

MODIS Land Products Quality Assurance Tutorial: Part-1

How to find, understand, and use the quality assurance information for MODIS land products

NASA LP DAAC, USGS EROS Center, Sioux Falls, SD (March 1, 2012)

Introduction

All MODIS land products include quality assurance (QA) information designed to help users understand and make best use of the data that comprise each product. This Part-1 document contains material for beginners as well as intermediate MODIS land product users to help educate them on how to correctly use the QA information. It includes the following five sections:

- 1. A general description of MODIS land products and QA**
- 2. Why is it important for users to consult the QA information?**
- 3. QA metadata sources in MODIS land products**
- 4. Information on Land Data Operational Product Evaluation (LDOPE) tools**
- 5. Links to MODIS QA-specific online information resources**

Part-2 of this document (in the near future) will provide a detailed delineation and examples of the pixel-level QA structure in three MODIS land product suites: land surface reflectance, vegetation indices, and BRDF and albedo.

Section-1: MODIS land products and their QA information

The MODIS Adaptive Processing System (MODAPS) facility at the Goddard Spaceflight Center (GSFC) in Greenbelt, MD routinely produces MODIS land products from data derived from twin MODIS instruments aboard the Terra and Aqua platforms that were launched in December 1999 and May 2002 respectively. These data are archived at and distributed from the Land Processes Distributed Active Archive Center (LP DAAC) at the USGS EROS Center in Sioux Falls, SD. The MODIS Land Science Team (MODLAND) is responsible for the MODIS land products in terms of their QA and validation. They help evaluate and document the science quality of the products that are intended to constantly inform the user community. The Land Data Operational Product Evaluation (LDOPE) facility, collocated with MODAPS at GSFC, is responsible for the overall coordination of the QA

activities in support of the MODIS Science Team. A fairly complex and laborious process, this includes the evaluation and documentation of the science quality of all MODIS land products, which is finally incorporated in the operational production code and carried within the products (at the pixel-level) and their metadata (at the file-level).

The MODIS land collections comprise over sixty-five products that include dailies to n-day composites. Over the eleven years since Terra's launch, MODIS data at the LP DAAC comprise ~26 million granules and ~575 Terabytes in volume (as of April 30, 2011). Over this time, the NASA data discovery interfaces to search and procure data have also continued to evolve, and the current interface called "Reverb" is the third incarnation, which replaces the Warehouse Inventory Search Tool (WIST). A number of other MODIS data search, access, and procurement methods exist as well.

Given the large number of MODIS land products (they include 16 daily, 1 four-day, 29 eight-day, 8 sixteen-day, 7 monthly, and 5 yearly products), the dependencies that exist between them, and the differences in the QA procedures that are applied to them, it is difficult to provide a generic description and approach that applies to all. Within the user-data interaction process, MODIS QA-related information has the potential to manifest itself at different times and locations. One of the primary goals of this document is to direct users to the best combination of QA information sources, and methods to tap them to help drive the data requirements for their research and applications.

NASA's Earth Observing System (EOS) manages one of the largest science data production and applications enterprises in the world, of which MODIS datasets comprise a leading component. QA has always been identified as very essential to the success of the real-world applications that MODIS datasets help support but its complexity has discouraged widespread use. The mechanisms to generate, publish, access, communicate, and interpret QA for diverse suites of MODIS products are very elaborate. This document is intended to expose users to basic information that initiates their interaction with the QA layers of MODIS land products.

Section-2: Why is it important for users to consult the QA information?

MODIS QA information provides vital clues regarding the usability and usefulness of the data products for any particular science application. *Usability* is the capability of being used for a particular purpose while *usefulness* refers to what extent something serves a purpose towards meeting a practical objective. Usability and usefulness address any of the

following requirements that MODIS QA information provide, which are not mutually exclusive:

- Are sufficiently enough cloud-free data available to meet the requirements of a particular science application?
- Do sufficiently enough data meet the nominal output specifications as expected by the product's algorithm?
- What proportion of data artifacts and anomalies present in the data are deemed to exist within a satisfactory threshold to proceed with a particular science application?
- Are there mitigating conditions under which we can rule certain science data layers within a product as more or less useful than others?
- Are the science data layers derived using the main algorithm deemed satisfactory compared to the back-up algorithm or vice-versa?

Part-2 of this document, scheduled for release later, provides details regarding the QA structure and implementation within three MODIS product suites. Please note that the examples provided from these product suites demonstrate how the QA characterization propagates to the higher-level products, and underscores the need to understand the data fidelity at the very beginning stage when users contemplate use of a particular MODIS data product.

