NASA STANDARDS FOR EARTH REMOTE SENSING DATA

Liping Di[*], Ben Kobler [**]
[*]Raytheon ITSS, USA
lpd@rattler.gsfc.nasa.gov
[**]NASA Goddard Space Flight Center, USA
koblerb@rattler.gsfc.nasa.gov

Working Group II/7

KEY WORDS: Standard, Remote Sensing, Data, NASA, Earth Science, Global Change.

ABSTRACT

Earth Observing System (EOS) is one of NASA's large programs for Earth science and global change research. In order to process, distribute, and archive data from EOS satellites and to facilitate interdisciplinary research, NASA has developed the EOS Data and Information System (EOSDIS) that defines a set of standards to deal with remote sensing data. In EOSDIS, data are grouped into three conceptual data models: grid, swath, and point. Each model defines the data object, the components of the object, and the relationship among the components. The EOSDIS conceptual data models are implemented in HDF by mapping the conceptual models into HDF internal data structures. The implementation of the conceptual data models in HDF is called HDF-EOS. In addition, EOSDIS has also developed metadata standards. There are two types of metadata in EOSDIS: structure and product. The structure metadata define the relationship among data components of a data model while the product metadata describe the data product. Both product and structure metadata are encoded in text format by using Object Description Language (ODL) and stored as attributes in HDF-EOS files. NASA provides HDF-EOS libraries for creating and accessing HDF-EOS files.

In order to broaden the use of EOS data and facilitate data interoperability, NASA has actively participated in the development of remote sensing data standards through U.S. Federal Geographic Data Committee (FGDC), ISO Technical Committee 211, and Open GIS Consortium (OGC). Currently NASA is leading the development of several remote sensing data standards for those standardization organizations.

1   INTRODUCTION

As a part of the U.S. Global Change Research Program, the National Aeronautics and Space Administration (NASA) is developing and implementing the Earth Observing System (EOS). The system provides the next generation of satellite remote-sensing instruments and platform hardware, a community of funded scientists, and the infrastructure to consolidate data and information from surface campaigns and remote-sensing satellites. It is planned that the EOS program will observe the Earth continuously for the next 15 years (Asrar and Greenstone, 1995).

In order to process, distribute, and archive data from the EOS satellites and to facilitate interdisciplinary research, NASA is developing the EOS Data Information System (EOSDIS) Core System (ECS). EOSDIS is central to the EOS program. It provides the environment in which scientists, educators, policy makers, and other users access and use reliable large-scale data sets describing geophysical and biological processes. It also facilitates interaction and communication among EOSDIS users (Asrar and Dozier, 1994). Because EOSDIS will process diverse data from field campaigns, aircraft and ship observations, satellite remote sensing, and the output of scientific models, it is unrealistic for EOSDIS to deal with data from individual sources individually. Therefore, EOSDIS has established a set of standards to deal with remote sensing data.

2   EOSDIS DATA STANDARDS

2.1   Conceptual Data Models

Common conceptual data models are the basis for developing data encoding and format standards and software interfaces. So far, NASA has developed three conceptual data models to describe data in EOSDIS, including:
- Grid: rectangular arrays or data structures for geocoded data (Figure 1).
- Swath: simple or complex remote-sensing swath data (Figure 2).
- Point: simple or indexed geolocated and/or time-tagged point observations or event data (Figure 3).

Each conceptual data model defines the data object (i.e., Swath, Grid, Point), the components of the object, and the relationship between the components. In Swath and Grid data models, there are two required components: data and geolocation. The description of how to relate the geolocation component to the data component has to be packaged with those two components. In the Point data model, data collected at a point on Earth can be a record in a table and the geolocation information for the record is either fields in the record or attached to the record through a pointer. The detailed descriptions of the three EOSDIS data models can be found at http://edhs1.gsfc.nasa.gov/.

At least three basic data types are needed for describing the EOSDIS data models:
- ASCII text--plain and formatted text for labels, descriptions or documents;
- Tables--standard tables for record-like data; index tables for indices and variable length counters for ragged arrays;
- n-dimensional array--array of records or array of scalars for heterogeneous or homogeneous multi-dimensional data.

For the Swath and Grid data models, all three basic data types are needed. For the Point data model, only the table and the ASCII text are needed.

