SCIENCE QUALITY ASSESSMENT IN SUPPORT OF EOS MISR DATA*

Scott A. Lewicki, Kathleen Crean, and Ralph Kahn
Jet Propulsion Laboratory
Pasadena, CA, USA

ABSTRACT

The Multi-angle Imaging SpectroRadiometer (MISR) instrument, scheduled to fly on board the first Earth Observing System (EOS) spacecraft, will have a data stream that produces over 100 GB of data per day, once in full operation, of information about the Earth's atmosphere, clouds, and surface. The processing of such a vast volume of data on a timely basis, requires innovative and efficient techniques for rapidly assessing data quality, as it flows though several levels of processing, and ultimately into the end products, to be released to the science community. This paper describes the production and analysis of quality assessment (QA) data supplied by the science algorithms which comprise the EOS MISR instrument data processing. A separate QA Summary file is produced for every MISR data product. Described also is the storage of QA data in an unique database, which will allow users to examine this data and make decisions about ordering product files based upon the information. QA data produced during MISR Level 1 processing reports detector health, radiometric conversion statistics, and geometric correction quality indicators. QA info data produced during MISR Level 2 processing includes science algorithm performance indicators, physical and climatological constraints on retrieval results, data processing path indicators, and statistical summaries of these quantities. Current plans for performing QA analyses of the MISR products include automated checks of QA information for all data, as well as using interactive visualization tools for in-depth QA analyses of selected subsets.

1.0 INSTRUMENT OVERVIEW

MISR (Diner et al., 1993) will be launched into polar orbit on the Earth Observing System (EOS) AM1 spacecraft in 1999. The instrument contains nine push-broom cameras to observe at fixed view angles, relative to the surface normal, of 0° (nadir), 26.1°, 45.6°, 60.0°, and 70.5° fore and aft of nadir using charge-coupled device (CCD) line arrays filtered to 443, 555, 670, and 865 nm. The line arrays consist of 1504 photoactive pixels plus 16 light-shielded pixels per array, each 21 μm square. The overlap swath width seen in common by all nine cameras is 360 km, which provides global multi-angle coverage of the entire Earth in 9 days at the equator, and 2 days at the poles. The cross-track instantaneous field of view (IFOV) and sample spacing of each pixel is 275 m for all of the off-nadir cameras, and 250 m for the nadir camera. Along-track IFOV's depend on view angle, ranging from 250 m in the nadir to 825 m at the most oblique angle. Sample spacing in the along-track direction is 275 m in all cameras.

2.0 QUALITY ASSESSMENT OVERVIEW

2.1 QUALITY ASSESSMENT CATEGORIES

Performing accurate data quality assessment (QA) is essential if the MISR data are to be scientifically meaningful to our users. Parts of the MISR QA activity need to occur at each of three sites: (1) in the Science Data Processing (SDP) software, (2) with the DAAC operator, and (3) at the Science Computing

*. Presented at the Thirteenth International Conference and Workshops on Applied Geologic Remote Sensing, Vancouver, British Columbia, Canada, 1-3 March 1999.

Facility (SCF). Data quality indicators are generated as part of the standard MISR processing software, and are reported as part of the standard MISR data products.

2.1.1 Quality Assessment within the SDP software

Routine QA processing will be automated. Human involvement will be limited to (1) spot checking of the data stream, and (2) investigating "anomalies." This puts most of the QA burden on the standard SDP software, which will create "indicators" of key aspects of the data quality and retrieval performance. A top-level diagram of this activity is shown in Figure 1. The indicators and statistics shown as part of the outputs are defined in the parameter tables in this document, and are established as an integral part of the algorithm and software development. These indicators and statistics are captured into an output Summary QA file, as well as the product file they are based upon.