Given the fact that any particular MODIS land product is the result of a fairly complex process that involves a science algorithm, inputs that include MODIS level-1B data, ancillary data, lookup tables, auxiliary inputs, and possibly, other derived MODIS data products, users run a serious risk by not consulting the QA information. Some of the known sources of error that impact data quality include data loss due to instrument contact errors, striping in the land surface reflectance data, geolocation errors traceable to instrument maneuvers, effects of solar eclipse on the data, and problems stemming from the cloud mask, especially as a function of latitude. MODIS land data product users are therefore strongly encouraged to consult the QA information before they decide to use their data.

Section-3: QA information sources in MODIS land products

Typically, a MODIS HDF dataset contains several Science Data Sets (SDS), one or more QA data layers, and metadata. The QA data layers provide pixel-level QA for the science data,

and the metadata describe summary statistics of certain attributes, and also a statement about the product QA.

Three sources of QA information exist within MODIS land products that represent those that users should consult, and include the following:

- ★ **File-level metadata**
- ★ **Pixel-level metadata**
- ★ **LDOPE Web information**

File-level metadata: File-level QA refers to metadata that summarizes the data quality within that file. *Please note that this file-level assessment largely helps the search and discovery process, and users should not solely rely on them for their science data application needs.* They include the following:

1. Additional attribute metadata returned from a user search on a data discovery interface (this includes an overall percent quality, and a percent-based assessment of product-specific variables)
2. Granule-level QA Stats and QA Flags metadata returned from a user search on a data discovery interface (this includes percent-based assessments of cloud-cover, missing-, interpolated-, and out-of-bounds data)
3. The encapsulated metadata that exist in the header of the HDF file
4. The external xml metadata file

Essentially, these metadata sources provide the same information that users encounter at different stages of the data search and acquisition processes. The first two are generally designed to help as part of the user search and screening processes, especially as users look for good quality, cloud-free datasets. Numbers 3 and 4 refer to information from acquired products that contain the same metadata represented in 1 and 2. This document's major emphasis is on pixel-level metadata, discussed next.

Pixel-level metadata: QA metadata that reside at the pixel-level is most valuable for applications that rely on consistent use of particular MODIS land products. For instance, pixel-level QA metadata may help applications based on time-series analyses to ensure that their data inputs remain consistently of reliable quality. Two kinds of pixel-level metadata implementations exist in MODIS land products.

1. The first includes a QA SDS that contains multiple information sources accomplished through binary encoding.
2. The second involves a QA SDS that contains a single information source, such as *pixel reliability* in the Vegetation Indices products, or *albedo quality* in the BRDF/Albedo products.

All MODIS land products contain one or more SDS devoted to QA among the multiple HDF arrays. These SDSs are critical to understand, parse, and interpret pixel-level QA. As users open the MODIS HDF dataset in any image processing software system, the one or more QA-specific SDSs are identifiable through the inclusion of “QA”, “QC”, or “Quality” in their name. [Table 1](#) identifies the QA SDS arrays for each of the MODIS land products.

