

JPL JPL D-104255
EMIT DRL XXX
ATBD-EMIT-02a

Earth Mineral dust source Investigation (EMIT)

EMIT L2A Algorithm: Surface Reflectance and Scene Content Masks

Theoretical Basis

David R. Thompson, Philip G. Brodrick, Robert O. Green, Olga Kalashnikova, Sarah Lundeen, Gregory Okin¹, Winston Olson-Duvall, Thomas Painter¹

Jet Propulsion Laboratory, California Institute of Technology

¹University of California, Los Angeles

Version 1.0

April 2020

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California 91109-8099



Change Log

Version	Date	Comments
0.1	Sept. 15, 2019	Initial Draft
0.2	Sept. 29, 2019	Candidate Science Peer Review Version
0.3	Dec. 6, 2019	Science Team Meeting at the Jet Propulsion Laboratory
0.4	Dec. 17, 2019	Changes attending to peer review and NASA template
0.5	Jan 2, 2020	Changes post-PDR
0.6	Jan 25, 2020	Added output products
0.7	Jan 30, 2020	Cloud mask dilation
0.8	Apr 7, 2020	Code repository
0.9	Apr 8, 2020	Input and output are no longer orthorectified
1.0	Apr 10, 2020	Pre-CDR Science Peer Review. Title Change.

TABLE OF CONTENTS

1. KEY TEAMMEMBERS 1

2. HISTORICAL CONTEXT AND BACKGROUND ON THE EMIT MISSION AND ITS INSTRUMENTATION 2

3 ALGORITHM RATIONALE AND PRIOR VALIDATION 4

4. ALGORITHM DESCRIPTION 5

4.1 Input data 5

4.2 Theoretical description 6

4.2.1 Radiative Transfer and Atmospheric Modeling 7

4.2.2 Model Inversion 9

4.2.3 Superpixel Segmentation 10

4.2.4 Empirical Line extrapolation 11

4.2.5 Cloud Masking 11

4.3 Practical Considerations 13

5. OUTPUT DATA 14

6. CALIBRATION, VALIDATION, AND FIELD MEASUREMENT 15

7. CONSTRAINTS AND LIMITATIONS 18

REFERENCES.....ERROR! BOOKMARK NOT DEFINED.

1. Key Teammembers

A large number of individuals contributed to the development of the algorithms, methods, and implementation of the L1b approach for EMIT. The primary contributors are the following:

- David R. Thompson (Jet Propulsion Laboratory) – EMIT Co-I, Instrument Scientist
- Robert O. Green (Jet Propulsion Laboratory) – Mission PI, Radiometric modeling
- Tom Painter (Jet Propulsion Laboratory) – Surface reflectance and BRDF
- Olga Kalashnikova (Jet Propulsion Laboratory) – Atmospheric Aersols
- Sarah Lundeen (Jet Propulsion Laboratory) – Science Data System Lead
- Randy Pollock (Jet Propulsion Laboratory) – Instrument Systems Engineer
- Philip Brodrick (Jet Propulsion Laboratory) – Algorithms Design and Implementation

In addition, the algorithms described are based on prior work that includes sponsorship by multiple research agencies and includes contributions by many individuals. These are associated with the papers and manuscripts listed throughout this text, and provided in references under the relevant topics.

2. Historical Context and Background on the EMIT Mission and its Instrumentation

Mineral dust radiative forcing is the single largest uncertainty in aerosol direct radiative forcing (USGCRP and IPCC). Mineral dust is a principal contributor to direct radiative forcing over arid regions, impacting agriculture, precipitation, and desert encroachment around the globe. However, we have poor understanding of this effect due to uncertainties in the dust composition. Dust radiative forcing is highly dependent on its mineral-specific absorption properties, and the current range of iron oxide abundance in dust source models (0 – 7 wt%) translates into a 460% uncertainty in regional radiative forcing predicted by Earth System Models (ESMs). Meanwhile, soil samples from North Africa regions - important sources of mineral dust - contain up to 30 wt% iron oxide. The National Aeronautics and Space Administration (NASA) recently selected the Earth Mineral Dust Source Investigation (EMIT) to close this knowledge gap. EMIT will launch an instrument to the International Space Station (ISS) to directly measure and map the mineral composition of critical dust-forming regions worldwide.

The EMIT Mission will use imaging spectroscopy across the visible shortwave (VSWIR) range to reveal distinctive mineral signatures, enabling rigorous mineral detection, quantification, and mapping. The overall investigation aims to achieve two objectives.

1. Constrain the sign and magnitude of dust-related RF at regional and global scales. EMIT achieves this objective by acquiring, validating and delivering updates of surface mineralogy used to initialize ESMs.
2. Predict the increase or decrease of available dust sources under future climate scenarios. EMIT achieves this objective by initializing ESM forecast models with the mineralogy of soils exposed within at-risk lands bordering arid dust source regions.

The EMIT instrument is a Dyson imaging spectrometer that will resolve the distinct absorption features of iron oxides, clays, sulfates, carbonates, and other dust-forming minerals with contiguous spectroscopic measurements in the visible to short wavelength infrared region of the spectrum. EMIT will map mineralogy with a spatial sampling to detect minerals at the one hectare scale and coarser, ensuring accurate characterization the mineralogy at the grid scale required by ESMs. EMIT's fine spatial sampling will resolves the soil exposed within hectare-scale agricultural plots and open lands of bordering arid regions, critical to understanding feedbacks caused by mineral dust arising from future changes in land use, land cover, precipitation, and regional climate forcing.