Science Data Processing software

--Instrument performance indicators

--Retrieval path indicators

--Algorithm performance indicators

--Physical constraint indicators

--Climatological likelihood indicators

--Statistical summary data at granule or higher level

→

Outputs

Indicators (subregion, region, domain, block)

Statistics (swath)

Anomaly list (QA Log)

Alarms (automatic e-mail to SCF, DAAC operator)

Figure 1. Automatic QA within the Science Data Processing software

A general description of SDP software QA indicator types is as follows:

(1) **Instrument Performance**: Instrument performance indicators that affect spectral, radiometric, and geometric performance are monitored for engineering purposes, and to effect updates to the instrument calibration parameters. The Level 1 data stream will produce summaries of instrument performance in three areas: (a) radiometric quality, (b) geometric quality, and (c) missing data. These metrics will be compared with sets of limits, and the relevant performance implications encoded into data quality indicators.

(2) **Retrieval Path Indicators**: Decisions made along the retrieval process, regarding which retrieval path to follow, are retained as part of the processing record. These include such choices as: whether an ocean, dense dark vegetation, or heterogeneous land retrieval of aerosol properties is attempted, and whether real-time inputs or climatologies are used for column ozone abundance.

(3) **Physical Constraints**: There are many physical constraints that can be applied to the retrieval results, some of which may be used as indicators of data quality. Some examples are: the requirement of non-negative radiances, and an upper bound on the total aerosol optical depth based on the darkest pixel in the scene.

(4) **Algorithmic Constraints**: Since keeping track of the assumptions and numerical behavior of the algorithm is part of the development effort, these constraints are relatively

easy to identify. They include such items as: (a) convergence characteristics of numerical methods (residuals and number of iterations), (b) the limits of intrinsic assumptions made in the parameterizations used (such as an ocean surface roughness model that is meaningful only within a certain range of wind speeds), (c) case limitations (such as treating pixels that may cross radically different terrain types (e.g., coasts) if the algorithm is designed to assume an "average" terrain type), and rejecting pixels that are too cloudy or with terrain too rough for the retrieval to work.

(5) **Climatological Constraints**: These are "statistical" constraints, which should appear as Warnings rather than as Errors. An "unlikely" result may mean a misinterpretation of the data, or a discovery. Indicators based upon such constraints will be very helpful for the first-order analysis of the MISR Level 2 results. The Aerosol Climatology Product, for example, may indicate that it is more likely to find biomass burning particles than mineral dust particles over a tropical rain forest. Placing climatological constraints on other MISR-retrieved physical parameters (surface albedo and view-dependent reflectances, cloud cover, etc.) requires similar climatologies for these quantities.

External inputs to the SDP software, such as atmospheric surface pressure and wind speed, will be obtained from a data assimilation model or other sources. Climatologies for all the external parameters needed by the MISR retrieval are included in the SDP software, to be used as default values if the external input data are unavailable, in the form of the Terrestrial Atmosphere and Surface Climatology (TASC) Dataset. Such cases will be flagged.

2.1.2 Quality Assessment at the DAAC

QA operations at the DAAC requiring human involvement will involve additional monitoring of alarms, and examining logs created by the real-time data stream, as shown in Figure 2.

Figure 2. QA Operations at the DAAC

The DAAC will also be involved in the operations of the QA Data Management System (DMS) described in §4.0.

2.1.3 Quality Assessment at the SCF

At the SCF, QA amounts to performing those tasks that require the attention of the MISR Instrument Team, Science Team, or Science Data System Team, and performing any processing steps that cannot be automated at the DAAC. The following QA activities are anticipated at the SCF:

(1) Examining the Summary QA File data and error logs produced routinely by the SDP software.

(2) Performing in-depth analysis of algorithm and software results on samples taken from the

MISR data products.

(3) Performing any special investigations indicated by the routine examination of Summary QA File data, by anomalies reported at the DAAC, or by data users at the SCF or elsewhere.

(4) Evaluating instrument and retrieval algorithm results using data from the MISR field validation program.

(5) Documenting SCF QA activities in the QA Log.