Radiation Budget Variables	Ecosystem Variables	Land Cover Characteristics
Land Surface Reflectance	Vegetation Indices	Thermal Anomalies & Fire
M*D09GA: 1 km Reflectance Data State QA 500 m Reflectance Band Quality	M*D13A1: 500 m 16 days Pixel Reliability QA 500 m 16 days VI Quality	M*D14: Algorithm QA
M*D09GQ: 250 m Reflectance Band Quality	M*D13A2: 1 km 16 days Pixel Reliability QA 1 km 16 days VI Quality	M*D14A1: QA
M*D09A1: 500 m Reflectance Band Quality 500 m State Flags	M*D13A3: 1 km Monthly Pixel Reliability QA 1 km Monthly VI Quality	M*D14A2: QA
M*D09Q1: 250 m Reflectance Band Quality	M*D13Q1: 250 m 16 days VI Quality 250 m 16 days Pixel Reliability QA	
M*D09CMG: Coarse Resolution QA Coarse Resolution State QA	M*D13C1: CMG 0.05° 16 days Pixel Reliability QA CMG 0.05° 16 days VI Quality	Land Cover
		MCD12Q1: Land Cover QC
	M*D13C3: CMG 0.05° Monthly Pixel Reliability QA CMG 0.05° Monthly VI Quality	MCD12Q2: Dynamics_QC
Land Surface Temp. & Emissivity		MCD12C1: Majority_Land_Cover_Type_1_QC
M*D11_L2: Daytime LST QC		
M*D11A1: Daytime LST QC Nighttime LS QC	LAI and FPAR	
	M*D15A2: FparLai_QC FparExtra_QC	
M*D11A2: QC for Daytime LST QC for Nighttime LST	MCD15A3: FparLai_QC FparExtra_QC	VCC and VCF
		MOD44A.004: Labeled LC Change Past 1 Year
M*D11B1: Daytime LST QC Nighttime LST QC QC for Retrieved Emissivities		MOD44B.005: Quality
M*D11C1: Daytime LST QC Nighttime LST QC QC for Retrieved Emissivities	GPP and NPP	Land Water Mask
	M*D17A2: PSN_QC_1km	MOD44W: Water Mask QA
M*D11C2: Daytime LST QC Nighttime LST QC	M*D17A3: NPP_QC_1km	
M*D11C3: Daytime LST QC Nighttime LST QC		
BRDF and Albedo		<i>An asterisk refers to both Terra & Aqua versions of the MODIS product. MCD refers to a "combined" product generated with Terra & Aqua MODIS inputs.</i>
MCD43A2 and MCD43B2: BRDF_Albedo_Quality Snow_BRDF_Albedo BRDF_Albedo_Ancillary BRDF_Albedo_Band_Quality		
MCD43C1, C2, C3, & C4: BRDF_Albedo_Quality		<i>The SDS names are either shortened or abbreviated, and may not exactly conform to those within the datasets.</i>

Table 1: QA Science Data Sets within the HDF arrays for each MODIS land product suite

Pixel-level QA is generated by the production code to evaluate product quality at the finest granularity. This information is useful to both users who need to decide the quality of retrieval at any particular pixel, and for LDOPE’s evaluation as well. LDOPE performs detailed analyses before decisions regarding this assessment are integrated into the final production code. Pixel-level QA provides information for each science parameter through the following two methods:

1. A binary representation of bit* combinations that characterize particular quality attributes.
2. An entire QA word may represent one summary quality information source such as pixel reliability in VI, and Albedo quality in BRDF/Albedo

Pixel-level QA varies between products and their levels. In general, there are two types of pixel-level QA metadata provided.

1. MODLAND-wide QA: The first is the MODLAND-wide QA bits that provide 1 or 2 generic QA (least-significant**) bits for each pixel of every product. Its purpose is to provide a consistent quality interpretation across all MODIS land products. Prior to Collection-5, a 2-bit QA was used to describe four potential conditions. [Table 2](#) describes the codes and their interpretation for this product collection-wide QA.

Pixel-level QA code	Interpretation
00	Pixel produced, good quality, not necessary to examine more detailed QA
01	Pixel produced, unreliable or unquantifiable quality, recommend examination of more detailed QA
10	Pixel not produced due to cloud effects
11	Pixel not produced primarily due to reasons other than cloud

Table 2: Pixel-level QA across all MODIS land products through Collection-4

Starting with Collection-5, some MODLAND products were implemented with a 1-bit generic product assessment ([Table 3](#)) rather than a 2-bit summary ([Table 2](#)) in part to reflect algorithm evolution. Land surface reflectance, land surface temperature &

* A bit (short for Binary Digit) is the smallest unit of information/memory in digital computing that can represent two possible values, represented by 0 and 1.

** The least-significant bit is the lowest bit in a series of numbers in binary notation, located at the far right of a string; also referred to as the right-most bit.

emissivity, and vegetation indices retained the 2-bit MODLAND QA as it continues to remain relevant to those products. The second bit describes the condition in which a pixel was not produced because of clouds or other effects. [Table 3](#) describes the 1-bit codes implemented in Collection-5.

Pixel-level QA code	Interpretation
0	Pixel produced, good quality, not necessary to examine more detailed QA
1	Other quality (produced or not produced; if produced unreliable or unquantifiable quality, examination of more detailed QA is recommended)

Table 3: Pixel-level QA across all MODIS land products implemented in Collection-5