The three data models are applicable to most of EOSDIS data products. However, with the increase of use of GIS in the global change research, there is an increasing need for handling vector-based GIS data in EOSDIS. Currently, NASA is studying the possibility of adding two more conceptual data models, namely Vector and Polygon, into EOSDIS.
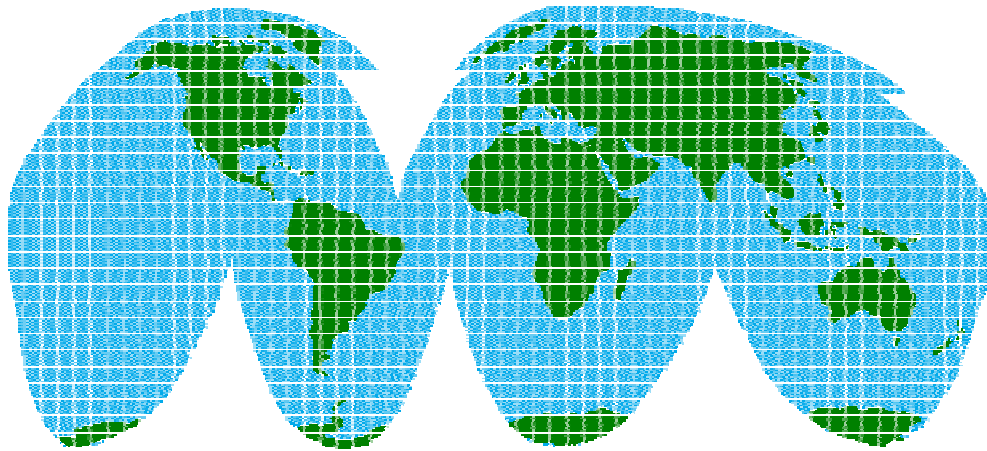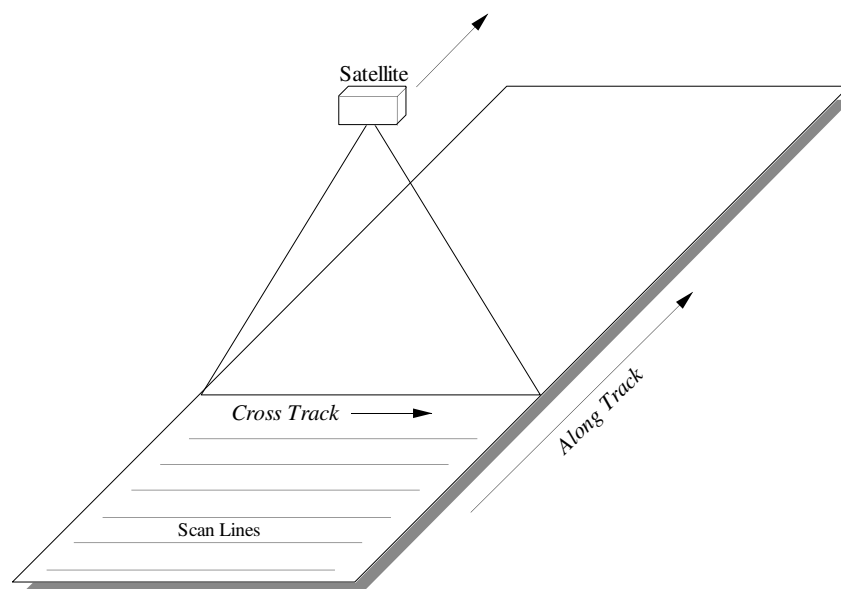