(6) Producing those routine statistical summaries that require the accumulation of time series of QA data, which might not be feasible to stage at the DAAC.

3.0 SUMMARY QA FILE CONTENTS AND FORMAT

Summary QA Files will be produced for each of the MISR Level 1 and Level 2 data products. Each file contains QA parameters derived from science parameters in the corresponding product files. The file format for all QA files is native Hierarchical Data Format (HDF), which was developed by the National Center for Supercomputing Applications (NCSA) [4].

Parameters in the QA files are grouped into HDF *vgroups* according to their spatial resolution. For example, all parameters which apply to an entire swath of data are grouped together in a swath vgroup, while all parameters which apply to a MISR 'block' (have we defined a MISR block yet?) are grouped together in a MISR block vgroup. Within each vgroup is defined an HDF *vdata* structure, which contains a set of records corresponding to each QA summary parameter.

As an example, the HDF file format of the summary QA file for the MISR Level 2 Aerosol product is shown in the figure and tables below.

MISR L2AS Aerosol QA HDF File

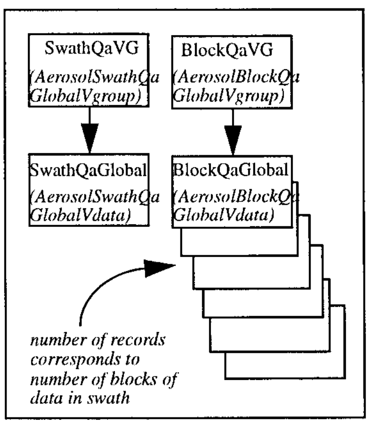

Table 1: HDF Vgroup and Vdata Constructs for the MISR Level 2 Aerosol Product File

HDF Construct	Aerosol QA File
Swath-level Constructs	
Vgroup name	SwathQaVG
Vgroup class	AerosolSwathQaGlobalVgroup
Vdata name	SwathQaGlobal
Vdata class	AerosolSwathQaGlobalVdata
Block-level Constructs	
Vgroup name	BlockQaVG
Vgroup class	AerosolBlockQaGlobalVgroup
Vdata name	BlockQaGlobal
Vdata class	AerosolBlockQaGlobalVdata

Table 2: HDF Vdata Record Definitions for the MISR Level 2 Aerosol QA File

	Data Field Name	Data Type	Field Dimensions
	Swath-level Constructs		
1.	NRegWAnyGoodDwSub	INT32	--
2.	NRegWNoGoodDwSub	INT32	--
3.	NSubWGoodData	INT32	--
	Block-level Constructs		
1.	BlockNumber	INT32	--
2.	ValidRecord	UINT8	--
3.	NRegWAnyGoodDWSub	INT32	--
4.	NRegWNoGoodDWSub	INT32	--
5.	NRegDwAlgProc	INT32	--
6.	NRegDwAlgSucc	INT32	--
7.	NRegDdvAlgProc	INT32	--

Table 2: HDF Vdata Record Definitions for the MISR Level 2 Aerosol QA File

	Data Field Name	Data Type	Field Dimensions
8.	NRegDdvAlgSucc	INT32	--
9.	NRegEofAlgProc	INT32	--
10.	NRegEofAlgSucc	INT32	--
11.	NRegObliqueSunAng	INT32	--
12.	NRegTopoComplex	INT32	--
13.	NRegReglCloudy	INT32	--
14.	NRegNoGoodSubData	INT32	--
15.	NRegOzDAOAvailOOB	INT32	--
16.	NRegOzDAONotAvail	INT32	--
17.	NRegWsDAOAvailOOB	INT32	--
18.	NRegWsDAONotAvail	INT32	--
19.	NRegPsDAOAvailOOB	INT32	--
20.	NRegPsDAONotAvail	INT32	--
21.	NRegTsDAOAvailOOB	INT32	--
22.	NRegTsDAONotAvail	INT32	--
23.	NRegGeopHtNotAvail	INT32	--
24.	NRegPtDAOCalcOOB	INT32	--
25.	NRegTtDAOCalcOOB	INT32	--
26.	NRegCpwModAvailOOB	INT32	--
27.	NRegCpwModNotAvail	INT32	--
28.	NRegCpwDAOAvailOOB	INT32	--
29.	NRegCpwDAONotAvail	INT32	--
30.	NRegCirTCAvail	INT32	--
31.	NRegCirModAvailOOB	INT32	--
32.	NRegCirModNotAvail	INT32	--
33.	NRegStrSagAvailOOB	INT32	--
34.	NRegStrSagNotAvail	INT32	--