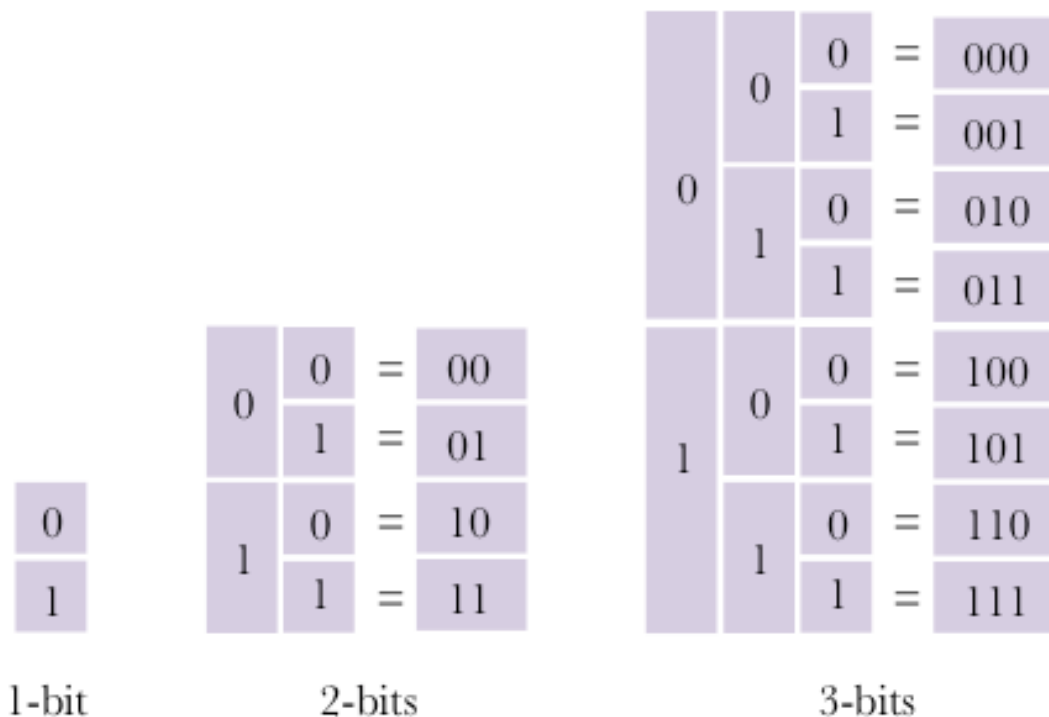
2. Product-specific QA: The second type of pixel-level QA addresses product-specific attributes. This metadata may address a variety of characteristic conditions that constitute a product’s elements. For instance, “products that can have meaningful error estimates assigned to them store per-pixel uncertainty estimates and/or ranges: for example, the land surface temperature product stores emissivity and temperature error estimates. Information on external factors known to affect product quality and consistency is also stored for each product. These data include atmospheric conditions (e.g., cloud cover); surface type (e.g., ocean, coast, wetland, inland water); scan, solar and viewing geometry; and whether dynamic ancillary data or backup estimates have been used as input (e.g., aerosol climatology estimates used to replace missing observations in the MODIS aerosol product)” (Roy et al., 2002).

Given the variety of MODIS land products, we cannot elaborate on all the different pixel-level QA attributes in this document. Links to relevant sources that offer both distilled versions of this information as well as complete product file specifications maintained by the MODIS Science Team are provided in [Section-5](#). Users are strongly encouraged to consult these sources to better understand and interact with their particular MODIS land products.

Generic description of the MODIS QA binary bits and bit-fields

Users often encounter problems with interpreting the binary encoded bitmap that represents useful product-specific QA metadata. This section provides a generic introduction to the basic elements of binary notation used to represent the pixel-level QA metadata.

The simplest way to parse the QA bits is by understanding the binary notation (built on base-2 rather than base-10) used to represent the values. A single bit represents two values (0, 1), while two bits represent four values, and three bits represent eight values, etc. as depicted below.



Hence, the number of bits and the number of values (or quality attribute meanings) they potentially represent, doubles with each step as shown below:

Number of Bits	Number of Values	Formula
1	2	2^1
2	4	2^2
3	8	2^3
4	16	2^4
5	32	2^5
6	64	2^6
7	128	2^7
8	256	2^8

The QA-specific SDS for each MODIS land product (listed in [Table 1](#)) generally breaks down into four columns. The order of these columns as depicted under the “Layers” tab in the product documentation on the LP DAAC Web page is described below. This order and the terminology used may vary slightly in the File Specifications and/or the User Guides, which are referenced in [Section-5](#):

- ★ The first column identifies the “Bit Number(s)”
- ★ The second column identifies the parameter “Bit-field name”
- ★ The third column identifies the bit values for each parameter
- ★ The fourth column provides the description of the bit-field

