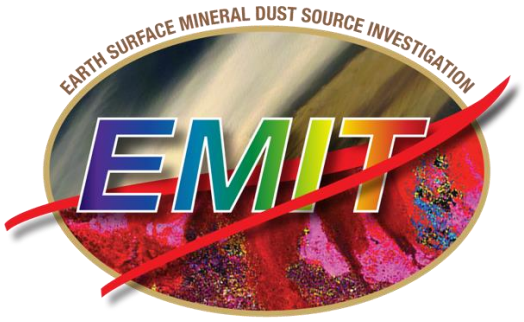

Earth Surface Mineral Dust Source Investigation (EMIT)

Level 2A Data Product User Guide

Initial Release

Paper copies of this document may not be current and should not be relied on for official purposes. The current version is V1.0.

May 9, 2022
JPL D- 107862

Jet Propulsion Laboratory
California Institute of Technology

Signature Page

Prepared by:

_____ _____
Nimrod Carmon Date
EMIT SDS Team Member

_____ _____
Philip Brodrick Date
EMIT SDS Algorithm Development Lead

Approved by:

_____ _____
Sarah Lundeen Date
EMIT SDS Manager

Change Log

Date	Sections Changed	Reason for Change	Revision
05/09/2022	All	All	Initial Release
11/18/2022	All	Updates from In Orbit Checkout	V1.0

EPDM Electronic Signatures

Snapshot of signatures from EPDM will be added upon release

Contacts

Readers seeking additional information about this product may contact the following:

Philip G. Brodrick
philip.brodrick@jpl.nasa.gov

David R. Thompson
david.r.thompson@jpl.nasa.gov

Nimrod Carmon
nimrod.carmon@jpl.nasa.gov

Table of Contents

1 Introduction

1.1 Identification

This document describes information about the file structure and datasets provided in the EMIT Level 2A data product. The algorithms and data content of the Level 2A data products are described briefly in this guide, with the purpose of providing the user with sufficient information about the content and structure of the data files to enable the user to access and use the data. For a more complete description of the algorithms used, please see the L2A algorithm theoretical basis document (ATBD).

1.2 Overview

The EMIT Project will deliver space-based measurements of surface mineralogy of the Earth's arid dust source regions. These measurements are used to initialize the compositional makeup of dust sources in Earth System Models (ESMs). The dust cycle, which describe the generation, lofting, transport, and deposition of mineral dust, plays an important role in ESMs. Dust composition is presently the largest uncertainty factor in quantifying the magnitude of aerosol direct radiative forcing. By understanding the composition of mineral dust sources, EMIT aims to constrain the sign and magnitude of dust-related radiative forcing at regional and global scales. During its one-year mission on the International Space Station (ISS), EMIT will make measurements over the sunlit Earth's dust source regions that fall within ±52° latitude. EMIT will schedule up to five visits (three on average) of each arid target region and will only downlink acquisitions not dominated by cloud cover. EMIT-based maps of the relative abundance of source minerals will advance the understanding of the current and future impacts of mineral dust in the Earth system.

The EMIT instrument is a Dyson imaging spectrometer that uses contiguous spectroscopic measurements in the visible to short wave infrared region of the spectrum to resolve absorption features of dust-forming minerals. From the instrument's focal plane array, on-board avionics read out raw detector counts at 1.6 Gbps, and then digitizes and stores these data to a high-speed Solid-State Data Recorder (SSDR). From there, the avionics software reads the raw uncompressed data, packages it into frames of 32 instrument lines, screens for cloudy pixels within the frames, and performs a lossless 4:1 compression of the science data before storing the processed, compressed data back onto the SSDR. The data is later read from the SSDR, wrapped in Consultative Committee for Space Data Systems (CCSDS) packets and then formatted as Ethernet packets for transmission over the ISS network and downlinked to the EMIT Instrument Operation System (IOS). Once on the ground, the EMIT IOS delivers the raw Ethernet data to the SDS where Level 0 processing removes the Huntsville Operations and Support Center (HOSC) Ethernet headers, groups CCSDS packet streams by Application Identifier (APID), and sorts them by course and fine time.

The Level 2A data product contains estimated surface reflectance, uncertainty, and mask data. In addition, the geolocation of all pixel centers is included as well as the calculation of observation geometry and illumination angles on a pixel-by-pixel basis. Each image line of the Level 2A data product is also UTC time-tagged.

All EMIT products are provided as Network Common Data Form (NetCDF) files (further information can be found at https://www.unidata.ucar.edu/software/netcdf/).