Table 2: HDF Vdata Record Definitions for the MISR Level 2 Aerosol QA File

	Data Field Name	Data Type	Field Dimensions
35.	NRegStrModAvailOOB	INT32	--
36.	NRegStrModNotAvail	INT32	--
37.	NSubWGoodData	INT32	--
38.	NSubHighValEqRefl	INT32	36
39.	CloudMaskSrc	INT32	--
40.	CloudShadowMaskSrc	INT32	--
41.	TopoShadowMaskSrc	INT32	--
42.	GeomDataQualInd	FLOAT64	--

4.0 QA DATA MANAGEMENT SYSTEM

The Summary QA Files, described above, will be ingested to a Data Management System (DMS) immediately after they are produced. The MISR QA DMS is being developed by the company ECOlogic in Washington, D.C., and deployed at the NASA Langley Research Center (LaRC) DAAC in Virginia. The QA DMS is based on Informix architecture and software. The primary users of the QA DMS are the MISR science team at the SCF and other facilities.

4.1 QA DMS FUNCTIONALITY

The implementation of the QA DMS has the following functionality:

(1) Ingest/Catalog Subsystem: The Ingest/Catalog Subsystem of the QA DMS will have an interface to the EOSDIS Core System (ECS) to facilitate the identification and location of the Summary QA files. The mechanism for this interface is through the ECS subscription service. Once the Summary QA files are located and delivered by ECS to a staging area accessible to the QA DMS, they will be ingested by the QA DMS and stored in an Informix database hosted at the LaRC DAAC.

(2) Search/Query/Delete GUI Subsystem: The Search/Query/Delete GUI Subsystem will provide users with a web-based interface to search for MISR data product names based on criteria included in the Summary QA files, to delete Summary QA files that are uninteresting or obsolete, and to perform ad-hoc queries using an open-query funciton.

5.0 ACKNOWLEDGMENTS

This research is being carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

6.0 REFERENCES

[1] Diner, D.J., C.J. Brugge, T. Deslis, V.G. Ford, L.E. Hovland, D.J. Preston, M.J. Shterenberg, E.B. Villegas, and M.L. White. "Development status of the EOS Multi-angle Imaging SpectroRadiometer (MISR)." *SPIE Proc.* 1939 (1993): 94-103.

[2] Kahn, R.A., Diner, D. J., Martonchik, J. V., Conel, J.E., Vargo, R. E., Wenkert, D., West, R.A., Bruegge, C. J., Abdou, W. A., Hansen, E. G., Paradise, S. R., Crean, K. A., Rheingans, B. E., and D. S. McDonald. "How Will We Choose Which Quality Flags and Constraints to Report for MISR Level 2 Data?", *The Earth Observer* (May/June 1995): 32-33.

[3] Kahn, R. A., Crean, K. A., Diner, D. J., Hansen, E. G., Martonchik, J. V., McMuldroch, S., Paradise, S. R., Vargo, R. E., and R. A. West. "Quality Assessment for MISR Level 2 Data." *The Earth Observer* (Jan./Feb. 1996): 19-21.

[4] HDF 4.1r2 User's Guide, NCSA HDF Development Group, July 21, 1998.