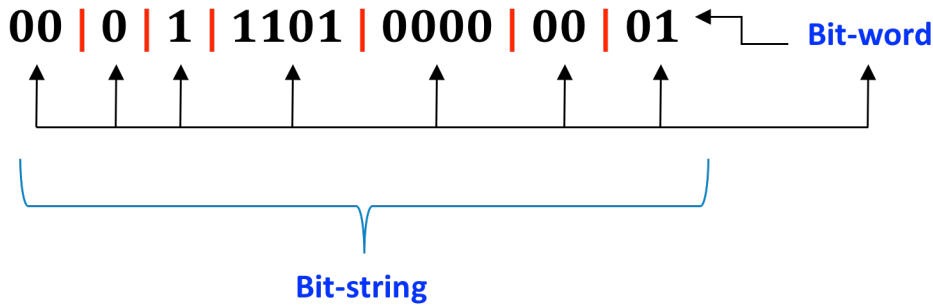
Users need to convert the pixel-level QA values from decimal to binary before they can parse and interpret them. Several online converters are available to make such a conversion, or users may convert* the desired QA bits to their binary format using LDOPE’s “unpack_sds_bits” utility (further details regarding LDOPE tools are available in [Section-4](#)). Once you have converted a pixel-level value from decimal to binary, you need to correctly parse the bits before you can interpret them.

The following four examples demonstrate how you handle a particular pixel-level QA value from four different products.

Example-1 (MOD09GQ)

A single pixel’s value of “**7425**” is derived from a 250 meters surface reflectance (MOD09GQ) product’s “QC_250m_1” SDS parameter. The **7425** decimal value converts to a binary value of **111010000001**. (Users should check for the *datatype*, which provides the number of bits in the QA word. The conversion should contain that many number of binary positions as defined in the number of bits). We need to add three zeros to the most-significant bit to complete the 16-bit string (i.e., to the left). This value, as assigned to the individual bit numbers, breaks down thus (based on the QA index specified in the MOD09GQ product’s file specifications):

*Users are reminded that all HDF-EOS products are written in the big-endian referencing scheme. The bits are always numbered from right (least-significant bit) to left (most-significant bit).



Please bear in mind that the binary bit-string is parsed from **right to left**, and the individual bits within a bit-word are read from **left to right**; the above string breaks down thus:

Bit Word	Description
01	Less than ideal quality some or all bands
00	Clear (Cloud state)
0000	Highest quality (Band-1 data quality)
1101	Correction out of bounds pixel constrained to extreme allowable value
1	Yes (Atmospheric correction performed)
0	No (Adjacency correction performed)
00	Spare (unused)

Example-2 (MOD09CMG)

The second example demonstrates how to parse a pixel-level value of **1075576832** from a single Coarse Resolution QA SDS of a land surface reflectance CMG (MOD09CMG) product:

The decimal value **1075576832** converts to the following 32-bit binary string that is separated into its bit-word components: **10000000001110000000000000000000**. A single zero is added to the most-significant bit to complete the 32-bit string (i.e., to the left).

0 | 1 | 0000 | 0000 | 0111 | 0000 | 0000 | 0000 | 0000 | 00

Reading from right to left, each bit-word is interpreted as follows:

Bit Word	Description
00	Correct product produced at ideal quality for all bands
0000	Highest quality (Band-1 data quality)
0000	Highest quality (Band-2 data quality)
0000	Highest quality (Band-3 data quality)
0000	Highest quality (Band-4 data quality)
0111	Noisy detector (Band-5 data quality)
0000	Highest quality (Band-6 data quality)
0000	Highest quality (Band-7 data quality)
1	Yes (Atmospheric correction performed)
0	No (Adjacency correction performed)

Example-3 (MCD43A2)

The third example is derived from a 500 meters, BRDF-Albedo Quality product (MCD43A2), which provides quality-specific information. The pixel value used in this example is derived from “BRDF_Albedo_Band_Quality,” which is one of four QA SDSs, a 32-bit unsigned integer. The pixel’s decimal value **70464307** converts to a binary value of **100001100110011001100110011**. Since this string does not account for the leading zeros, we need to add five zeros to the most-significant bit (i.e., to the left): **00000100001100110011001100110011**. This binary string breaks down into the following bit-words:

0 | 000 | 0100 | 0011 | 0011 | 0011 | 0011 | 0011 | 0011

This QA SDS is unique in that the above values are not directly interpreted for the first seven 4-bit values. The dataset attributes for the MCD43A2 product provide the decimal values that equate to certain binary strings. For instance, the “Band 1 Quality” attributes provide the following values:

