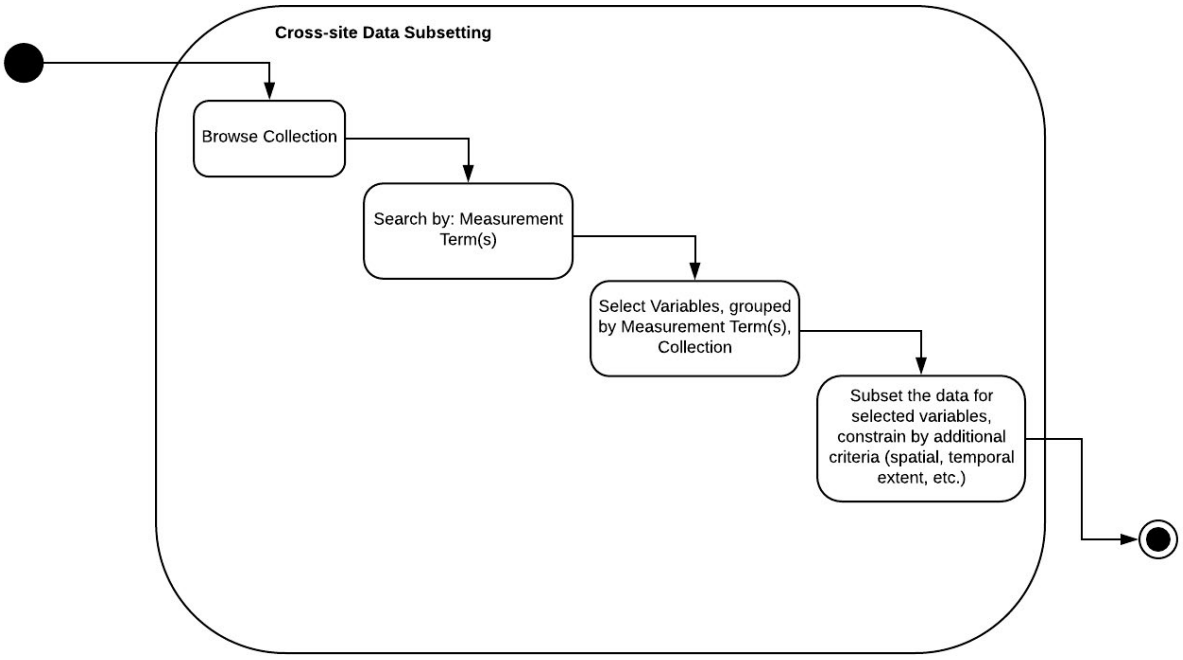

Figure E.18: Activity Diagram: Cross-site Data Subsetting

This following depicts the case where collections are searched for across DAACs. See the workflow sequence diagram below.

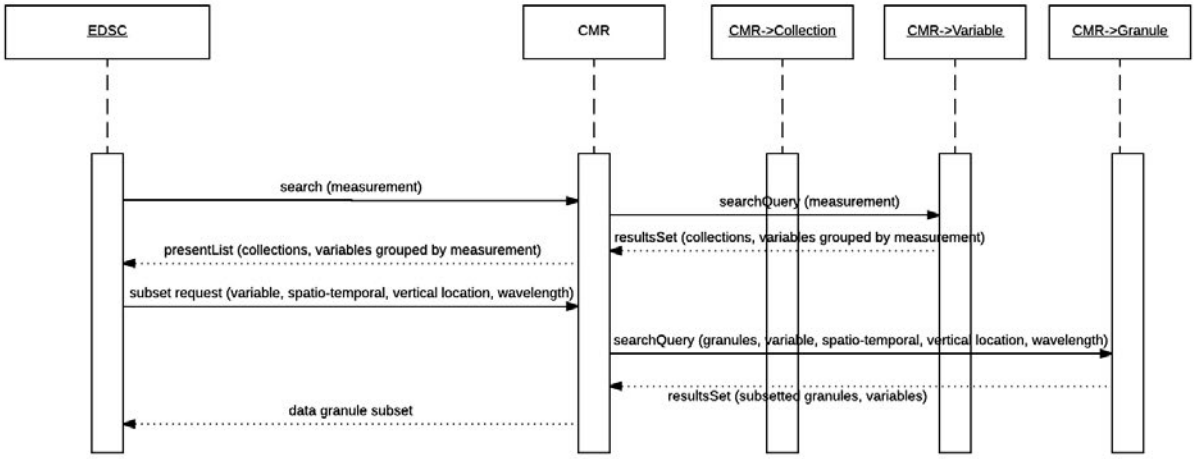

Figure E.19: Sequence Diagram: Cross-site Data Subsetting

E.2.2.6 Access Variables, including Ancillary Variables (extension of the Cross-Site Subsetting Use Case).

Scenario: The user starts with a list of science variables, e.g., {"sea surface height", "10.7 micron band", "NIR radiances", ... }, and wants to know which collections contain the specified variables (and may also want to know about data quality, instrument calibration, brightness temperature). These variables may be needed to fully understand the data.

Outcomes: Enables a user without any knowledge of the available variables to locate those variables of interest, and constrain these by collection metadata, e.g. calibration, spatial extent, temporal extent, spacecraft orbit location, etc. which contain a variable containing properties selected from the initial list of properties. Allows the subsequent discovery of associated variables, e.g. ancillary, calibration, geolocation, data quality variables, which are directly related to the science variable selected.

See the use case diagram below.

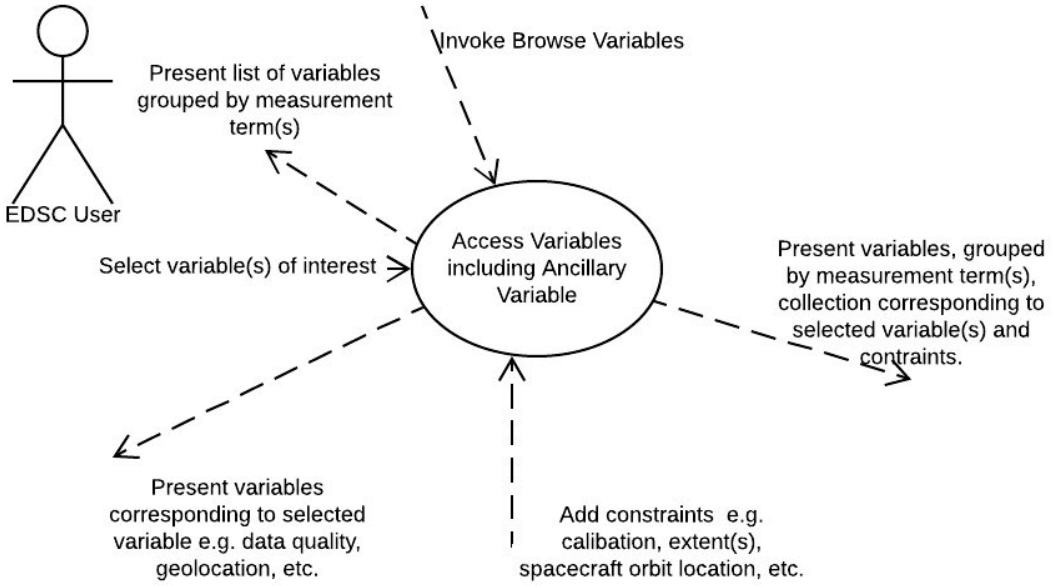

Figure E.20: Use Case: Access Variables including Ancillary Variable

See the user experience activity diagram below.

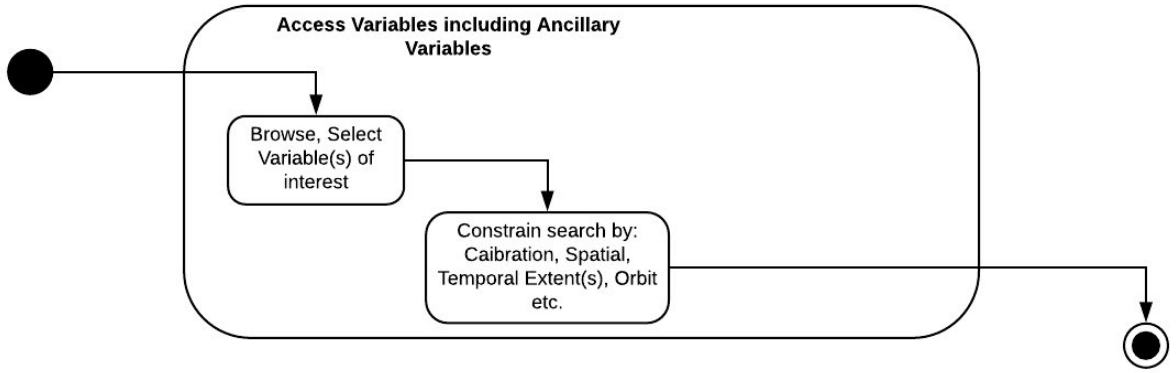

Figure E.21: Activity Diagram: Access Variables including Ancillary Variable

See the workflow sequence diagram below.

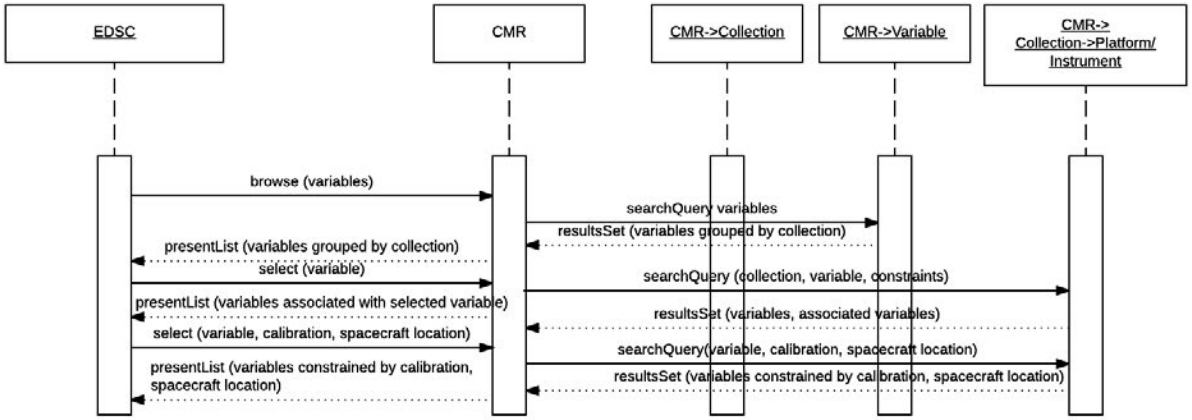

Figure E.22: Sequence Diagram: Access Variables, including Ancillary Variable

E.2.2.7 Integrating GIBS with Web-Based Clients

Scenario: As a user of a Global Imagery Browse Services (GIBS) client (e.g. EDSC, Worldview, GloVIS - https://glovis.usgs.gov/), I can view a pre-generated visualization for a specific granule or a daily composite of multiple granules with a specific collection and variable. Through CMR, I can locate the granules, the corresponding data variable from which the layer was generated, and any ancillary variables that need to go along with that variable (geolocation, data quality, etc.). Ideally, I can transform that

information into a set of subsetting request URLs that will fetch just those data variables from the appropriate granules.

Outcomes: Allow users of a GIBS client to fetch data subsets based on their layer selections, and any associated variables.

Note: It may be a prerequisite for this use case to have a way to invoke "show layers" for a collection, and selected variable(s), per the GIBS client. Also, providers may be providing a single image per granule to GIBS, e.g. if it's a L3/L4 product. But for L1/L2 products, this is not always the case. For one provider, LP DAAC, the only non-composited L1/L2 granule imagery available is for the AST_L1T product.

See the use case diagram below.