Figure 1. Grid Concept
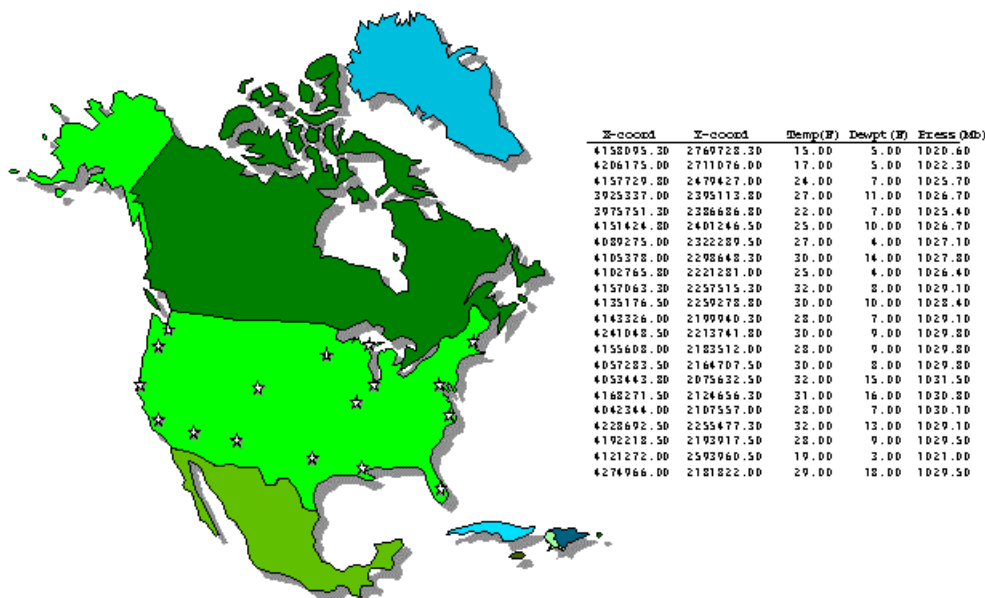

Figure 2. Swath Concept

X-coord	Y-coord	Temp(F)	Dewpt(F)	Press(Mb)
4158095.3D	2769728.3D	15.DD	5.DD	1D2D.6D
42D6175.DD	2711D76.DD	17.DD	5.DD	1D22.3D
4157729.8D	2479427.DD	24.DD	7.DD	1D25.7D
3925337.DD	23D5113.8D	27.DD	11.DD	1D26.7D
3975751.3D	2386686.8D	22.DD	7.DD	1D25.4D
4151424.8D	24D1246.5D	25.DD	1D.DD	1D26.7D
4D89275.DD	2322289.5D	27.DD	4.DD	1D27.1D
41D5378.DD	2298648.3D	3D.DD	14.DD	1D27.8D
41D2765.8D	2221281.DD	25.DD	4.DD	1D26.4D
4157D63.3D	2257515.3D	32.DD	8.DD	1D29.1D
4135176.5D	2259278.8D	3D.DD	1D.DD	1D28.4D
4143326.DD	2199940.3D	28.DD	7.DD	1D29.1D
4241D48.5D	2213741.8D	3D.DD	9.DD	1D29.8D
4155608.DD	2183512.DD	28.DD	9.DD	1D29.8D
4D57283.5D	2164707.5D	3D.DD	8.DD	1D29.8D
4D53443.8D	2D75632.5D	32.DD	15.DD	1D31.5D
4168271.5D	2124656.3D	31.DD	16.DD	1D3D.8D
4D42344.DD	21D7557.DD	28.DD	7.DD	1D3D.1D
4228692.5D	2255477.3D	32.DD	13.DD	1D29.1D
4192218.5D	2193917.5D	28.DD	9.DD	1D29.5D
4121272.DD	2593960.5D	19.DD	3.DD	1D21.DD
4274966.DD	2181822.DD	29.DD	18.DD	1D29.5D

Figure 3. Point Concept

2.2 Metadata

There are two groups of metadata in EOSDIS -- structure metadata and the product metadata. Both product and structure metadata are encoded in text format by using Object Description Language (ODL). The structure metadata define the relationship between data components, as well as how the data components are stored in data files. The content standard for the structure metadata is defined in the EOSDIS conceptual data models. Normally, the structure metadata are generated, encoded and stored automatically by data production system when a data product is created.

The product metadata describe the data product. EOSDIS has developed its own metadata model and metadata standard (Heller, 1994). The metadata standard is compatible with the FGDC Content Standard for Geospatial Metadata (FGDC, 1998), e.g., the equivalencies of all metadata elements required by the FGDC standard can be found in the EOSDIS metadata standard. However, the EOSDIS standard provides more metadata elements in the remote-sensing domain, such as platform information, sensor information, the history of data processing, and the spatial representation of sensor data.