00–03 Band 1 Quality

0 = Best quality, full inversion (WoDs, RMSE majority good)

1 = Good quality, full inversion

2 = Magnitude inversion (numobs >= 7)

3 = Magnitude inversion (numobs >=3&<7)

4 = Fill value

The individual bit-words, delineated above, are parsed again to derive their decimal values. The final interpretation yields the following information:

QA bit value	Description
0011 = 3	Magnitude inversion (numobs >= 3 & < 7) (Band-1 data quality)
0011 = 3	Magnitude inversion (numobs >= 3 & < 7) (Band-2 data quality)
0011 = 3	Magnitude inversion (numobs >= 3 & < 7) (Band-3 data quality)
0011 = 3	Magnitude inversion (numobs >= 3 & < 7) (Band-4 data quality)
0011 = 3	Magnitude inversion (numobs >= 3 & < 7) (Band-5 data quality)
0011 = 3	Magnitude inversion (numobs >= 3 & < 7) (Band-6 data quality)
0100 = 4	Fill value (Band-7 data quality)
000	Unassigned – TBD
0	Not fill-value (QAFill)

Example-4 (MOD15A3)

The final example is a pixel from a 4-day LAI-FPAR (MCD15A3) product. The decimal value of 107 from the FparLai QC SDS converts to a binary value of **1101011** which, following addition of a single zero to complete the 8-bit string, produces the following bit-words:

011 | 01 | 0 | 1 | 1

The above bit-words are interpreted to reveal the following:

QA bit value	Description
1	Other quality (Back-up algorithm or fill-values) (MODLAND QA)
1	Aqua (Sensor)
0	Detectors apparently fine for up to 50% of channels 1, 2
01	Significant clouds were present (Cloud State)
011	Main algorithm failed due to problems other than geometry, empirical algorithm used (Science Computing Facility QC)

LDOPE Web information: Refer to [Section-5](#)

Section-4: Land Data Operational Product Evaluation (LDOPE) QA tools

LDOPE develops and maintains a number of software tools designed to manipulate, visualize, and analyze MODIS data. A subset of LDOPE QA tools is available to the user community to help parse and interpret the QA SDS layers. These tools, numbering about two dozen, are provided as source code and command-line executables that run on a limited number of operating systems (Linux, IRIX, Solaris, and Windows). Written in C, they are executed either from the command-line or invoked via scripts. One of the most commonly used routines (referred in the earlier section) is “unpack_sds_bits” which helps parse and interpret the bit-packed QA attributes.

Interested users may register and download these LDOPE QA tools and related documentation from the following site:

https://lpdaac.usgs.gov/lpdaac/tools/ldope_tools

Section-5: Links to QA-specific online information sources

This section provides links to all QA-related online information that users should find useful.

LDOPE Web site: http://landweb.nascom.nasa.gov/cgi-bin/QA_WWW/newPage.cgi

LDOPE maintains the “MODIS Land Quality Assessment” Web site that provides a plethora of QA-specific as well as other information. Under the “Quality” drop-down menu, there are “Terra C5”, “Aqua C5”, and “Terra+Aqua C5” hyperlinks that lead to product-specific QA documentation. This includes Science Quality Flag values, their explanation, and related comments for each product’s discrete acquisition time ranges. Hyperlinks also exist for “Terra Known Issues”, “Aqua Known Issues”, and “Terra+Aqua Known Issues.” They provide detailed descriptions of past problems and current issues under investigation for each MODIS land product.

The “Docs” drop-down menu contains hyperlinks to the following resources for each product suite:

- ★ **User Guides**
- ★ **Algorithm Theoretical Basis Documents**
- ★ **File Specifications** (*provides QA bit descriptions*)
- ★ **List of the Earth Science Data Types (ESDT) and Science Data Sets**
- ★ **List of MODIS product interdependencies**

A number of other links germane to MODIS land data quality assessment exists as well.

LP DAAC Web site: https://lpdaac.usgs.gov/lpdaac/products/modis_products_table

The above link to the LP DAAC’s Web site provides a table of all MODIS land products, each of which leads the user to product-specific documentation. QA-specific attributes and description information are provided under the “Layers” tab. This information is extracted from the “File Specifications” for each product. The “Links” tab provides links to the product-specific User Guide, ATBD, and the MODIS Validation Web site.

References

Roy, D.P., Borak, J.S., Devadiga, S., Wolfe, R.E., Zheng, M., Desloîtres, J. (2002) The MODIS Land Quality Assessment Approach. *Remote Sensing of Environment*, 83: 62–76.