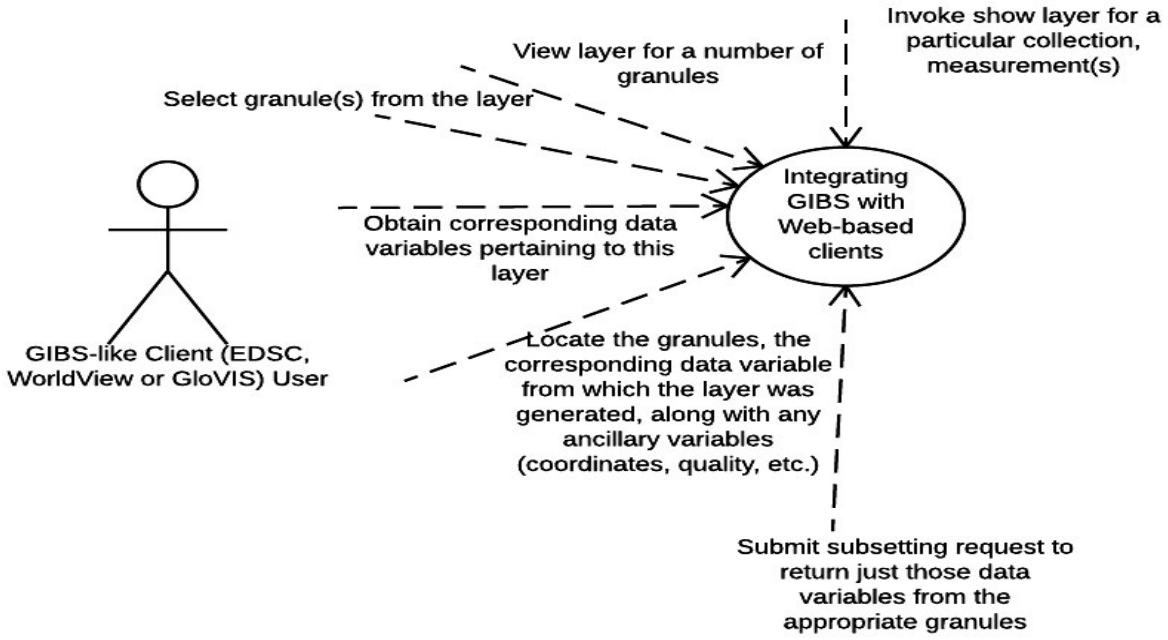

Figure E.23: Use Case: Integrating GIBS with web-based clients

See the user experience activity diagram below.

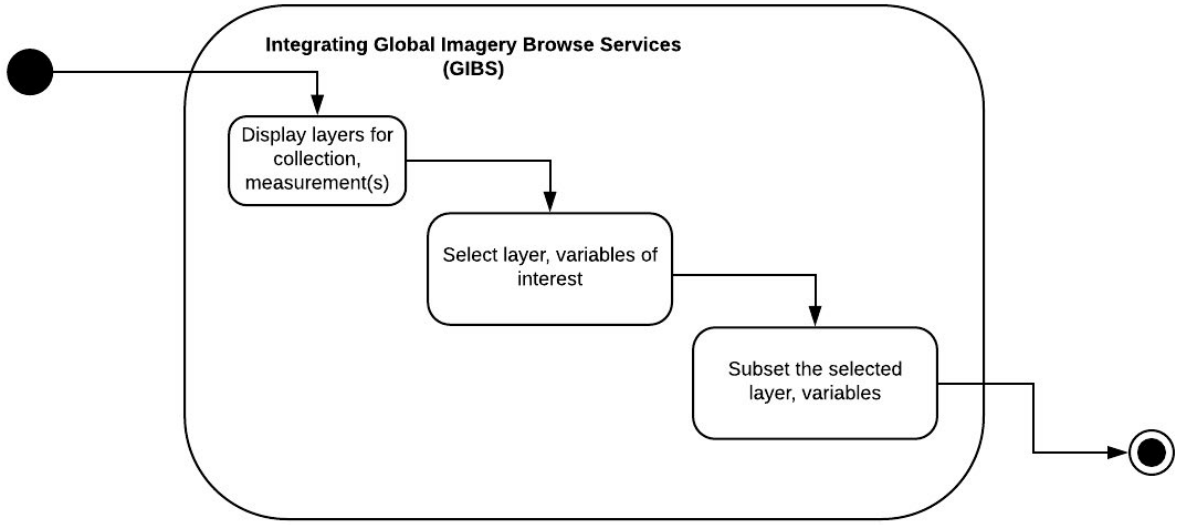

Figure E.24: Activity Diagram: Integrating GIBS with web-based clients

See the workflow sequence diagram below.

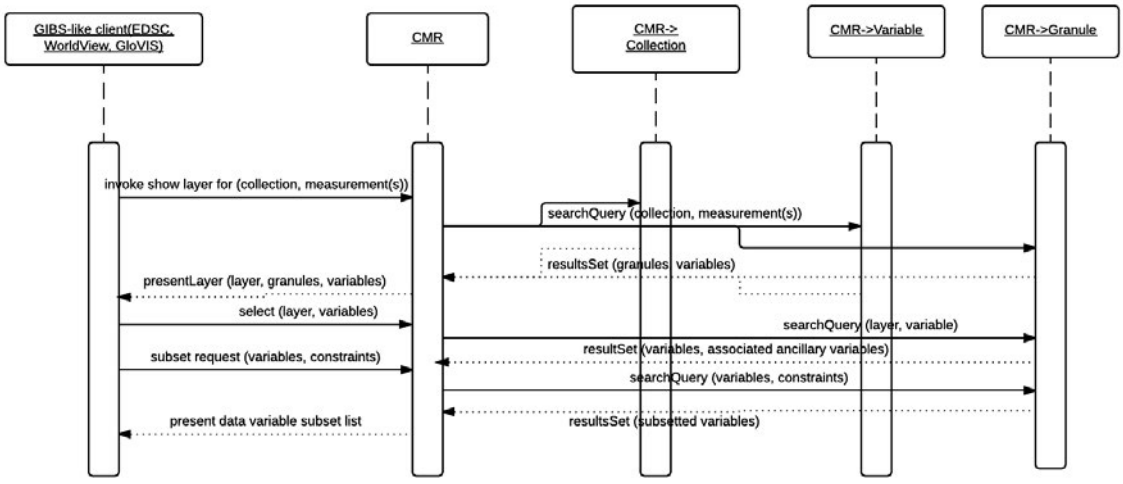

Figure E.25: Sequence Diagram: Integrating GIBS with web-based clients

E.2.2.8 Measurement Comparison of two in-situ measurements of Species X

Scenario: Measurement Comparison of two in-situ measurements of Species X, which is critical to understanding of tropospheric chemistry. In this scenario, the measurement techniques are not well established. Find

collections (and variables) which are tagged with measurement terms for Species X (e.g. X = Nitrous Oxide).

Outcomes: User obtains collections (and variables) for Principal Investigator (PI)-related data files containing Species X from instrument A and instrument B.

See the use case diagram below.

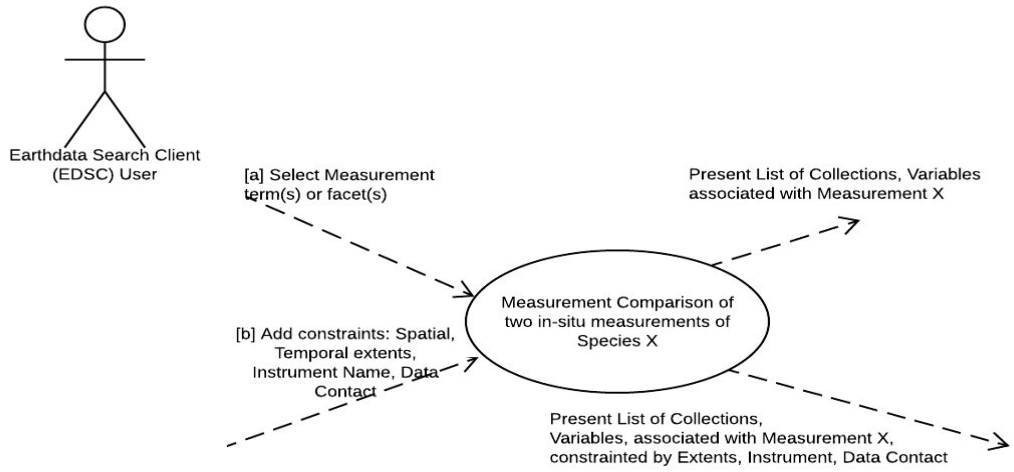

Figure E.26: Use Case, Measurement Comparison of two in-situ measurements of Species X

See the user experience activity diagram below

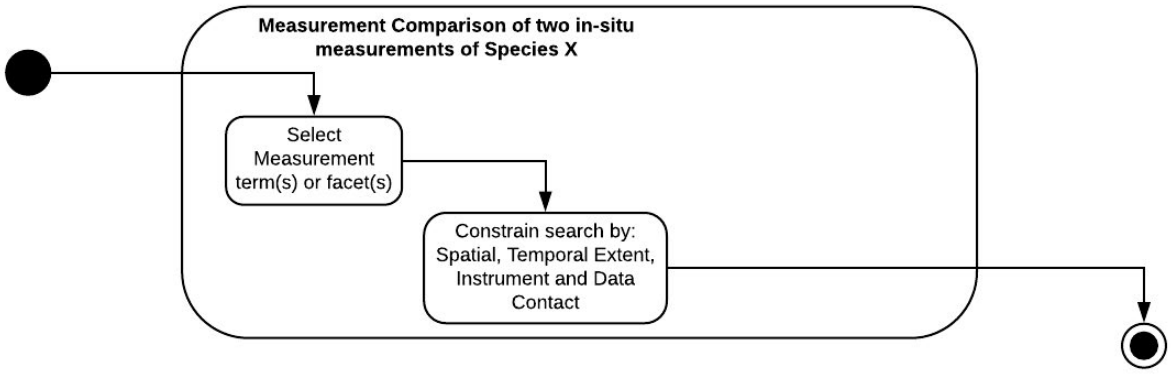

Figure E.27: Activity Diagram: Measurement Comparison of two in-situ measurements of Species X

See the workflow sequence diagram below.

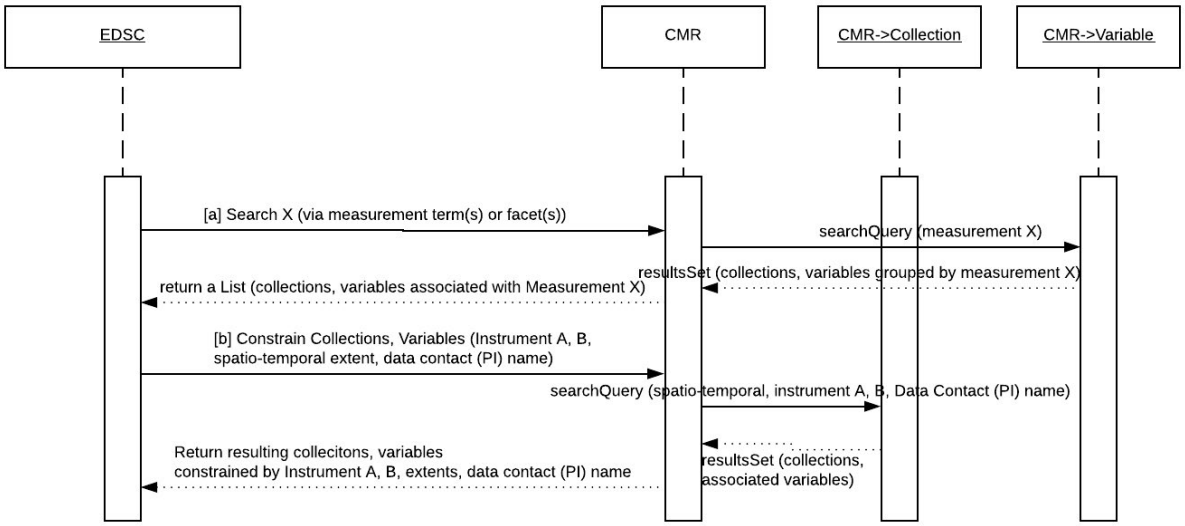

Figure E.28: Sequence Diagram: Measurement Comparison of two in-situ measurements of Species X

E.2.3 UMM-Var Metadata Model

As shown in Figure E.29, the UMM-Var Metadata Model asserts that a Variable metadata instance is related to one or more Collections, one or more Granules, one or more Variables (e.g. a Science Variable may have a related Quality or Ancillary Variable). The remaining classes: Characteristics, ScienceKeywords, Measurements, Sets, FillValues and Dimensions 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.