2.3 Levels of Data Products

The degree of data processing applied to a data product determines many important characteristics of the product. It also determines if some specific metadata elements or data services are applicable to the product. For example, geographic coordinate -based data subsetting can be easily implemented in the georectified raster data, but the same type of service is difficult to be provided to raw remote sensing images. In order to facilitate the data management and standardize the metadata and data services, data products in EOSDIS are classified into five levels according to the degree of processing—the higher the level, the higher the degree of processing (Asrar and Greenstone 1995):

- Level 0-Reconstructed, unprocessed data directly from the instruments at full resolution.
- Level 1A-Reconstructed, unprocessed instrument data at full resolution, time referenced, and annotated with ancillary information, including geometric and radiometric calibration coefficients and georeferencing parameters.
- Level 1B-Level 1A data that have been processed to sensor units.
- Level 2-Derived geophysical variables at same resolution and location as the Level 1 source data.
- Level 3-Variables mapped on uniform space-time grid scales, usually with some completeness and consistency.
- Level 4-Model output or results from analyses of lower level data.

Among the above data levels, most data at level 0-2 can fit into the Swath data model while level 3 and level 4 data fit into the Grid data model. Although the above classification of data levels is the standard for EOSDIS, the classification, some missions in EOS may also use other level names that are not defined in the EOSDIS classification of data levels, due to historical or practical reasons. For example, Landsat 7 designated data level 1G as the product that is georectified with pixel values still in sensor units.

2.4 Data Format

To implement the conceptual data models in EOSDIS, one or several physical data formats have to be selected or developed as standard format(s) to host the data models. As the result of an investigation, NASA decided that selecting a single physical format as the EOSDIS standard format would allow Earth science data users to access data more easily and efficiently than multiple formats for performing inter-disciplinary global change research. The advantages of using a single standard format in EOSDIS include:

- promoting distribution, exchange, and sharing of Earth science data;
- simplifying data manipulation, analysis, visualization, and management;
- minimizing the cost of system development and maintenance;
- optimizing system performance;
- supporting long-term preservation of the scientific data.

However, accommodating all the data models in a single data format is not easy. The format must be flexible enough to store all possible EOSDIS data types. Through detailed study, NASA selected the Hierarchical Data Format (HDF) (NCSA, 1994) as the basis for developing the EOSDIS standard format, HDF-EOS.

2.4.1 Hierarchical Data Format (HDF)

HDF has been developed for facilitating data access in heterogeneous computing environments by the National Center for Supercomputing Applications (NCSA), University of Illinois at Urbana-Champaign (NCSA, 1994) since the late 1980s. It has six internal data models, including:

- The 8-bit raster model. Stores and retrieves 8-bit raster images, their dimensions, and palettes.
- The palette model. Stores and retrieves 8-bit palettes outside the 8-bit raster model.
- The 24-bit raster model. Stores and retrieves 24-bit images and their dimensions.
- The scientific data model (Scientific Data Sets, SDS). Store and retrieves multi-dimensional arrays of integer, floating-point, or n-bit numbers, their dimensions, number type, and attributes.
- The annotation model. Stores and retrieves the text strings used to describe a file or any of the data elements it contains.
- The virtual data model table (Vdata and Vgroup). Vdata is a record based structure wherein record fields may be defined, named, and typed individually. Vdatas are one-dimensional arrays of records. Each Vdata may be given an optional name and/or class string for identification purposes. Vgroup is a structure for associating sets of data objects. Vgroups may "contain" any HDF objects, including other Vgroups. Like Vdatas, Vgroups may be given an optional name and/or class string for identification purposes.

For any of the above data models, metadata (called attributes in HDF) can be attached. Attributes are stored as named values or lists of values, all in same data type. An attribute can be global (pertaining to the entire file) or local (associated with an individual data object).

HDF is an object-based data format. Each data object in an HDF file can have a user-defined name. To support the HDF format, NCSA provides a set of libraries for accessing and managing HDF files. In traditional formats, data users have to know the exact physical location of an object in the file before the object can be accessed. However, for accessing data in HDF, users do not need to know how a data object is stored in the file and where its physical location is. An application program can directly access a data object through its name by calling HDF library routines. All details of data accesses are handled by HDF libraries. This procedure greatly simplifies user access to the complicated data structures. The HDF libraries and documents can be obtained at no cost from ftp://ftp.ncsa.uiuc.edu/.

The above description of HDF is based on HDF major version 4. NASA and NCSA plan to continuously maintain this version of HDF. In addition, NCSA has developed HDF version 5, originally called BigHDF. HDF 5 has totally different internal data structure from HDF 4. Basically HDF 4 and HDF 5 are two different data formats.