1.3 File Formats

1.3.1 1.3.1. Metadata Structure

\ EMIT operates from the ISS, orbiting Earth approximately 16 times in a 24-hour period. EMIT starts and stops data recording based on a surface coverage acquisition mask. The top-level metadata identifier for EMIT data is an orbit, representing a single rotation of the ISS around Earth. A period of continuous data acquisition within an orbit is called an orbit segment, where each orbit segment can cover up to thousands of kilometers down-track depending on the acquisition mask map. Each orbit segment is broken into scenes of 1280 down-track lines for convenience, though scenes may be seamlessly reassembled into orbit segments. To prevent a very small number of lines in any scene, the last scene can extend up to 2559 lines.

EMIT scenes are processed to different data levels. The bundle of data and metadata for a given scene at a given level is referred to as a granule. Each EMIT granule contains one or more NetCDF files as well as a browse PNG file. The EMIT EMITL2ARFL Reflectance collections contain a three NetCDF files, Reflectance, Reflectance Uncertainty, and Mask, and a quick look PNG file (Browse):

Table 1-1: EMITL2ARFL collection file list and naming convention

Collection: EMITL1BRFL
Reflectance: EMIT_L2A_RFL_<VVV>_<YYYYMMDDTHHMMSS>_<OOOOOOO>_<SSS>.nc
Reflectance Uncertainty: EMIT_L2A_RFLUNCERT _<VVV>_<YYYYMMDDTHHMMSS>_<OOOOOOO>_<SSS>.nc
Mask: EMIT_L2A_MASK_<VVV>_<YYYYMMDDTHHMMSS>_<OOOOOOO>_<SSS>.nc
Browse: EMIT_L2A_RFL_<VVV>_<YYYYMMDDTHHMMSS>_<OOOOOOO>_<SSS>.png

<YYYYMMDDTHHMMSS> is the UTC timestamp of the first line of the scene, e.g., 20220101T083015
<OOOOOOO> is the orbit identification number represented as <YY><DOY>NN> where YY = Year, DOY = Day of Year, NN = Incrementing number for each day (two digits)
<SSS> is the scene identification number, e.g., 007. Within an orbit, scene numbers begin at 001 and increment by 1 for each new scene.

1.3.2 L2A Data Products

The EMIT L2A Reflectance collection contains surface reflectance data in spatially raw, non-orthocorrected, format. L2A reflectance match the orientation of L1B radiance, which have been 'flipped' in both the spectral and spatial dimensions of the focal plane array, so that images appear as they would on the Earth's surface, and wavelengths start from lower to higher with increasing channel/band number. EMIT Level 2A Reflectance collection (EMITL2ARFL) contains three separate NetCDF files: Reflectance (EMIT_L2A_RFL), Reflectance Uncertainty (EMIT_L2A_RFLUNCERT), and Reflectance Mask (EMIT_L2A_MASK). The reflectance NetCDF file contains surface reflectance maps and the reflectance uncertainty NetCDF file contains uncertainty estimates about the reflectance captured as per-pixel, per-band, posterior standard deviations. The mask NetCDF file contains atmospheric state estimates and binary flags described later in this document. All NetCDF files in this collection contain a location group storing a geometric lookup table (GLT), an orthocorrected image that provides relative x and y reference locations from the raw scene to facilitate fast projection, as well as latitude, longitude and elevation datasets.

Table 1-1: EMIT L2A Data Products Summary

Earth Science Data Type	Product Level	Data Dimension	Spatial Resolution	Swath Width	Map Projection
Collection	*EMITL2ARFL*				
Reflectance	L2	• 1242 cross-track • 1280** down-track • 285 bands	60 m*	75 km*	Non-orthocorrected, latitude and longitude tagged (WGS-84)
Reflectance Uncertainty	L2	• 1242 cross-track • 1280** down-track • 285 bands	60 m*	75 km*	Non-orthocorrected, latitude and longitude tagged (WGS-84)
Mask	L2	• 1242 cross-track • 1280** down-track • 8 bands primary • 32 bands	60 m*	75 km*	Non-orthocorrected, latitude and longitude tagged (WGS-84)

* Nominal at equator

** Images have at least 1280 rows. To prevent small scenes, scenes at the end of an orbit segment will have up to 2559 rows.

1.4 Product Availability

The EMIT L1B products will be available at the NASA Land Processes Distributed Active Archive Center (LP DAAC, https://lpdaac.usgs.gov/) and through NASA Earthdata (https://earthdata.nasa.gov/).