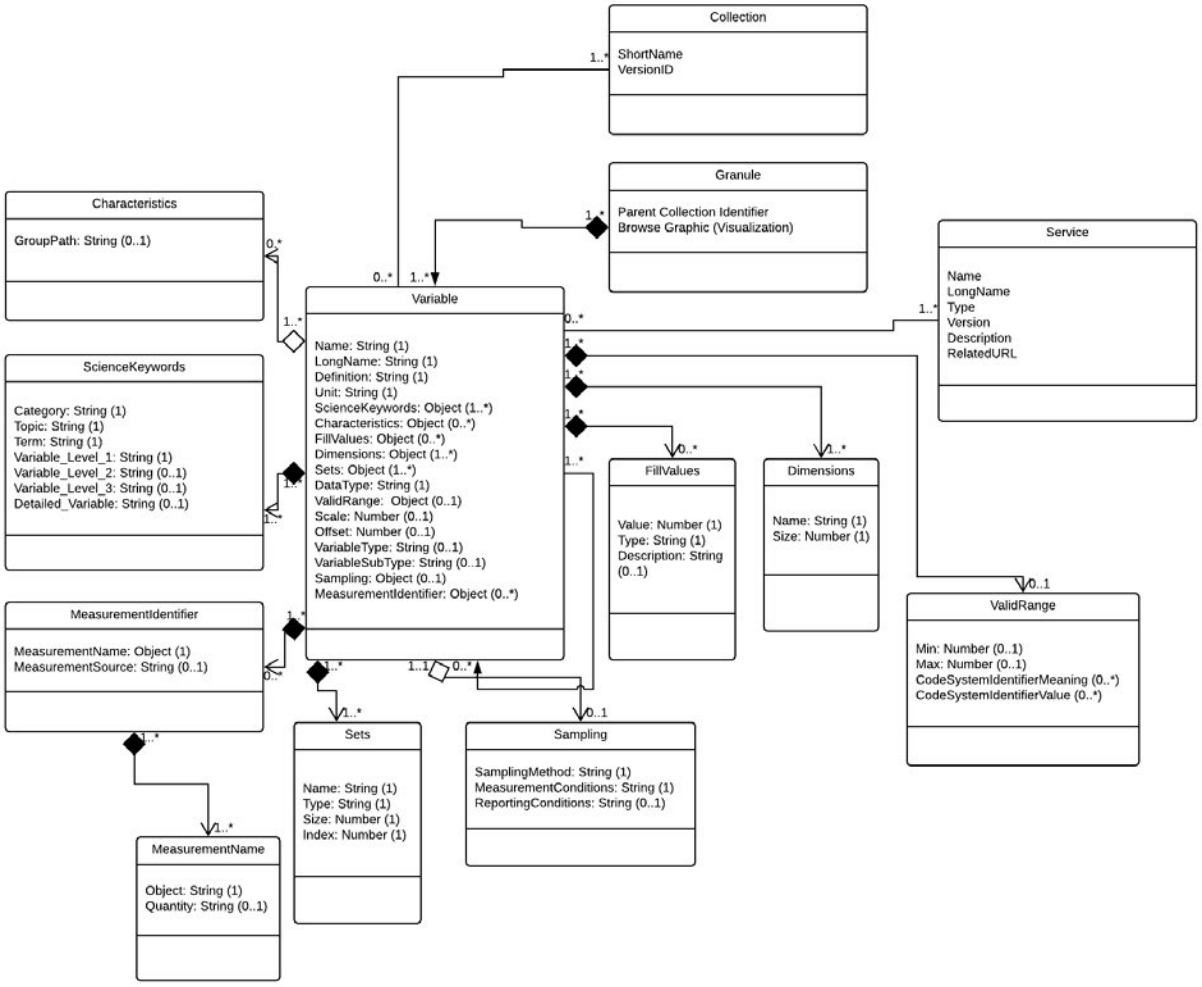

Figure E.29: UMM-Var Metadata Model

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

1. A Collection has zero or more Variables.
2. A Granule aggregates one or more Variables. Note the Variable class lifecycle is dependent on the Granule class instance lifecycle. Meaning that a Granule is defined by the Science Team first, and then each Variable which is aggregated by the Granule is defined. The metadata within the CMR simply reflects this.
3. A Variable may be related to zero or more Variables. For example, a Variable with VariableType: Science may have a related Variable(s) with VariableType: Quality and/or a Variable with VariableType: Ancillary.

4. The elements within the Characteristics section apply to a Variable. Not all Variables have all the elements contained within the Characteristics class.
5. The elements of the ScienceKeywords section also apply to a Variable. The ScienceKeywords may be sourced from GCMD Keywords. (See Appendix C).
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 D).
7. Information in the Characteristics section should be derived only from the Granule's data file. Granule selected should be from a collection that is associated with the variable.
8. A Variable record may be created / updated via the MMT GUI or XML file.
9. A Variable's record should answer all parts of the following question: What measurement type, collections, variables, granules, visualizations are associated with the Variable?
10. A mechanism within the UMM-S model enables the association of Services with specific Variables.

E.2.3.1 Variable

Elements
Variable [1..N]
Variable/Name [R]
Variable/LongName [R]
Variable/Definition [R]
Variable/Unit [R]
Variable/DataType [R]
Variable/Scale
Variable/Offset
Variable/VariableType

Description
The required fields: Name, LongName, DataType are derived from the Variable fields in the data set. The non-mandatory fields. Units, ValidRange, Scale and Offset are derived from the data set, if available. The Definition and the VariableType fields are set by the metadata curator via the MMT client.

Name [R]
Element Specification
Variable/Name (1)

Description
A variable short name given by the data provider.

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 below.

The VIIRS_SST_NPP L3C-GHRSST-SST Data Set structure is represented as:

Name	Long Name	Type
20160426130000-OSISAF-L3C_GHRSST-SSTsubskin-VIIRS_SST_NPP...	2016042613000...	Local File
adi_dtime_from_sst	time difference of ...	Geo2D
aerosol_dynamic_indicator	aerosol dynamic i...	Geo2D
dt_analysis	deviation from SST...	Geo2D
l2p_flags	L2P flags	Geo2D
lat	latitude	2D
lon	longitude	2D
or_latitude	original latitude of...	Geo2D
or_longitude	original longitude ...	Geo2D
quality_level	quality level of SST...	Geo2D
satellite_zenith_angle	satellite zenith angle	Geo2D
sea_ice_fraction	sea ice fraction	Geo2D
sea_surface_temperature	sea surface subski...	Geo2D
solar_zenith_angle	solar zenith angle	Geo2D
sources_of_adi	sources of aerosol...	Geo2D
sses_bias	SSES bias estimate	Geo2D
sses_standard_deviation	SSES standard dev...	Geo2D
sst_dtime	time difference fro...	Geo2D
wind_speed	10m wind speed	Geo2D

Figure E.30: 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 as the following figure. Note the dimensionality of the variable is: time=1, nj=3072 and ni=4096.

Variable "sea_surface_temperature"

```
short sea_surface_temperature(time=1, nj=3072, ni=4096);
  :_FillValue = -32768S; // short
  :long_name = "sea surface subskin temperature";
  :standard_name = "sea_surface_subskin_temperature";
  :units = "kelvin";
  :add_offset = 273.15; // double
  :scale_factor = 0.01; // double
  :valid_min = -300S; // short
  :valid_max = 4500S; // short
  :coordinates = "lon lat";
  :grid_mapping = "polar_stereographic_proj";
  :depth = "1 millimeter";
  :source = "VIIRS";
  :comment = "Temperature of the subskin of the ocean";
  :_ChunkSizes = 1, 1536, 2048; // int
```

Figure E.31 : The sea_surface_temperature variable structure

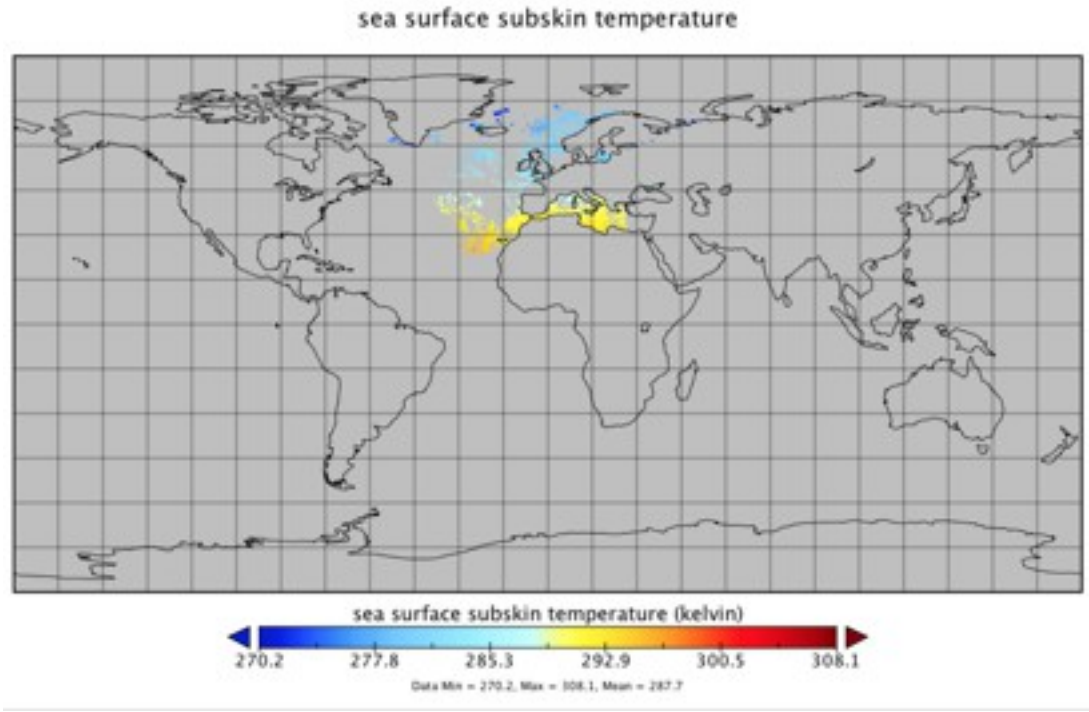

Figure E.32 : A sea_surface_temperature variable plot

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

Variable "quality_level"

```
byte quality_level(time=1, nj=3072, ni=4096);
  :_FillValue = -128B; // byte
  :long_name = "quality level of SST pixel";
  :valid_min = 0B; // byte
  :valid_max = 5B; // byte
  :flag_meanings = "no_data bad_data worst_quality low_quality acceptable_quality best_quality";
  :flag_values = 0B, 1B, 2B, 3B, 4B, 5B; // byte
  :coordinates = "lon lat";
  :grid_mapping = "polar_stereographic_proj";
  :comment = "These are the overall quality indicators and are used for all GHRSST SSTs";
  :_ChunkSizes = 1, 1536, 2048; // int
```

Figure E.33: The quality_level variable structure

Conversely, the LST variable contained within the MOD11A1 Data Set Structure is shown as:

Name	Long Name	Type
MOD11A1.A2009172.h16v05.006.2016014073638.hdf	MOD11A1.A2009172.h16v05.006....	Local File
ArchiveMetadata.0	ArchiveMetadata.0	—
CoreMetadata.0	CoreMetadata.0	—
MODIS_Grid_Daily_1km_LST	MODIS_Grid_Daily_1km_LST	—
_HDFEOS_CRS	_HDFEOS_CRS	—
Data_Fields	MODIS_Grid_Daily_1km_LST/Data_Fi...	—
Clear_day_cov	day clear-sky coverage	Geo2D
Clear_night_cov	night clear-sky coverage	Geo2D
Day_view_angl	View zenith angle of daytime Land-...	Geo2D
Day_view_time	Time of daytime Land-surface Tem...	Geo2D
Emis_31	Band 31 emissivity	Geo2D
Emis_32	Band 32 emissivity	Geo2D
LST_Day_1km	Daily daytime 1km grid Land-surfac...	Geo2D
LST_Night_1km	Daily nighttime 1km grid Land-surf...	Geo2D
Night_view_angl	View zenith angle of nighttime Land...	Geo2D
Night_view_time	Time of nighttime Land-surface Te...	Geo2D
Projection	Projection	—
QC_Day	Quality control for daytime LST and ...	Geo2D
QC_Night	Quality control for nighttime LST and...	Geo2D
XDim	x coordinate	1D
YDim	y coordinate	1D
StructMetadata.0	StructMetadata.0	—

Figure E.34: he LST_Day_1km variable highlighted within the MOD11A1 Data Set

The LST_Day_1km variable structure is represented as shown below. Note the dimensionality of the variable is: YDim=1200 and XDim=1200.