The main reasons for NASA to select HDF include the following:
1. Multiple data type support - All data types in EOSDIS can be supported by HDF.
2. Portability - Data in HDF can be ported to many different platforms.
3. Easy to use and implement - HDF provides software libraries and good documentation.
4. HDF software and documents are free and available on anonymous FTP.
5. Availability of software tools for manipulating and visualizing data in HDF. Most of the tools are free.

2.4.2 HDF-EOS

The EOSDIS conceptual data models have been implemented in HDF 4 by mapping the conceptual models to HDF internal data models and structures. Among the HDF data models, the 8-bit raster, 24-bit raster, and palette models are

used in EOSDIS mostly in browse and visualization products. The annotation model is somewhat obsolete and is less useful in EOSDIS. Among all the data models and objects in HDF, the Vdata, SDS, Vgroup, and their associated attribute structures are most relevant to the EOSDIS data models. Based on these HDF data models, EOSDIS has developed three more HDF data models—the Swath, Grid, and Point—and the associated application program interfaces for implementing the three EOSDIS conceptual data models. The three data models together are called HDF-EOS. Currently NASA is evaluating the possibility of implementing HDF-EOS in HDF 5. Even if NASA decides to implement HDF-EOS in HDF 5, the application programming interface of HDF-EOS will not be changed so that no code change is need for all HDF-EOS-based application programs.

Figures 4, 5, and 6 show the three HDF-EOS models, which are the implementation of the three corresponding EOSDIS conceptual data models. In both Swath and Grid models, instrument data are stored in arrays or tables. The major difference between the two is the implementation of georeferencing. In Swath model, scanlines have the associated geolocation records in a geolocation component. The records contain information about the georeferencing-related information, such as geolocation of certain pixels within a scanline, the start and end time of a scanline, orbit parameters, etc. Because of the geocoded nature of the Grid data, the geolocation information in Grid model is much simpler than that in Swath model. For each grid, only one geolocation Vdata is required. The Vdata contains information about the map projection, which can be used to project grid-cell coordinates back to geographic coordinates. The point data model is implemented by using mainly Vdata.

One important requirement for EOSDIS is to keep the metadata with the data so that when a user obtains a data product from EOSDIS, s/he will automatically get all necessary metadata information about the product. In HDF-EOS, both product metadata and the structure metadata are encoded in Object Description Language (ODL) and stored as attributes in HDF-EOS files.
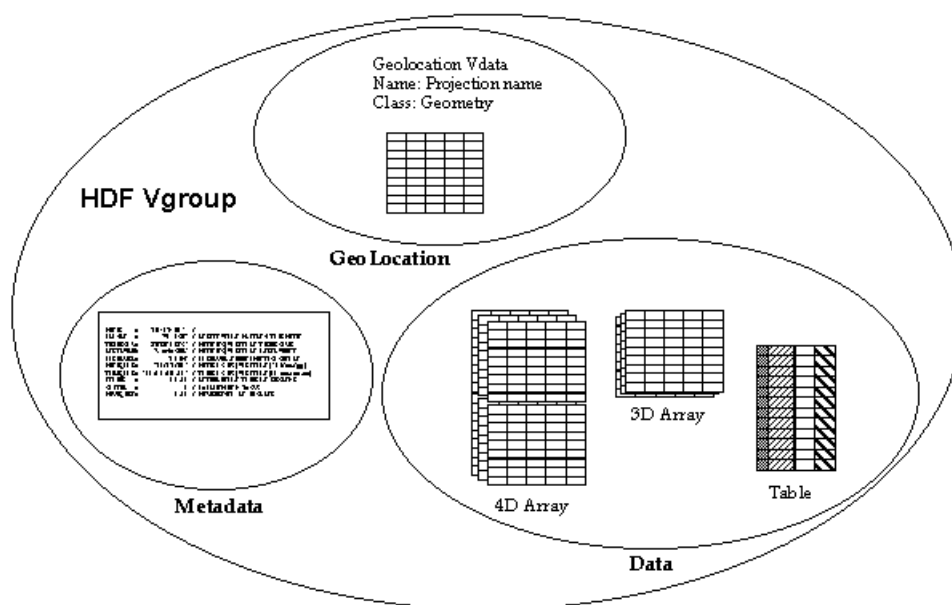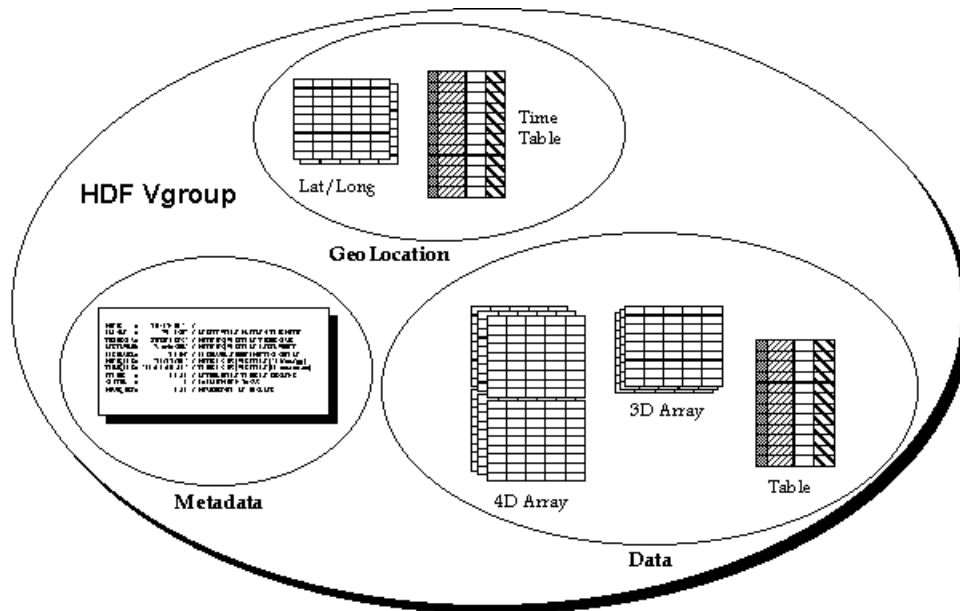