2 Reflectance Retrieval (Atmospheric Correction)

EMIT's Level 2A is a surface reflectance product, providing the hemispherical-directional reflectance factor (HDRF) for every pixel in the scene. Reflectance is estimated from at-sensor radiance (Level 1B) using an optimal estimation (OE) based atmospheric correction procedure, fully described in the Level 2A ATBD. The OE algorithm produces two maps for each pixel; surface reflectance, and reflectance uncertainty. The reflectance uncertainty map provides the diagonal elements of the posterior covariance matrix, square-rooted, to provide a spectrum of uncertainty about the reflectance estimate in standard deviations units. Together, these two products define the posterior probability of the surface reflectance given the at-sensor radiance measurement, captured as a multivariate normal distribution (*Reference Pending Peer Review*). Uncertainty-aware downstream analysis of the reflectance map can leverage both products, using the reflectance uncertainty as error bars over the reflectance estimate.

3 File Structure

3.1 Dimensions and Global Attributes

The NetCDF files contain metadata information describing the dimensions of the datasets (down-track, cross-track, bands, ortho_y, and ortho_x), where down-track and cross-track describe the dimensions of the non-orthocorrected datasets, and ortho_y and ortho_x describe the dimensions of the orthocorrected (EPSG:4326 projected) GLT bands.

Each NetCDF file contains a list of global attributes describing mission relevant information.

3.2 Reflectance

The L2A Reflectance product provides surface reflectance maps in raw spatial format (unorthocorrected). Reflectance values are estimated using an optimal estimation (OE) based atmospheric correction algorithm (see L2A ATBD for details). The nodata value is -9999, and deep atmospheri water features – where reflectance is not estimated – are set to -0.01.

Table 3-1: L2A Reflectance File Structure and Content

Field Name	Type	Units	Comments
Group	***Root***		

Reflectance	float32, Multiband array (1280*x1242x285)	Spectral Reflectance	
Group	***location***		
GLT-X	int32	Index	
GLT-Y	int32	Index	
Lat	float64	Decimal Degree	
Lon	float64	Decimal Degree	
Elevation	float64	Meters	
Group	***/instrument_band_parameters***		
Wavelength	float32	nm	Wavelength grid
FWHM	float32	nm	Full width at half maximum

* Scenes have at least 1280 rows. To prevent small scenes, scenes at the end of an orbit segment will have up to 2559 rows.

3.3 Reflectance Uncertainty

The L2A Reflectance Uncertainty product provides estimated uncertainty maps in raw spatial format (unorthocorrected). Reflectance uncertainty values are given in standard deviation units and are estimated using an optimal estimation (OE) based atmospheric correction algorithm (See chapter 2).

Table 3-2: L2A Reflectance Uncertainty File Structure and Content

Field Name	Type	Units	Comments
Group	***root***		
Reflectance Uncertainty	Multiband array (1280*x1242x285)	Standard Deviations	
Group	***location***		
GLT-X	int32	Index	
GLT-Y	int32	Index	

Lat	float64	Decimal Degree	
Lon	float64	Decimal Degree	
Elevation	float64	Meters	
Group	***/instrument_band_parameters***		
Wavelength	float32	nm	Wavelength grid
FWHM	float32	nm	Full width at half maximum

* Scenes have at least 1280 rows. To prevent small scenes, scenes at the end of an orbit segment will have up to 2559 rows.

3.4 Masks

The Mask Bands dataset contains 6 binary flag bands (1-5, 8), and two data bands (6,7). The mask flag bands identify features which should be excluded from the analysis. More information is available in the EMIT Level 2A ATBD sections 4.2.5 and 4.2.6. The two data bands hold estimated atmospheric state values for aerosol optical depth (band 6), and water vapor (band 7). These maps are given in raw spatial, non-orthocorrected, format, and are provided in super-pixel segments (full description in EMIT Level 2A ATBD section 4.2.3). **Note that the Mask variable dataset is provided as a multiband array on the root group level without individual band labels.** The labels are provided in the instrument band parameters group.

A second provided field – band_mask - indicates whether or not any given wavelength of any given pixel is interpolated. Interpolation occurs either due to a focal plane array bad pixel, or from saturation. Data is provided as a packed unsigned integer array with 36 elements

Table 3-3: L2A Mask File Structure and Content

Field Name	Type	Units	Comments
Group	***Root***		
mask	float32		
band_mask	uint8		
Group	***location***		
GLT-X	int32	Index	

GLT-Y	int32	Index	
Lat	float64	Decimal Degree	
Lon	float64	Decimal Degree	
Elevation	float64	Meters	
Group	***/instrument_band_parameters***		
mask_band	str Array	Labels	Array of strings indicating the name of each mask band

The contents of each mask band are shown in the table below.