Variable "LST_Day_1km"

```
short LST_Day_1km(YDim=1200, XDim=1200);
  :_Unsigned = "true";
  :long_name = "Daily daytime 1km grid Land-surface Temperature";
  :units = "K";
  :Number_Type = "uint16";
  :valid_range = 7500S, -1S; // short
  :_FillValue = 0S; // short
  :LST = "LST data * scale_factor";
  :scale_factor = 0.02; // double
  :scale_factor_err = 0.0; // double
  :add_offset_err = 0.0; // double
  :calibrated_nt = 5; // int
```

Figure E.35 : The LST_Day_1KM variable structure

LST_Day_1km plot

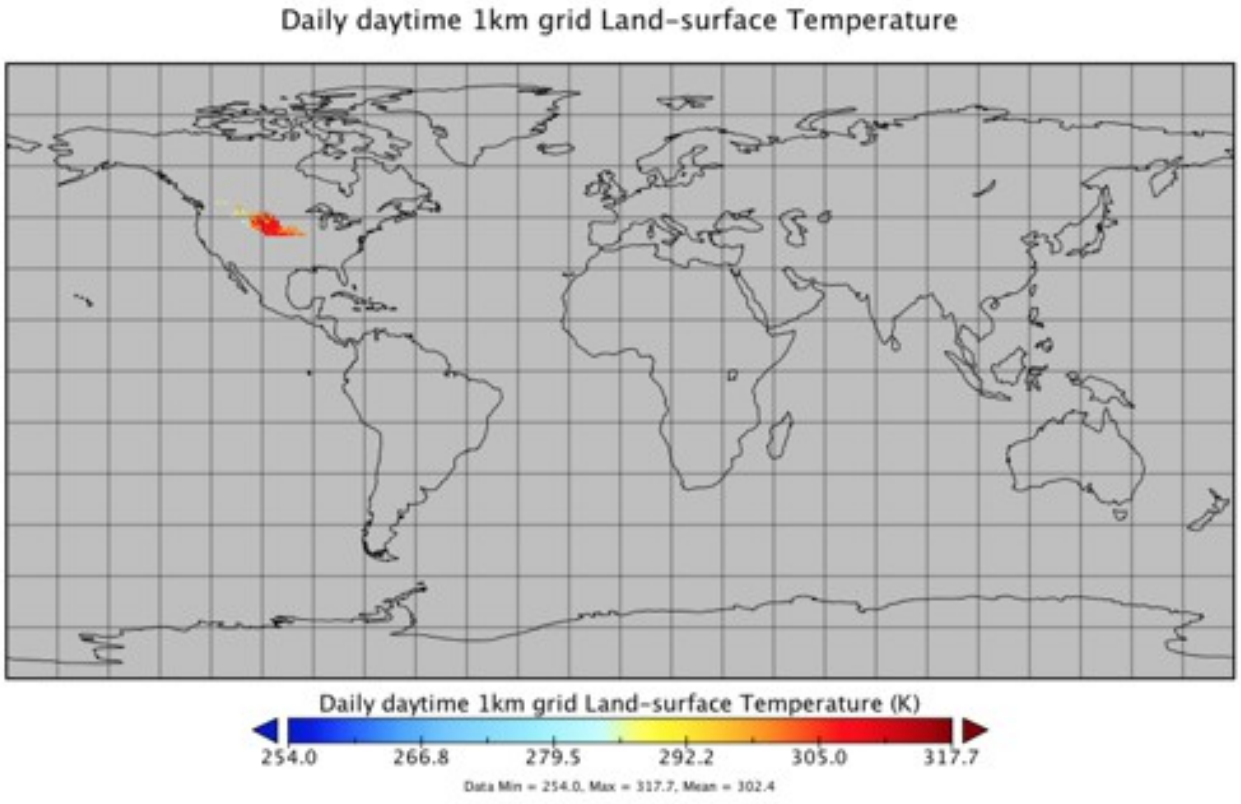

Figure E.36 : A LST_Day_1km plot

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

Variable "QC_Day"

```
byte QC_Day(YDim=1200, XDim=1200);
  :_Unsigned = "true";
  :long_name = "Quality control for daytime LST and emissivity";
  :Number_Type = "uint8";
  :valid_range = 0B, -1B; // byte
```

Figure E.37 : The QC_Day variable structure

CER_BDS_Aqua-FM3_Edition1 Data Set structure

Name	Long Name	Type
CER_BDS_Aqua-FM3_Edition1-...	CER_BDS_Aqua-FM3_Edition1-CV...	Local File
Azimuth_Position_Count	Azimuth Position Count	2D
CERES_metadata	CERES_metadata	—
CERES_Relative_Azimuth_at...	CERES Relative Azimuth at Surface	2D
CERES_Relative_Azimuth_at...	CERES Relative Azimuth at TOA - ...	2D
CERES_Solar_Zenith_at_Surf...	CERES Solar Zenith at Surface	2D
CERES_Solar_Zenith_at_TOA...	CERES Solar Zenith at TOA - Geoc...	2D
CERES_SW_Filtered_Radianc...	CERES SW Filtered Radiance, Upw...	2D
CERES_TOT_Filtered_Radian...	CERES TOT Filtered Radiance, Up...	2D
CERES_Viewing_Zenith_at_S...	CERES Viewing Zenith at Surface	2D
CERES_Viewing_Zenith_at_T...	CERES Viewing Zenith at TOA - G...	2D
CERES_WN_Filtered_Radianc...	CERES WN Filtered Radiance, Up...	2D
Clock_Angle_Rates	Clock Angle Rates	2D
Clock_Angles	Clock Angles	2D
Colatitude_of_CERES_FOV_a...	Colatitudes of CERES FOV at Surface	2D
Colatitude_of_CERES_FOV_a...	Colatitudes of CERES FOV at TOA	2D
Cone_Angle_Rates	Cone Angle Rates	2D
Cone_Angles	Cone Angles	2D
Converted_Azimuth_Angles	Converted Azimuth Angles	2D
Converted_Elevation_Angles	Converted Elevation Angles	2D
Converted_Instrument_Stat...	Converted_Instrument_Status_Data	—
Converted_Temperatures	Converted_Temperatures	—
Converted_Voltages_and_T...	Converted_Voltages_and_Torques	—
Count_Conversion_Constants	Count_Conversion_Constants	—
Count_Conversion_SW_Sam...	Count Conversion SW Sample Offs...	2D
Count_Conversion_TOT_Sa...	Count Conversion TOT Sample Of...	2D
Count_Conversion_WN_Sa...	Count Conversion WN Sample Off...	2D

Figure E.38: 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 as shown below. Note the dimensionality of the variable is: Records=13091 and Samples=660.

Variable "CERES_SW_Filtered_Radiances_Upwards"

```
float CERES_SW_Filtered_Radiances_Upwards(Records=13091, Samples=660);
  :long_name = "CERES SW Filtered Radiance, Upwards";
  :units = "Watts per square meter per steradian";
  :format = "32-BitFloat";
  :coordsys = "not used";
  :valid_range = -10.0f, 510.0f; // float
  :_FillValue = 3.4028235E38f; // float
```

Figure E.39: The CERES_SW_Filtered_Radiances_Upwards variable structure

CERES_Solar_Zenith_at_Surface variable structure

Variable "CERES_Solar_Zenith_at_Surface"

```
float CERES_Solar_Zenith_at_Surface(Records=13091, Samples=660);
  :long_name = "CERES Solar Zenith at Surface";
  :units = "deg";
  :format = "32-BitFloat";
  :coordsys = "not used";
  :valid_range = 0.0f, 180.0f; // float
  :_FillValue = 3.4028235E38f; // float
```

Figure E.40: CERES_SYN_1km Data Set structure

Figure E.41: The SW_TOA_Clear-Sky variable highlighted within the CERES_SYN_1km Data Set structure

SW_TOA_Clear-Sky variable structure. 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.

```
Variable "SW_TOA_Clear-Sky"
float SW_TOA_Clear-Sky(Mean_&_Stdev=2, Synoptic_Hours_(1,4,7,10,13,16,19,22)=8, 1.0_deg._regional_colat._zones=180, 1.0_deg._regional_long._zones=360);
  :_FillValue = 3.4028235E38f; // float
  :long_name = "1.0 Degree Regional MonthObserved TOA Fluxes";
  :units = "Watt per square meter";
```

Figure E.42 : The SW_TOA_Clear-Sky variable structure

SW_TOA_Clear-Sky variable plot

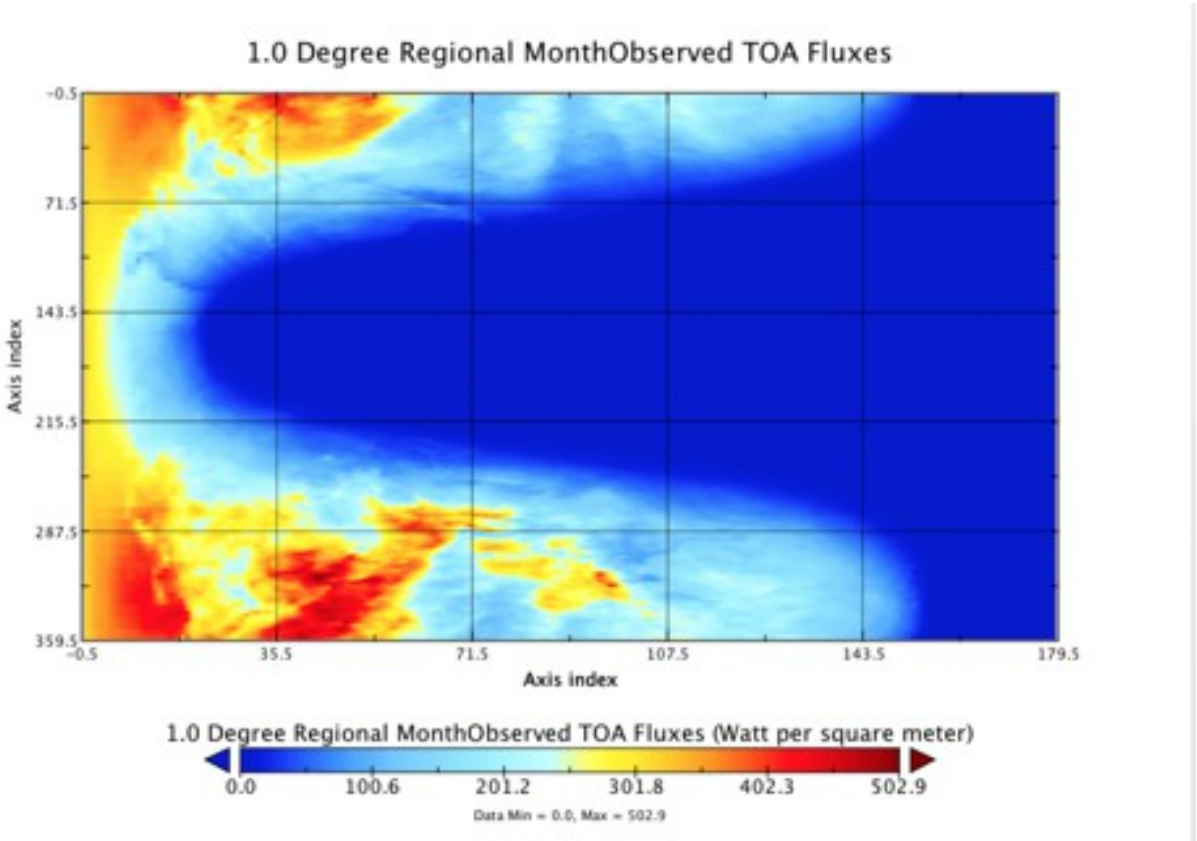

Figure E.43 : A SW_TOA_Clear-Sky variable

AIRS.2012.02.09.L3.CO2Std008 data set structure

Name	Long Name	Type
AIRS.2012.02.09.L3.CO2Std008.v5.9.14.0.X12089141044.hdf	AIRS.2012.02.09....	Local File
CO2	CO2	—
Data_Fields	CO2/Data_Fields	—
Latitude	Latitude	2D
Longitude	Longitude	2D
mole_fraction_of_carbon_dioxide_in_free_troposphere	mole_fraction_of_c...	Geo2D
mole_fraction_of_carbon_dioxide_in_free_troposphere_count	mole_fraction_of_c...	Geo2D
mole_fraction_of_carbon_dioxide_in_free_troposphere_sdev	mole_fraction_of_c...	Geo2D
Grid_Attributes	CO2/Grid_Attributes	—

Figure E.44 : 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. Note the dimensionality of the variable is: LatDim=91, LonDim=144.