Figure 4. Grid model in HDF-EOS

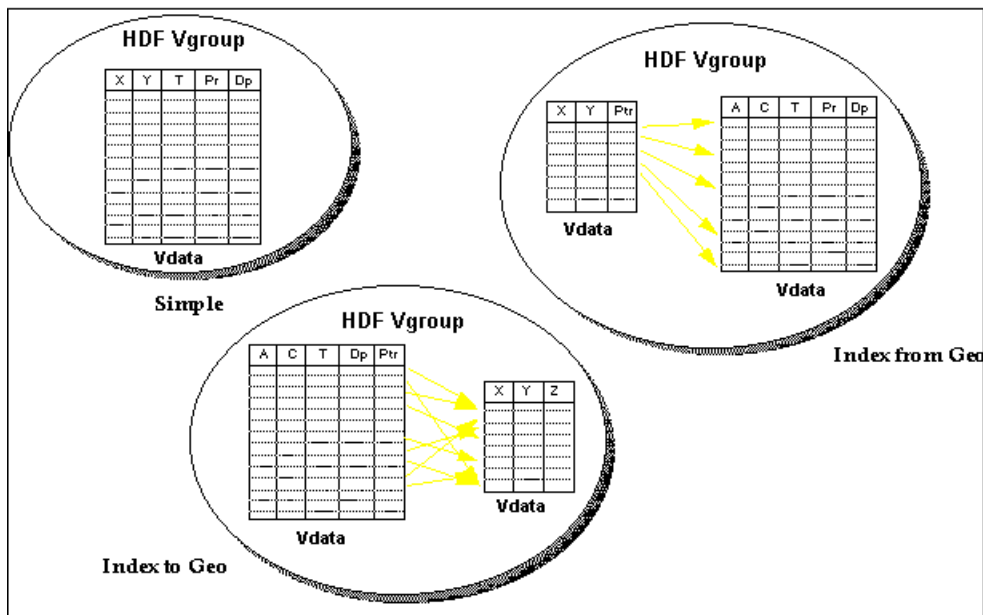Figure 5. Swath model in HDF-EOS

HDF File

Figure 6. Point model in HDF-EOS

3   PROMOTING THE USE OF NASA EARTH REMOTE SENSING STANDARDS

Global change research requires integrated analysis of huge amount of multi-disciplinary and multi-source data. International cooperation and data sharing is one of the keys for successes in the global change studies. The data for a research project may come from many different sources. Therefore, making data and information in EOSDIS interoperable with other players in the global change research community are very important for EOSDIS. To achieve the interoperability, either NASA EOSDIS standards are adopted by other players as their standards, or EOSDIS standards are compatible with other's standards, or NASA EOSDIS adopts other's standards. Since NASA has already invested large amount of money in EOSDIS, changing the data standards in EOSDIS is vary unlikely both politically and economically. Therefore, promoting the use of NASA EOSDIS standards by broad Earth science community is the best way for NASA to achieve the data interoperability. For this purpose, NASA is actively participating in the U.S.