Table 3-4 Mask Bands

IDX	Band Name	Band Description
1	Cloud Flag	Cloud Coverage
2	Cirrus Flag	Dense Cirrus clouds
3	Water Flag	Water bodies
4	Spacecraft Flag	Spacecraft or space station components that intersect the EMIT field of view
5	Dilated Cloud Flag	Cloud coverage + buffer
6	AOD550 (unitless)	AOD at 550nm estimates
7	H2O (g cm^{-2})	Water Vapor estimates
8	Aggregate Flag	Aggregated binary flag of bands 1 through 5

3.4.1 Geometry Lookup Table (GLT)

EMIT's radiance and observation data are provided in non-orthocorrected spatially raw format (termed 'instrument'). To conveniently project the instrument data onto a gridded geographical map, a geometry lookup table (GLT) is provided. The GLT dataset is an orthocorrected product with a fixed pixel size projected into a North-up WGS-84 system that contains the information about which original pixel occupies which output pixel in the final product. The GLT file contains two parameters – instrument cross-track (e.g. column, sample, or x) index, and instrument down-track (e.g. row, line, or y) index. These two numbers provide the one-based relative reference from

the upper left corner of the instrument-space dataset. The GLT uses one-based indexing, and 0 is the nodata value.

Table 3-5: GLT File Structure and Content

Field Name	Type	Units	Comments
GLT cross-track lookup	int32	Index	Cross-track is the column, or x direction
GLT down-track lookup	int32	Index	Down-track is the row, or y direction

3.4.2 Location

EMIT data product files contain location information in the orientation they were collected (without orthocorrection) in order to preserve the maximum information content. However, for user convenience, products are 'flipped' from the orientation in which they were acquired from the focal plane array in order to match the spatial and spectral orientation of the ground (this means an image rotation and translation would be sufficient to approximate the true ground location, though a a more sophisticated process is used and outlined in the Level 1B ATBD). The following location data are provided: latitude, longitude, elevation. The latitude and longitude coordinates are given in EPSG:4326 (coordinates on the WGS-84 ellipsoid). The longitude values are Easting (values increasing Eastward from Greenwich). The elevation dataset is sourced from the Shuttle Radar Topography Mission (SRTM v3, void filled with ASTER v2), and is resampled to EMIT's spatial resolution.

Table 3-6: Location File Structure and Content

Field Name	Type	Units	Comments
Latitude	BINARY 32-bit signed long integer	Degree	Degrees increasing Eastward from Greenwich
Longitude	BINARY 32-bit signed long integer	Degree	
Elevation	BINARY 32-bit signed long integer	m	Estimated ground elevation at pixel center

4 References

Appendix A: Acronyms

Term	Definition
ADC	Analog to Digital Converter
APID	Application Identifier
ASCII	American Standard Code for Information Interchange
BIL	Band Interleaved by Line
CCSDS	Consultative Committee for Space Data Systems
DAAC	Distributed Active Archive Center
DCID	Data Collection Identifier
DN	Digital Number
EMIT	Earth Mineral dust source InvesTigation
ENVI	Environment for Visualizing Images
ESDIS	Earth Science Data and Information System
ESM	Earth System Model
FPA	Focal Plane Array
FPGA	Field Programmable Gate Array
FPIE	Focal Plane Interface Electronics
FPIE-A	Focal Plane Interface Electronics - Analog
FSW	Flight Software
Gbps	Gigabits per second
GLT	Geometry Lookup Table
HOSC	Huntsville Operations and Support Center
ICD	Interface Control Document
IOS	Instrument Operations System
ISS	International Space Station
JPL	Jet Propulsion Laboratory
kHz	Kilohertz
L0	Level 0 (compressed, raw packets)
L1A	Level 1A (reconstructed, uncompressed data reassembled into scenes)
L1B	Level 1B (calibrated radiances with geolocation parameters)
L2A	Level 2A (atmospherically-corrected surface reflectance)
L2B	Level 2B (mineral feature depth maps)
L3	Level 3 (gridded global map of mineral composition and abundances)
L4	Level 4 (model runs of GISS ModelE2 and NCAR CESM)
LP DAAC	Land Processes Distributed Active Archive Center
LSB	Least Significant Bit
MSB	Most Significant Bit
NASA	National Aeronautics and Space Administration
NetCDF	Network Common Data Format
PGE	Product Generation Executable
PLRA	Program Level Requirements Appendix
ROIC	Readout Integrated Circuit
SDS	Science Data System
SIS	Software Interface Specification

SSDR	Solid State Data Recorder
UTC	Universal Time Coordinated