Variable "mole_fraction_of_carbon_dioxide_in_free_troposphere"

```
float mole_fraction_of_carbon_dioxide_in_free_troposphere(LatDim=91, LonDim=144);
  :_FillValue = -9999.0f; // float
```

Figure E.45 : The mole_fraction_of_carbon_dioxide_in_free_troposphere variable structure

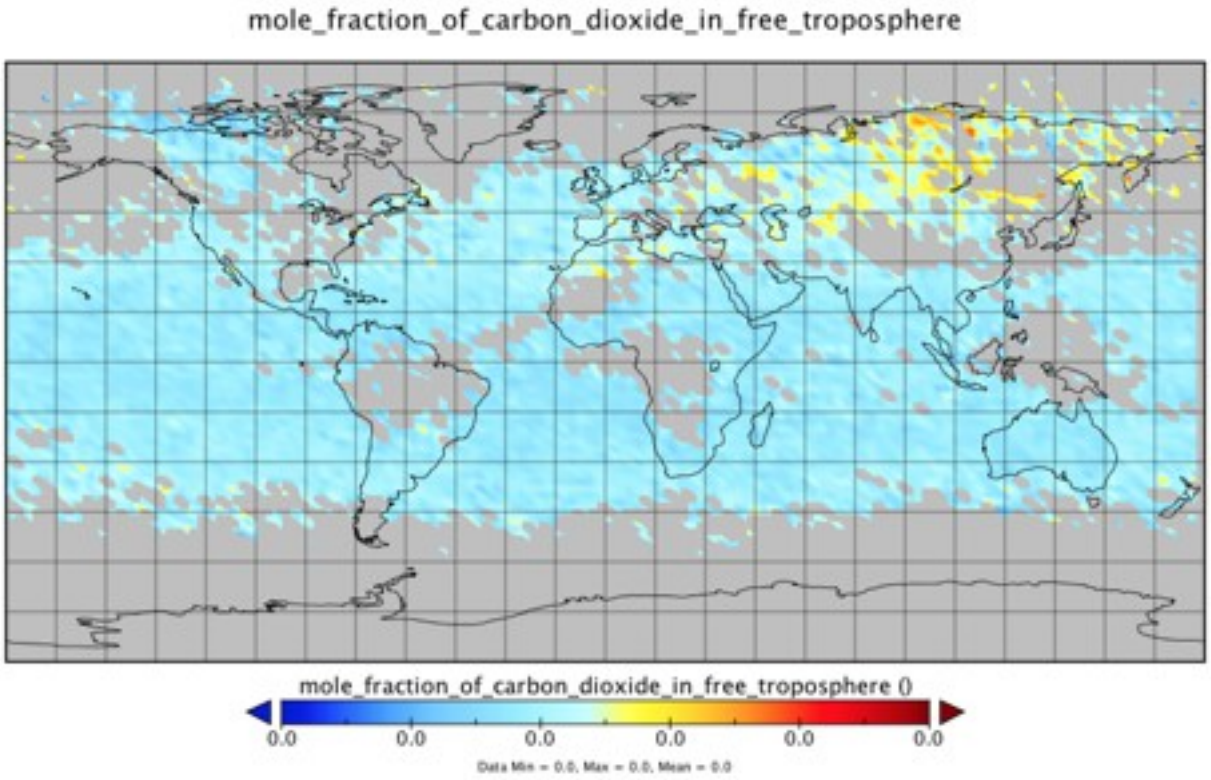

Figure E.46 : A mole_fraction_of_carbon_dioxide_in_free_troposphere plot

Sample Values (given in **bold**) below:

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_at_Surface (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)
Sur_Refl_b01 (surface reflectance band 1)
WDB_L3MCA10 (Aerosol Optical Depth 550nm (Land Only))

Tags

Required, Free Text Search

Sample Mappings
ISO 19115-1:
/
mdb:MD_Metadata/mdb:identificationInfo/mri:MD_DataIdentification/mri:citation/cit:CI_Citation/cit:identifier/mcc:MD_Identifier[mcc:description/gco:CharacterString[.='Name']]/mcc:code/gco:CharacterString

LongName [R]
Element Specification

Variable/LongName (1)

Description
The expanded or long name given by the data provider.

Sample values (given in **bold**) below:

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

Sample Mappings
ISO 19115-1:
/
mdb:MD_Metadata/mdb:identificationInfo/mri:MD_DataIdentification/mri:ci

tation/cit:CI_Citation/cit:identifier/mcc:MD_Identifier[mcc:description/gco:CharacterString[.='LongName']]/mcc:code/gco:CharacterString

Definition [R]
Element Specification

Variable/Definition (1)

Description

The definition 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 help a scientist to understand what is particular to this variable. See the Sampling 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, $\tau a = \beta \lambda \alpha$ where α is Angstrom exponent (β = aerosol optical thickness at 1 μm)".

Tags

Required

Sample Mappings

ISO 19115-1:

/
mdb:MD_Metadata/mdb:identificationInfo/mri:MD_DataIdentification/mri:abstract/gco:CharacterString

Unit [R]
Element Specification

Variable/Unit (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 E.2: 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 (e.g., 1 Jan 1970, TAI93, start of mission, etc.).	Time is in ISO- 8601 format. seconds since 1992-10-08T15:15:42.5-6:00 days since 1970-01-01T00:00:0

Tags

Required

Sample Mappings

ISO 19115-1:

/mdb:MD_Metadata/mdb:contentInfo/mrc:MD_CoverageDescription/mrc:attributeGroup/mrc:MD_AttributeGroup/mrc:attribute/mrc:MD_SampleDimension/mrc:units/gml:UnitDefinition/gco:uomName

with

/mdb:MD_Metadata/mdb:contentInfo/mrc:MD_CoverageDescription/mrc:attributeGroup/mrc:MD_AttributeGroup/mrc:attribute/mrc:MD_SampleDimension/mrc:units/gml:UnitDefinition/gco:uomSymbol

with

/mdb:MD_Metadata/mdb:contentInfo/mrc:MD_CoverageDescription/mrc:attributeGroup/mrc:MD_AttributeGroup/mrc:attribute/mrc:MD_SampleDimension/mrc:units/gco:UomAngle/gco:measureType/gco:MeasureType/gco:CharacterString

DataType [R]

Element Specification

Variable/DataType (1) <"byte", "string", "float", "float32", "float64", "double", "uchar", "uchar8", "short", "long", "int8", "int16", "int32", "int64", "uint8", "uint16", "unit32", "uint64", "ubyte", "ufloat", "ufloat32", "ufloat64", "OTHER">

Description

Specifies the basic computer science data type of a variable. These types can be either short, long, character, binary, etc.

Table E.3: 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 E.4: 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
Required

Sample Mappings
ISO 19115-1:
/
mdb:MD_Metadata/mdb:contentInfo/mrc:MD_CoverageDescription/mrc:attributeGroup/mrc:MD_AttributeGroup/mrc:attribute/mrc:MD_SampleDimension/mrc:sequenceIdentifier/gco:MemberName/gco:attributeType/gco:TypeName/gco:aName/gco:CharacterString

Scale
Element Specification
Variable/Scale (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. May be used together with Offset. 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)

Ref:
https://support.hdfgroup.org/release4/doc/UG_PDF.pdf (Section 3.10.6 Calibration Attributes)

Also,
https://cdn.earthdata.nasa.gov/conduit/upload/495/netcdf_UG_3.6.3.pdf (See Appendix B Attribute Conventions)

Note: the additive offset formula is the standard one, with the subtractive being non-standard, and rarely used. Exceptions include science variables from: 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 value: 0.00100000004749745

Tags
Recommended

Sample Mappings
ISO 19115-1:
/
mdb:MD_Metadata/mdb:contentInfo/mrc:MD_CoverageDescription/mrc:attributeGroup/mrc:MD_AttributeGroup/mrc:attribute/mrc:MD_SampleDimension/mrc:scaleFactor/gco:Real

Offset
Element Specification
Variable/Offset (0..1)

Description
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. May be used together with Scale. 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)

Ref:
https://support.hdfgroup.org/release4/doc/UG_PDF.pdf (Section 3.10.6 Calibration Attributes)

Also,
https://cdn.earthdata.nasa.gov/conduit/upload/495/netcdf_UG_3.6.3.pdf (See Appendix B Attribute Conventions)

Note: the additive offset formula is the standard one, with the subtractive being non-standard, and rarely used. Exceptions include science variables from: 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 value: 0.0

Tags
Recommended

Sample Mappings
ISO 19115-1:
/
mdb:MD_Metadata/mdb:contentInfo/mrc:MD_CoverageDescription/mrc:attributeGroup/mrc:MD_AttributeGroup/mrc:attribute/mrc:MD_SampleDimension/mrc:offset/gco:Real

VariableType
Element Specification
Variable/VariableType (0..1) <"SCIENCE_VARIABLE", "QUALITY_VARIABLE", "ANCILLARY_VARIABLE", "OTHER">

Description
Specifies the basic type of a variable. These types can be either: "SCIENCE_VARIABLE", "QUALITY_VARIABLE", "ANCILLARY_VARIABLE", "OTHER". This field is selected by the Metadata Curator via a suitable GUI, i.e. MMT, or an alternate metadata curation tool.

Sample value: "SCIENCE_VARIABLE".

Tags

Recommended

Sample Mappings

ISO 19115-1:
/
mdb:MD_Metadata/mdb:contentInfo/mrc:MI_CoverageDescription/mrc:attributeGroup/mrc:MD_AttributeGroup/mrc:contentType/mrc:MD_CoverageContentTypeCode and
/
mdb:MD_Metadata/mdb:contentInfo/mrc:MI_CoverageDescription/mrc:attributeGroup/mrc:MD_AttributeGroup/mrc:contentType/mrc:MD_CoverageContentTypeCode/@codeListValue

VariableSubType
Element Specification
Variable/VariableSubType (0..1) <"SCIENCE_SCALAR", "SCIENCE_VECTOR", "SCIENCE_ARRAY", "SCIENCE_EVENTFLAG", "OTHER">

Description
Specifies the sub type of a variable. This field is selected by the Metadata Curator via a suitable GUI, i.e. MMT, or an alternate metadata curation tool. There are different types of science variables and this information is variable specific and important for data use.

Sample values:
"SCIENCE_SCALAR"
"SCIENCE_VECTOR"
"SCIENCE_ARRAY"
"SCIENCE_EVENTFLAG"
"OTHER"

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.

Tags
Recommended

Sample Mappings
ISO 19115-1:
TBD

E.2.3.2 Characteristics

Elements

Characteristics (0..1)

Characteristics/GroupPath

Description

The elements of this section apply to a Variable.

Tags

Recommended

E.2.3.2.1 GroupPath

Element Specification

Characteristics/GroupPath (0..1)

Description

The full path to the variable within the Granule structure. The main purpose of this field is to capture the full path of the variable from within the granule file structure. Sets of variables which are nested a levels below the "/' root level can be located correctly.

In the example shown here, the set named "/Data_Fields" is nested in a path called '/MODIS_Grid_Daily_1km_LST". This important structural information is not lost once the Variable records have been ingested into the CMR.

Sample Value: '/MODIS_Grid_Daily_1km_LST/Data_Fields'

Tags

Recommended

Sample Mappings

ISO 19115-1:

/

mdb:MD_Metadata/mdb:contentInfo/mrc:MD_CoverageDescription/mrc:attributeGroup/mrc:MD_AttributeGroup/mrc:attribute/mrc:MD_SampleDimension/mrc:otherProperty:Structure/gco:CharacterString

E.2.3.3 ScienceKeywords

Elements

ScienceKeywords (0..N)

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 selected by a Metadata Curator via a suitable GUI or directly in the metadata file. Science Keywords are derived from a controlled source; in this case the GCMD Keywords. Science Keywords are provided to enable better searches by the use of human-readable measurement terms. Note that GCMD Keywords have a more complex structure than the Measurement class.

ScienceKeywords are hierarchical with the higher level keywords, e.g. Category, Topic, Term required and the lower level keywords, e.g. 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 field "DetailedVariable" for a collection which possesses Methane variables, e.g. 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.