federal, national, and international standards activities for promoting NASA standards and for sharing NASA's experiences and expertise on the development of Earth remote sensing data standards. Currently NASA has been involved in standard development activities of U.S. Federal Geographic Data Committee (FGDC), the National Committee on Information Technology Standards (NCITS) Technical Committee for Geographic Information (L1), Open GIS Consortium (OGC), and International Organization for Standardization Technical Committee 211 (ISO/TC211).

3.1   U.S. FGDC Standards Based on NASA EOSDIS Standards

The United States Federal Geographic Data Committee (FGDC) was established under U.S. President Executive Order 12906 to coordinate the development of the National Spatial Data Infrastructure (NSDI). The FGDC develops geospatial data standards for implementing the NSDI. NASA, as a member of FGDC, has actively participated in the development of FGDC standards for geospatial data. The EOSDIS data models are in the process of becoming FGDC content standards for remote sensing data. The FGDC Content Standard for Remote Sensing Swath Data (Di and Carlisle, 1998; FGDC 1999), developed by the NASA-led FGDC Imagery subgroup, is based on the EOSDIS Swath conceptual data model. The Extensions for Remote Sensing to the FGDC Content Standard for Digital Spatial Metadata is currently under development by the FGDC Imagery Subgroup (NASA, 1998). These extensions are based on the EOSDIS Core Metadata Standard with additional input from U.S. and international organizations.

3.1.1   Content Standard for Remote Sensing Swath Data

The Imagery Subgroup of the FGDC Standard Working Group (SWG) has developed the Content Standard for Remote Sensing Swath Data (FGDC, 1999). The FGDC Standards Reference Model (FGDC 1997) specifies that content standards provide semantic definitions of a set of objects and the relationships among them. The Swath standard defines the minimum content for remote sensing swath data, based upon the HDF-EOS Swath data model with generalizations and extensions.

A swath is produced when an instrument scans approximately perpendicular to a moving point. As illustrated in Figure 2, the path of the moving point is called the *track* dimension and that of the scan is the *cross-track* dimension. Another application of the swath data model is ground-based radar, which sweeps in an azimuthal direction (the cross-track dimension) beginning and ending at different points along a particular radial path (the track direction). In the swath data model, geolocation information is stored at selected (track, cross-track) grid locations and the data elements are mapped to (track, cross-track) points in the geolocation grid system to match the data elements with the geolocation parameters. Data elements can be either sensor measurements or geophysical parameters. Geolocation information has to be provided in at least one of the three following ways:

- Date/Time and Latitude/Longitude or Colatitude/Longitude pairs
- Date/Time, platform attitude and position
- Analytic fit of Latitude/Longitude or Colatitude/Longitude to (track, cross-track index position).

The major difference between the EOSDIS Swath data model and the FGDC standard is that the FGDC standard is more flexible in providing the geolocation information. Therefore, the EODIS Swath data model can be considered as a profile of the FGDC standard.

3.1.2   Content Standard for Digital Geospatial Metadata: Extensions for Remote Sensing Metadata

In June 1998, the FGDC issued version 2 of the Content Standard for Digital Geospatial Metadata (FGDC 1998). This standard has been used widely by many data providers and was the basis for the draft ISO TC 211 metadata standard. However this standard did not address much of the metadata content needed for describing Earth remote sensing data. Therefore, NASA proposed to develop extensions for remote sensing metadata to the FGDC standard based on the EOSDIS metadata standard in October 1998. In April 1999 FGDC approved the proposal, and the development of the extensions is in progress, under the FGDC Imagery Subgroup. Currently, a draft standard has been produced by a team consisting of members form the FGDC Imagery Subgroup and a variety of agencies, organizations, and institutions. It is now in internal committee draft form. It is expected that the draft will become an FGDC standard within year 2000.