Tags
Recommended

E.2.3.3.1 ScienceKeywords
Element Specification

ScienceKeywords (0..N)

ScienceKeywords/Category [R]

ScienceKeywords/Topic [R]

ScienceKeywords/Term [R]

ScienceKeywords/Variable_Level1

ScienceKeywords/Variable_Level2

ScienceKeywords/Variable_Level3

ScienceKeywords/Detailed_Variable

Description
These elements describe controlled science keywords describing the measurements/variables. The controlled vocabulary for Science Keywords is maintained in the Keyword Management System (KMS). These will be sourced from GCMD Keywords. See Appendix C.

Sample Values:
"Category": "EARTH SCIENCE", "Topic": "ATMOSPHERE", "Term": "ATMOSPHERIC CHEMISTRY", "Variable_Level1": "NITROGEN COMPOUNDS", "Variable_Level2": "Peroxyacyl Nitrate".

Tags
Recommended, Controlled Vocabulary

Sample Mappings
ISO 19115-1:
/
mdb:MD_Metadata/mdb:identificationInfo/mri:MD_DataIdentification/mri:descriptiveKeywords/mri:MD_Keywords[mri:type/mri:MD_KeywordTypeCode/@codeListValue='Keyword']/mri:keyword/gco:CharacterString

E.2.3.4 FillValues

Element Specification

FillValues [0..N]

FillValues/Value [R]

FillValues/Type [R]

FillValues/Description

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.

Tags
Recommended

E.2.3.4.1 Value [R]
Element Specification
FillValues/Value (1)

Description
The fill value of the variable in the data file.

Sample values:
-1
-9999

Tags
Required

Sample Mappings
ISO 19115-1:
/
mdb:MD_Metadata/mdb:identificationInfo/mri:MD_DataIdentification/mri:citation/cit:CI_Citation/cit:identifier/mcc:MD_Identifier[mcc:description/gco:CharacterString[.='FillValue']]/mcc:code/gco:CharacterString = Value:

E.2.3.4.2 Type [R]
Element Specification

FillValue/Type (1) <"SCIENCE_FILLVALUE", "QUALITY_FILLVALUE", "ANCILLARY_FILLVALUE", "OTHER">

Description
Type of fill value of the variable in the data file.

Sample values:
SCIENCE_FILLVALUE

Tags
Required

Sample Mappings
ISO 19115-1:

/
mdb:MD_Metadata/mdb:identificationInfo/mri:MD_DataIdentification/mri:citation/cit:CI_Citation/cit:identifier/mcc:MD_Identifier[mcc:description/gco:CharacterString[.='FillValue']]/mcc:code/gco:CharacterString = Type:

E.2.3.4.3 Description

Element Specification

FillValue/Description (0..1)

Description

Description of the fill value of the variable in the data file.

Sample values:

Valid Science Fill Value

Tags

Recommended

Sample Mappings

ISO 19115-1:

/
mdb:MD_Metadata/mdb:identificationInfo/mri:MD_DataIdentification/mri:citation/cit:CI_Citation/cit:identifier/mcc:MD_Identifier[mcc:description/gco:CharacterString[.='FillValue']]/mcc:code/gco:CharacterString = Description:

E.2.3.5 Dimensions [R]

Element Specification

Dimensions [1..N]

Dimensions/Name [R]

Dimensions/Size [R]

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'. Variables are rarely one dimensional. Mode commonly, they are two or three dimensional.

Sample values:

For the sea_surface_temperature variable, the dimensionality is: time=1, nj=3072 and ni=4096.
For the quality_level variable variable, the dimensionality is: time=1, nj=3072 and ni=4096.
For the LST_Day_1KM variable, the dimensionality is: YDim=1200 and XDim=1200.
For the QC_Day variable, the dimensionality is: YDim=1200 and XDim=1200.
For the CERES_SW_Filtered_Radiances_Upwards variable, the dimensionality is: Records=13091 and Samples=660.
For the SW_TOA_Clear-Sky variable, 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.
For the mole_fraction_of_carbon_dioxide_in_free_troposphere variable, the dimensionality of the variable is: LatDim=91, LonDim=144

Tags
Recommended

E.2.3.5.1 Name [R]
Element Specification

Dimensions/Name (1..N)

Description
The name of the dimensions of the variable represented in the data field. For example, 'XDim.

Sample values:
YDim, lat

Tags
Required

Sample Mappings
ISO 19115-1:

/
mdb:MD_Metadata/mdb:contentInfo/mrc:MD_CoverageDescription/mrc:attributeGroup/mrc:MD_AttributeGroup/mrc:attribute/mrc:MD_SampleDimension/mrc:sequenceIdentifier/gco:MemberName/gco:attributeType/gco:TypeName/gco:aName/gco:CharacterString

E.2.3.5.2 Size [R]
Element Specification

Dimensions/Size (1..N)

Description

The size of the dimensions of the variable represented in the data field. For example, '1200'.

Sample values:
1200, 3600

Tags
Required

Sample Mappings
ISO 19115-1:
/
mdb:MD_Metadata/mdb:contentInfo/mrc:MD_CoverageDescription/mrc:attributeGroup/mrc:MD_AttributeGroup/mrc:attribute/mrc:MD_SampleDimension/mrc:sequenceIdentifier/gco:MemberName/gco:attributeType/gco:TypeName/gco:aName/gco:CharacterString

E.2.3.6 ValidRange

Element Specification
Variable/ValidRange (0..1)
Variable/ValidRange/Max (0..1)
Variable/ValidRange/Min (0..1)
Variable/ValidRange/CodeSystemIdentifierMeaning (0..N)
Variable/ValidRange/CodeSystemIdentifierValue (0..N)

Description
ValidRange specifies the minimum and maximum valid values of the variable represented in the data field.

Sample values:

Variable/ValidRange/Max: 5000
Variable/ValidRange/Min: -100

Optionally, if the valid range is not continuous, a code system can be defined. An example of such a code system is shown below:

Sample values:

Variable/ValidRangeCodeSystemIdentifierMeaning: <no_data, bad_data, worst_quality, low_quality, acceptable_quality, best_quality>
Variable/ValidRange/CodeSystemIdentifierValue: <0B, 1B, 2B, 3B, 4B, 5B>

Tags
Recommended

Sample Mappings
ISO 19115-1:

/
mdb:MD_Metadata/mdb:contentInfo/mrc:MD_CoverageDescription/mrc:attributeGroup/mrc:MD_AttributeGroup/mrc:attribute/mrc:MD_SampleDimension/mrc:minValue/gco:Real

and

/
mdb:MD_Metadata/mdb:contentInfo/mrc:MD_CoverageDescription/mrc:attributeGroup/mrc:MD_AttributeGroup/mrc:attribute/mrc:MD_SampleDimension/mrc:maxValue/gco:Real

E.2.3.6.1 Min
Element Specification

ValidRange/Min (0..1)

Description

The minimum value of the variable represented in the data field. For example, '0.0'.

Sample values:

0.0

Tags

Recommended

E.2.3.6.2 Max
Element Specification

ValidRange/Max (0..1)

Description

The maximum value of the variable represented in the data field. For example, '1.0'.

Sample values:

1.0

Tags

Recommended

E.2.3.6.3 CodeSystemIdentifierMeaning
Element Specification
ValidRange/CodeSystemIdentifierMeaning (0..N)

Description
This element can be used to specify a code system identifier meaning. For example, Open Shrubland corresponds to '7'.

This element can be used for variables which do not use a typical continuous range of valid values. If there is a discrete number system used for the data values, then there needs to be a code system identifier.

An example of each meaning is shown here:

For the GHRSST SST variable, it has a discrete code system used to identify the quality of each grid cell.

Meaning: <no_data, bad_data, worst_quality, low_quality, acceptable_quality, best_quality>
Value: <0B, 1B, 2B, 3B, 4B, 5B>

Examples are cloud masks, or land surface classification variables. (e.g. if the code system is IGBP, and the value is set to 7, then there is a way of determining that this is Open Shrubland). If the code system identifier for NODATA is -1, then, this field would be set to -1.

Sample values:

7

Tags

Recommended

E.2.3.6.4 CodeSystemIdentifierValue

Element Specification

Variable/ValidRange/CodeSystemIdentifierValue (0..N)

Description

ValidRange specifies valid values of the variable represented in the data field using the defined code system.

The code system identifier value is the textual or numerical value assigned to each meaning.

An example of each value is shown here:

For the GHRSST SST variable, it has a discrete code system used to identify the quality of each grid cell.

Meaning: <no_data, bad_data, worst_quality, low_quality, acceptable_quality, best_quality>
Value: <0B, 1B, 2B, 3B, 4B, 5B>

Tags

Recommended

E.2.3.7 Measurements [O]

Elements

Measurements (0..*)

Measurements/MeasurementName

Measurements/MeasurementName/Object [R]

Measurements/MeasurementName/Quantity [R]

Measurements/MeasurementSource

Description
Elements in this category are used for search purposes. Measurement names are added to the metadata by a Metadata Curator via the MMT, an alternate metadata curation tool, using bulk updates, or via a JSON metadata file. The measurement name is structured according to the form defined by Scott Peckham. This is: <<object, quantity>>

Therefore, the specification of the MeasurementName class will be:
Measurements/MeasurementName (0..1)
Measurements/MeasurementName/Object [R]
Measurements/MeasurementName/Quantity [R]

and the source of the names can be identified by the following field:

Measurements/MeasurementSource

Every standard name has an object part that describes a particular object and a quantity part that describes a particular attribute of the object that can be quantified. These names are sorted alphabetically and other sorting methods can be added later. When using the MeasurementSource field to identify measurement names source from CSDMS system, use: CSDMS. When sourcing measurement names from other sources, e.g. CF convention, use "CF".

Membership criteria are given by:
Object = "Aerosol"
Quantity = "Optical Depth"

More discussion on MeasurementName valid values is given in the MeasurementName object specification. In consultation with the GCMD team, it is recommended that Measurement's valid values should be enumerations in KMS and not keywords. It is also recommended that the Measurement's valid values will eventually be managed via the current ESO process but not until the valid values are matured.

Tags
Recommended

E.2.3.7.1 MeasurementName
Element Specification

Measurements/MeasurementName (0..1)

Measurements/MeasurementName/Object [R]

Measurements/MeasurementName/Quantity [R]

Description

The names of the measurement may be taken from a variety of sources. These include, but are not limited to, CSDMS Cross Domain Naming Conventions or CF Standard Name Convention, British Oceangraphic Data Centre (BODC).

According to the CSDMS Basic Rules, every standard name has an object part that describes a particular object and a quantity part that describes a particular attribute of that object that can be quantified with a number. These names are sorted alphabetically, but other sorting methods can be added later.

Names are of the form: <object>__<quantity>.
Names shall contain only lowercase letters and numbers along with the Standard Names separator characters (_, -, ~, __).
The Standard Names separators:
_: delimiter separate words of a name.
-: join multi-word objects, quantities, adjectives, etc.
~: join an adjective to a noun (the noun comes first following by or more adjectives).
__: separate an object from a quantity.
_of_: apply a math operation to the subsequent quantity.
Qualifiers that make an object or quantity more specific are added to the left of the base object or quantity (with increasing specificity).

CSDMS Standard Names may be further grouped by category: Atmosphere, Oceans, Radiation, Sea Ice, Soil, Snow, Topography.

This is defined more fully in the CSDMS WIki: https://csdms.colorado.edu/wiki/CSN_Basic_Rules

See Appendix C for more details.

Sample Values:

"land_subsurface_water_sat-zone_top", "CSDMS"

"land_surface", ;"CSDMS"

"land_surface_air", "CSDMS"

"land_surface_air_flow", "CSDMS"

"land_surface _air_heat-incoming-latent", "CSDMS"

"land_surface_air-incoming-sensible", "CSDMS"

"specific_humidity", "CSDMS"

"specific_humidity-standard_error", "CSDMS"

"specific_humidity-detection_minimum" "CSDMS"

This list may be supplemented further by standard names sourced from the CF Standard Names: http://cfconventions.org/Data/cf-standard-names/docs/guidelines.html

These can also be expressed in the form: <object>__<quantity> with care.

Species, e.g.vapor, sulfur, can be quantified by terms: e.g. at_cloud_top, at_convective_cloud_top, at_cloud_base, at_convective_cloud_base, at_freezing_level, at_ground_level

Fluxes, e.g. radiative_flux, can be quantified by terms e.g. at_top_of_atmosphere_model, at_sea_level, can be expressed as:

"radiative_flux-at_top_of_atmosphere_model", "CF"

"radiative_flux-at_sea_level, "CF"

Physical Quantities, e.g. temperature, pressure, humidity, entropy, which are commonly used in mathematics, science and engineering, can be expressed using CF convention.

e.g. electrical charge, or scientific symbol, q, and quantified by terms, e.g. error_limit, detection_limit, can be expressed as:

"q-error_limit", "CF"

"q-detection_limit"."CF"

Note: MeasurementNames' values will come from KMS which is a controlled list

Tags

Recommended, Controlled Vocabulary

Sample Mappings

ISO 19115-1:

/
mdb:MD_Metadata/mdb:identificationInfo/mri:MD_DataIdentification/mri:descriptiveKeywords/mri:MD_Keywords[mri:type/mri:MD_KeywordTypeCode/@codeListValue='MeasurementNames']/mri:keyword/gco:CharacterString

Object [R]
Element Specification

Measurements/MeasurementName/Object (1)

Description

The name of the object of measurement.

Sample Values:

land_subsurface_water-sat, land_surface, land_surface_air, land_surface_air_flow, land_surface_air_heat, specific_humidity, radiative_flux, q.

These represent the named object term in the <<object, quantity>> structure.

"**land_subsurface_water_sat** -zone_top",

"**land_surface** ",

"**land_surface_air** ",

"**land_surface_air_flow** ",

"**land_surface _air_heat** -incoming-latent",

"**land_surface_air** -incoming-sensible",

"**specific_humidity** ",

"**specific_humidity-** standard_error",

"**specific_humidity-** detection_minimum",

"**radiative_flux** -at_top_of_atmosphere_model"

"**radiative_flux** -at_sea_level"

"**q**-error_limit",

"**q**-detection_limit".

Tags

Required

Quantity [R]
Element Specification

Measurements/MeasurementName/Quantity (0..1)

Description

The name of the quantity of measurement.

Sample Values:

zone-top, incoming-latent, incoming-sensible, standard_error, detection_minimum, at_top_of_atmosphere_model, at_sea_level, error_limit, detection_limit

These represent the named quantity term in the <<object, quantity>> structure.

"land_subsurface_water_sat-**zone_top** ",

"land_surface_air_heat-**incoming-latent** ",

"land_surface_air-**incoming-sensible** ",

"specific-humidity-**standard_error** ",

"specific-humidity-**detection_minimum** ",

"radiative_flux-**at_top_of_atmosphere_model** "

"radiative_flux-**at_sea_level**"

"q-**error_limit** ",

"q-**detection_limit** ".

Tags

Required

E.2.3.7.2 MeasurementSource

Element Specification

Measurements/MeasurementSource (0..1) <"CSDMS", "CF", "BODC">

Description

The source of the measurement names include, but are not limited to: CSDMS Cross Domain Naming Conventions, CF Standard Name Convention, or British Oceanographic Data Centre. See Appendix C for more on the sources of measurement names and the recommended governance approach.

Sample Values:

CSDMS,

CF,

BODC

Tags

Recommended, Controlled Vocabulary

Sample Mappings

ISO 19115-1:

/
mdb:MD_Metadata/mdb:identificationInfo/mri:MD_DataIdentification/mri:descriptiveKeywords/mri:MD_Keywords[mri:type/mri:MD_KeywordTypeCode/@codeListValue='MeasurementNames']/mri:keyword/mri:thesaurusName/cit:CI_Citation/cit:citedResponsibleParty/cit:CI_Responsibility[cit:role/cit:CI_RoleCode@codeListValue='resourceProvider']/cit:party/cit:CI_AbstractCI_Party/cit:name/gco:CharacterString

E.2.3.8 Sets

Elements

Sets [1..N]

Sets/Name [R]

Sets/Type [R]

Sets/Size [R]

Sets/Index [R]

Description

Elements in this category are used to group variables into sets.

The set information of a variable. The variable is typically grouped within a set (essentially, an indexed array of variables). The set is defined by the name, type, size, and index. The Set class is flexible enough to be expanded to include compound variables (variable that groups related variable together to describe a phenomenon).

The current UMM-Var schema supports compound variables using the Set class.

Sets can be named and typed according to the grouping or phenomenon.

Example (variables common to the 'Data_Fields' group, within the MOD11A1 collection, so the set class would be populated in the following way for the variable named 'LST_Day_1km'):

"Sets": [
{

"Name": "Data_Fields ",

"Type": "MODIS 1km gridded",

"Size": 15 ,

"Index": 7

}
]

Example (variables common to the AIRX3STD gridded data field group', within the AIRX3STD collection, so the set class would be populated in the following way for the variable named 'EmisIR_A_ct'):

"Sets": [
{

"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 13th variable in a set of 867 variables. (14th if the numbering starts at 0).

For a phenomenon example, take the MOD08 v006 collection,

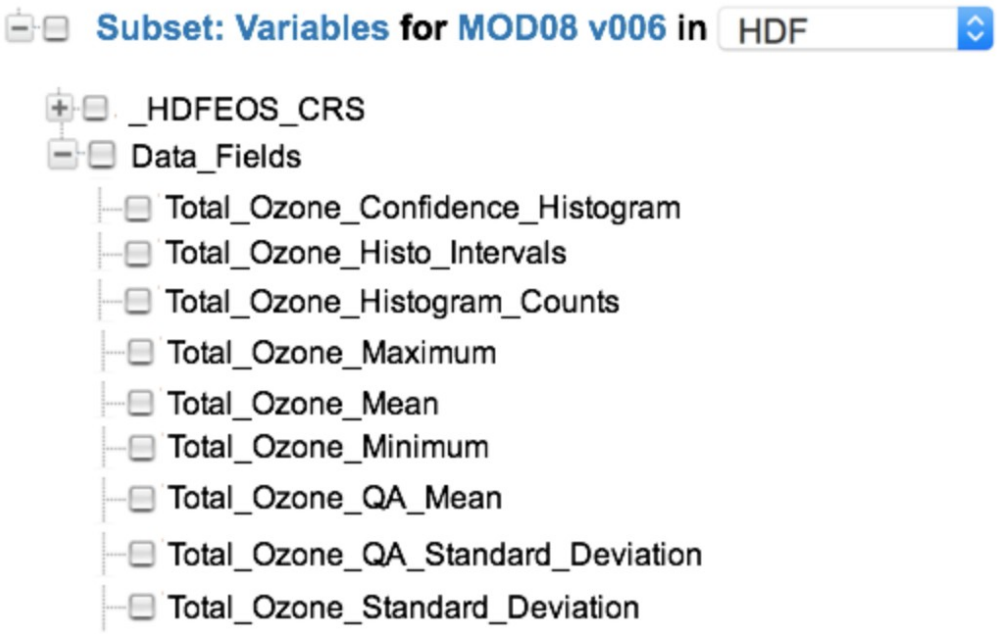

Figure E.47: Subset variable choices for the MOD08 collection

we can group the variables 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 example, the Variables for this SMAP_L3_SM_P data set are organized into two sets. The first set contains the variables representing the morning (AM) crossing and the second set contains variables representing the afternoon (PM) crossing.

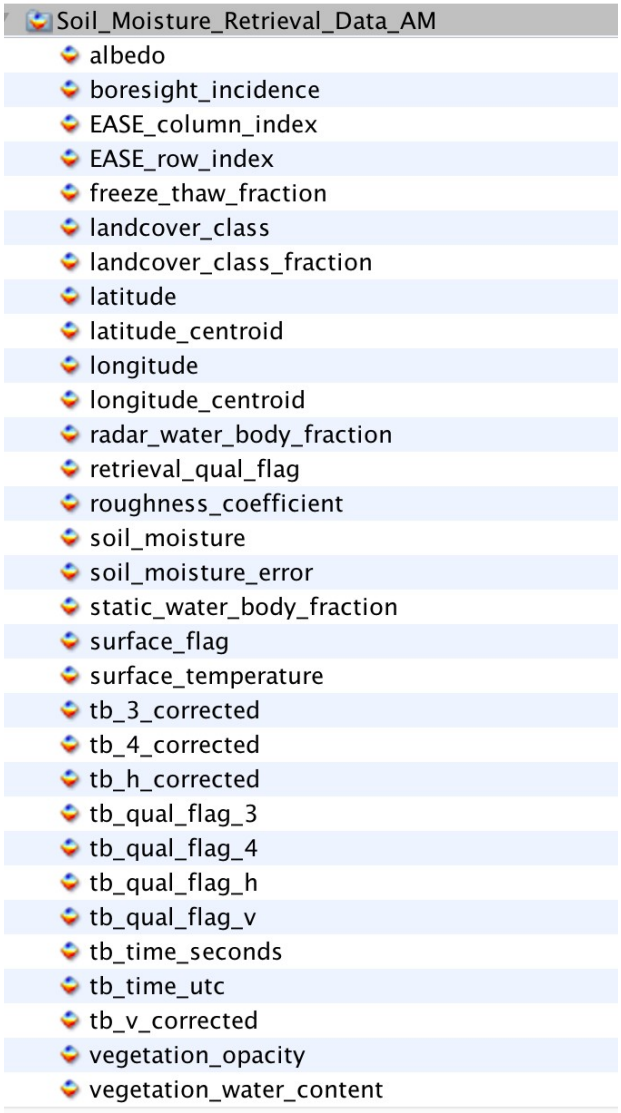

Figure E.48: SMAP_L3_SM_P variables representing the morning (AM) crossing

The second set represents the variables containing the afternoon (PM) crossing.

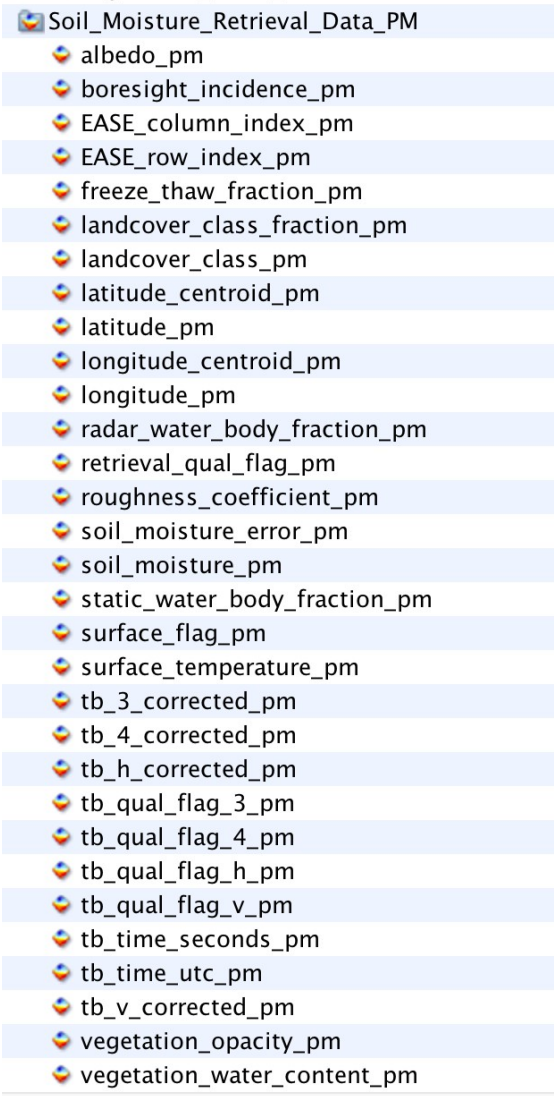

Figure E.49 : SMAP_L3_SM_P variables representing the afternoon (PM) crossing

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

E.2.3.8.1 Name [R]

Element Specification

Sets/Name (1)

Description

This element enables specification of set name. For example, 'Data_Fields'.

Sample Value

"Data_Fields".

Tags

Required

Sample Mappings

ISO 19115-1:

/
mdb:MD_Metadata/mdb:identificationInfo/mri:MD_DataIdentification/mri:citation/cit:CI_Citation/cit:identifier/mcc:MD_Identifier[mcc:description/gco:CharacterString[.='Sets']]/mcc:code/gco:CharacterString = Name:

E.2.3.8.2 Type [R]

Element Specification

Sets/Type (1)

Description

This element enables specification of set type. For example, if the variables have been grouped together based on a particular theme, such as

wavelength, then the type should be set to that theme, otherwise it should be set to 'General'.