In the extensions, most of metadata elements are taken from the ECS metadata standard (Heller, 1994). However, metadata elements are also included from other sources, such as ISPRS Working Group II/7. The major extensions of metadata are in the following areas:

- The relation of a data set or collection of data to component parts or to a larger collection to which it belongs.
- The production of higher level data from lower level data (section 3.1) and the algorithms used.
- User's guides, scientific papers, and other documents describing such aspects of the data as the organization, production, algorithm, and usage.
- The properties of the measuring instrument.

- Position of the measuring platform and its time dependence.
- Instrument attitude

The detailed description of this FGDC metadata extension project can be found in Di, et. al., (2000).

3.2   NASA's Involvement in the Development of NCITS L1, ISO TC 211, and OGC Standards

NCITS L1 is the U.S. national standardization body responsible for geographic information standards. It is also the U.S. representative organization to the ISO TC 211. NASA is a voting member of NCITS L1. NASA's participation of ISO TC 211 activities is through the NCITS L1. NASA staff and contractors serve as members of U.S. delegation and as U.S. experts to ISO TC 211. Currently NASA has been actively involved in the ISO TC 211 Imagery and Gridded Component project, the Coverage Geometry standard, the Metadata standard, and the geographic data service standards. Since NASA has significant experience in the development and implementation of remote sensing standards, NASA's participation can make significant contribution to ISO TC 211 for developing practical and implementable international standards in remote sensing-related areas.

The Open GIS Consortium, Inc.(OGC) is a not-for-profit membership organization that address the lack of interoperability among systems that process georeferenced data, and between these systems and mainstream computing systems. One of the major activities of OGC is to develop technology standards and business process innovations that support widespread adoption and use of georeferenced data and services throughout the global economy (OGC, 1999). NASA is one of the principal members and the major sponsors of OGC. Currently, NASA has been involved in development of OGC specifications on Catalog, Web Mapping, Geographic Markup Language (GML), Coverage Geometry, and others.

4   CONCLUSION

NASA has developed and implemented a series of Earth remote sensing data standards through its EOSDIS project. Since EOS is one of the largest Earth remote sensing missions in the world, huge amount of remote sensing data will be produced according to EOSDIS standards. In the coming years, global change researchers and Earth scientists will encounter many EOS data products compliant with EOSDIS data standards. NASA's promotion of EOSDIS standards will broaden the user community of EOS data and make the interoperability of the data easier. During the development of EOSDIS standards, NASA has gained experiences on developing remote sensing data standards. NASA is sharing these experiences with remote sensing communities by participating in the standard development activities of U.S. federal, national, and international standardization bodies.

ACKNOWLEDGEMENTS

REFERNENCES

Asrar, G., and Dozier, J., 1994. EOS: Science Strategy for the Earth Observing System. American Institute of Physics Press, Woodbury, New York 11797.

Asrar, G., and Greenstone, R., 1995. MTPE EOS Reference Handbook. NASA Goddard Space Flight Center, Greenbelt, MD 20771.

Di, L., B. M. Schlesinger, and B. Kobler, 2000. U.S. FGDC Content Standard for Digital Geospatial Metadata: Extensions for Remote Sensing Metadata. In: International Archives of Photogrammetry and Remote Sensing, Amsterdam, The Netherlands, Vol XXXIII, Part B.

Di, L. and Carlisle, C., 1998. The Proposed FGDC Content Standard for Remote Sensing Swath Data. Proceedings, ASPRS/RTI 1998 Annual Conference, Tempa, Florida (CD-ROM, 12 pages).

FGDC, 1999. The FGDC Content Standard for Remote Sensing Swath Data. FGDC-STD-009-1999. Reston, Virginia.

FGDC, 1998. Content Standard for Digital Geospatial Metadata. FGDC-STD-001-1998, Federal Geographic Data Committee, Reston, VA-USA.

Heller, D., 1994. Proposed ECS Core Metadata Standard, Release 2.0. Hughes Applied Information System, Landover, MD-USA. http://edhs1.gsfc.nasa.gov/

NCSA, 1994. HDF Reference Manual. University of Illinois at Urbana-Champaign.

NASA, 1998. Proposal for a National Spatial Data Infrastructure Standards Project: Content Standard for Digital Geospatial Metadata: Extensions for Remote Sensing Metadata. http://www.fgdc.gov/standards/documents/proposals/csdgm_rs_pr.html

OGC, 1999. OGC's On-Line Corporate Brochure. http://www.opengis.org/info/brochure/brochure0599.pdf.