Sample Value

"SWIR Bands"

Tags

Required

Sample Mappings

ISO 19115-1:

/ mdb:MD_Metadata/mdb:identificationInfo/mri:MD_DataIdentification/mri:citation/cit:CI_Citation/cit:identifier/mcc:MD_Identifier[mcc:description/gco:CharacterString[.='Sets']]/mcc:code/gco:CharacterString = Type:

E.2.3.8.3 Size [R]

Element Specification

Sets/Size (1)

Description

This element specifies the number of variables in the set. For example, if the number of variables in the set is fifteen, the size should be recorded as '15'.

Sample Value

"15"

Tags

Required

Sample Mappings

ISO 19115-1:

/
mdb:MD_Metadata/mdb:identificationInfo/mri:MD_DataIdentification/mri:citation/cit:CI_Citation/cit:identifier/mcc:MD_Identifier[mcc:description/gco:CharacterString[.='Sets']]/mcc:code/gco:CharacterString = Size:

E.2.3.8.4 Index [R]
Element Specification

Sets/Index (1)

Description

This element specifies the index value within the set for this variable. For example, if this variable is the third variable in the set, the index value should be '3'.

Sample Value

"3"

Tags

Required

Sample Mappings

ISO 19115-1:

/
mdb:MD_Metadata/mdb:identificationInfo/mri:MD_DataIdentification/mri:citation/cit:CI_Citation/cit:identifier/mcc:MD_Identifier[mcc:description/gco:CharacterString[.='Sets']]/mcc:code/gco:CharacterString = Index:

E.2.3.9 Sampling

Elements

Sampling (0..*)

Sampling/SamplingMethod [R]

Sampling/MeasurementConditions

Sampling/ReportingConditions

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. MeasurementConditions and ReportingConditions are useful metadata for field campaign data sets.

Tags

Recommended

Sample Mappings

ISO 19115-1:

mrl:LI_Lineage/mrl:LE_ProcessStep/mrl:description/gco:CharacterString

and

mrl:LI_Lineage/mrl:LE_ProcessStep/mrl:rationale/gco:CharacterString

E.2.3.9.1 SamplingMethod [R]

Element Specification

Sampling/SamplingMethod (1)

Description

The name of the sampling method used for the measurement. For example, 'radiometric detection within the visible and infra-red ranges of the electromagnetic spectrum'.

Tags

Required

Sample Mappings

ISO 19115-1:

TBD

E.2.3.9.2 MeasurementCondition

Element Specification

Sampling/MeasurementCondition (0..1)

Description

Conditions at the time the observation or measurement was recorded. For example, 'Sampled Particle Size Range: 90 - 600 nm'.

Tags

Recommended

Sample Mappings

ISO 19115-1:

/
mdb:MD_Metadata/mdb:contentInfo/mrc:MD_CoverageDescription/mrc:attributeGroup/mrc:MD_AttributeGroup/mrc:attribute/mrc:MD_SampleDimension/mrc:otherProperty:MeasurementCondition/gco:CharacterString

E.2.3.9.3 ReportingCondition

Element Specification

Sampling/ReportingCondition (0..1)

Description

Conditions over which the observation or measurement are valid. For example, 'STP: 1013 mb and 273 K'.

Tags

Recommended

Sample Mappings

ISO 19115-1:

/
mdb:MD_Metadata/mdb:contentInfo/mrc:MD_CoverageDescription/mrc:attributeGroup/mrc:MD_AttributeGroup/mrc:attribute/mrc:MD_SampleDimension/mrc:otherProperty:ReportingCondition/gco:CharacterString

Appendix E-1 Tags Glossary

The following table lists all tags used in this model and provides a description of the tags' usage.

Table E.5: Tags 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 E-2 Inputs

1. Discussions with Kathleen Baynes, Jon Pals (ECS Science Office), Dan Pilone (EED Chief Technologist), Abe Taaheri (HDF-EOS Lead Developer), Christopher Lynnes (NASA/ESDIS), Mahabal Hegde (NASA/GSFC)
2. Inputs from Kathleen Baynes

   • Original Appfriendly EOSDIS Science Information Retriever (AESIR)

   proposal:
   https://docs.google.com/document/d/1H0EAPV3Iv3WtheOAVZrJ1eY89NneyiweCcHAz5eI9a0/edit#heading=h.8r4j53o9wk3j

   • Variable data use case: http://giovanni.gsfc.nasa.gov/giovanni/

   • Variable response from AESIR: AESIRresponse.xml file

   • ODISEES tool link: http://odisees.larc.nasa.gov/

   (Point of Contact for this is Beth Huffer:
   https://wiki.earthdata.nasa.gov/display/~bhuffer)

   • More information about the GIBS API:
   https://wiki.earthdata.nasa.gov/display/GIBS/GIBS+API+for+Developers

   • Existing/Preliminary UMM-V Model:
   https://docs.google.com/spreadsheets/d/1Zm5i_Mln0jGGHDS9wC8vzNZAqZ034bu5aSk58Bn77Ig/edit#gid=0 (to be refined by the GIBS Technical Working Group)

   • GIBS TWG Wiki Space:
   https://wiki.earthdata.nasa.gov/display/GIBS/GIBS+Technical+Working+Group

   • Current Status of UMM-V:
   https://wiki.earthdata.nasa.gov/display/GIBS/Visualization+Metadata

3. Inputs from Edward Seiler regarding AESIR questions and content

   • Giovanni application: http://giovanni.gsfc.nasa.gov/giovanni/

   • The information for each variable is populated via an editor named EDDA at

http://dev-ts1.gesdisc.eosdis.nasa.gov/EDDA/index.html

• The help page for EDDA at

http://dev-ts1.gesdisc.eosdis.nasa.gov/EDDA/EDDA_help.html

4. SMAP Data – sample 'Freeze Thaw' variable data from Abe Taaheri: SMAP_L3_FT_A_20140122_R11160_001.h5
5. CF Conventions and CF Standard Variable Names: http://cfconventions.org/Data/cf-standard-names/27/build/cf-standard-name-table.html
6. CSDMS Standard Variable Names: http://csdms.colordo.edu/wiki/CSN_Examples
7. The CSDMS Standard Names: Cross-Domain Naming Conventions for Describing Process Models, Data Sets Their Associated Variables, S.D. Peckham, University of Colorado, INSTAAR, 1560 30th Street, Boulder, CO (Scott.Peckham@colorado.edu) http://www.iemss.org/sites/iemss2014/papers/iemss2014_submission_263.pdf
8. CMR Data Partner User Guide: https://wiki.earthdata.nasa.gov/display/CMR/CMR+Data+Partner+User+Guide
9. Towards Unifying NASA Earth Science Enterprise-Wide Metadata Around International Standards: Study Results and Recommendations, S.J.S. Khalsa, CIRES, University of Colorado, Boulder, CO 80309 USA – sjsk@nsidc.org, S.F. Browdy, OMS Tech, Orlando, Florida, USA – steveb@omstech.com, B.H. Weiss, Jet Propulsion Laboratory, Pasadena, CA – Barry.h.weiss@jpl.nasa.gov http://www.isprs.org/proceedings/2011/ISRSE-34/211104015Final00852.pdf
10. The Dataset Interoperatbility WG https://wiki.earthdata.nasa.gov/display/ESDSWG/Dataset+Interoperability+Working+Group
11. The Unified Modeling Language Reference, Rumbaugh, Jacobson, I, Booch, G, 199, Addison Wesley
12. Summary of British Oceanographic Data Centre Holdings: https://www.bodc.ac.uk/resources/inventories/holdings/search/

Appendix E-3 Keywords and Measurements Governance Structure

The Governance Structure shown in Figure E.50 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.

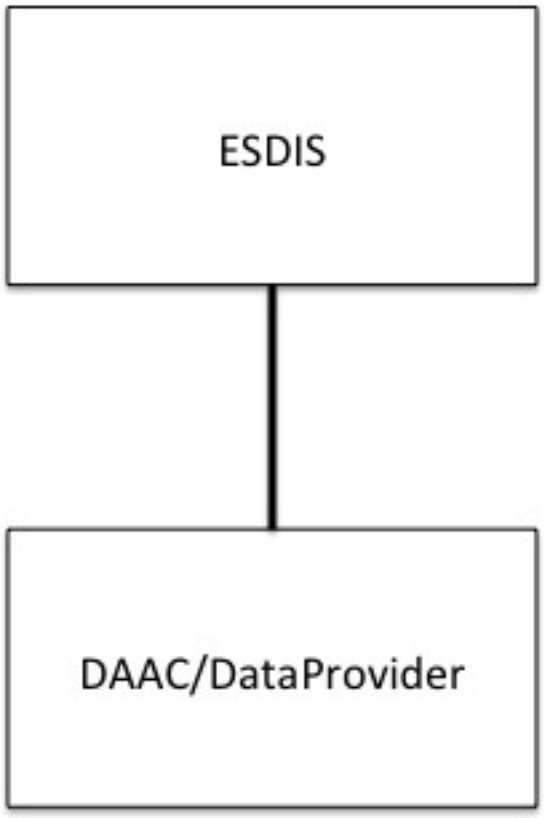

Figure E.50 : Suggested Governance Structure

Adding Keywords or Measurements are expect 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, e.g., MMT. 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: http://gcmd.nasa.gov/learn/rules.html
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, e.g. CSDMS, CF Conventions, etc.

The process by which measurements may be selected is simple.

1. Determine level: i.e. Atmosphere, Oceans, Land (highest) or Atmosphere Air Temperature (mid), or Atmosphere Air Temperature Saturated Adiabatic Lapse Rate (lowest), etc.

2. 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.

3. Select the most appropriate measurement to suit the need. If the measurement does not exist, apply crosswalk to another standard, i.e. CSDMS to CF convention Standard Names.

4. Add to the measurements list stored in the CMR so that all future users can use this measurement.

5. Map these to the appropriate variables.

6. Include a definition of the uncontrolled measurement.

Appendix E-4 Analysis of CSDMS and CF Standard Names as a Source of Tagging

Analysis of CSDMS Standard Names as a source of tagging

The CSDMS (Community Surface Dynamics Modeling System) modeling framework provides mechanisms that allow models and data sets from different contributors (i.e. from different geoscience domains: hydrology, oceanography, meteorology, seismology). The framework defines an approach to the semantic mediation problem. It offers a unique approach to solving this problem by offering a set of standardized and precise descriptions of each variable. It provides a holistic approach to solve the semantic mediation problem 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.

CSDMS Standard Names are grouped, i.e. by Variables Names for: the 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 are groups are suitable for EOS: Variables Names for: the Atmosphere, Oceans, Radiation, Sea Ice, Soil, Snow, Topography.

Potentially, other groups are suitable for EOS: Variable Names for: Basins, Channel, Earthquakes, Glaciers, Planets, River Deltas, Sea Floor Debris.

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 + 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, precipitation mass flux.

The first example of combining Object Name + 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 COARDS conventions.

Most of the CF standard names have been derived from guidelines which have drawn on ECMWF, and NCEP GRIB tables, the PCMDI and GCMD.
CF standard names consist of lower-letters, digits and underscores, and begin with a letter. Upper case is not used.
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 E-5 Abbreviations and Acronym List

AESIR	Application friendly EOSDIS Science Information Retriever
BODC	
CERES	
CF	Climate and Forecast metadata
CH20	
CMR	Common Metadata Repository
COARDS	Cooperative Ocean/Atmosphere Research Data Service
CN	
CSDMS	
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
EDSC	
EOS	Earth Observing System
EOSDIS	Earth Observing System Data and Information System
ESDIS	Earth Science Data and Information System
ESI	EOSDIS Service Interface
ESO	Earth Science Office
GCMD	Global Change Master Directory
GES DISC	
GHRSST	
GIBS	Global Imagery Browse Services
GRIB	GRIdded Binary file format
GUI	
HDF	
IGBP	
ISO	International Organization for Standardization
KM	
KMS	Keyword Management System
LAADS	
LST	
M	
MENDS	

MMT	
NASA	National Aeronautics and Space Administration
NCEP	National Centers for Environmental Prediction
NETCDF	
NIR	
NM	
NO	
NO2	
O3	
PCMDI	Program for Climate Model Diagnosis and Intercomparison
PI	
SERF	Service Entry Resource Format
SO2	
SST	
SW	
TBS	To Be Supplied
TOA	
UI	
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-P	
UMM-S	Unified Metadata Model - Services
UMM-Var	Unified Metadata Model - Variables
UMM-Vis	Unified Metadata Model - Visualization
URI	Uniform Resource Identifier
URL	Uniform Resource Locator
XML	Extensible Markup Language
XPath	XML Path Language