Data Product	Description	Initial Availability	Median Latency Post-delivery	NASA DAAC
Level 0	Raw collected telemetry	4 months after IOC	2 months	LP DAAC
Level 1a	Reconstructed, depacketized, uncompressed data, time referenced, annotated with ancillary information reassembled into scenes.	4 months after IOC	2 months	LP DAAC
Level 1b	Level 1a data processed to sensor units including geolocation and observation geometry information	4 months after IOC	2 months	LP DAAC
Level 2a	Surface reflectance derived by screening clouds and correction for atmospheric effects.	8 months after IOC	2 months	LP DAAC
Level 2b	Mineralogy derived from fitting reflectance spectra, screening for non-mineralogical components.	8 months after IOC	2 months	LP DAAC

EMIT Level 2a ATBD

Level 3	Gridded map of mineral composition aggregated from level 2b with uncertainties and quality flags	11 months after IOC	2 months	LP DAAC
Level 4	Earth System Model runs to address science objectives	16 months after IOC	2 months	LP DAAC

Table 1: EMIT Data Product Hierarchy

The EMIT Project is part of the Earth Venture-Instrument (EV-I) Program directed by the Program Director of the NASA Earth Science Division (ESD). EMIT is comprised of a Visible/Shortwave Infrared Dyson imaging spectrometer adapted for installation on the International Space Station (ISS). It will be installed on Flight Releasable Attachment Mechanism (FRAM) of an ExPRESS Logistics Carrier (ELC) on the ISS, in a site formally designated ELC 1 FRAM 8. NASA has assigned management of the Project to the Jet Propulsion Laboratory of the California Institute of Technology. The EMIT Payload is scheduled to be installed on the ELC 1 FRAM 8 in 2021. Table 1 above describes the different data products to which the EMIT Mission will provide to data archives. This document describes the “Level 2A” stage.

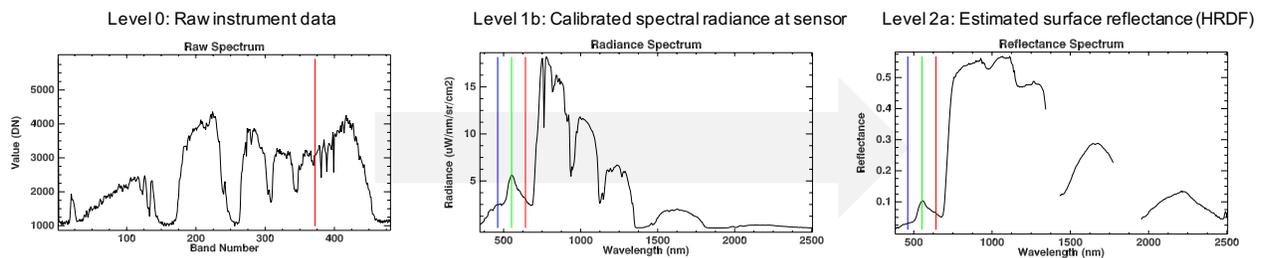


Figure 1: Representative spectra from the EMIT analysis, data product levels 0, 1b, and 2a.

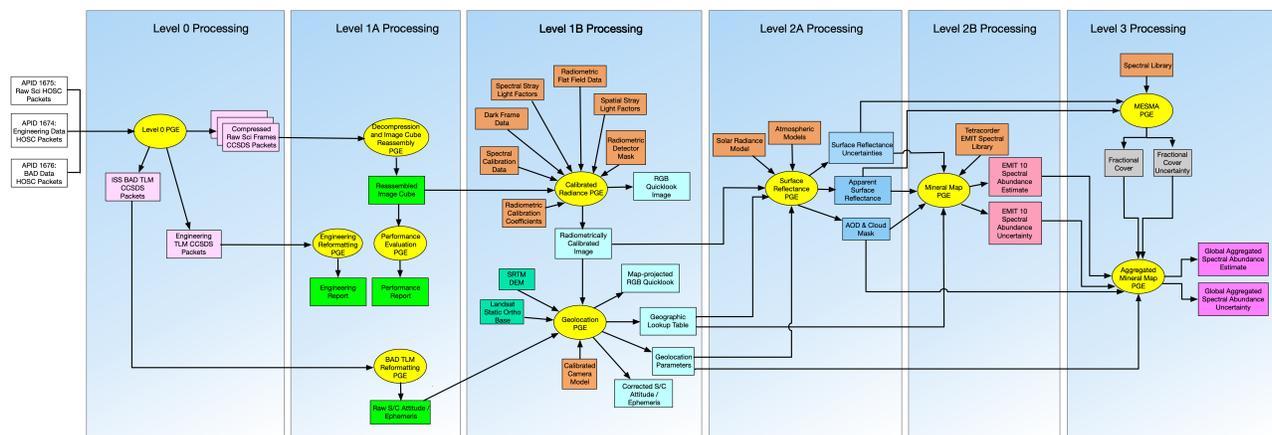


Figure 2. High-level workflow of the EMIT science data system.

This document describes the theoretical basis for the algorithm producing EMIT’s “Level 2a” product. Figure 1 shows examples of the spectrally-defined quantities leading up to this analysis, drawn from an airborne precursor analogue instrument. Figure 2 is a diagram of the Science Data System workflow, including all analysis stages and dependencies. The system begins with a “Level 0” raw data product that records the raw sensor output in digital numbers. The EMIT Science Data System (SDS) applies spectral and radiometric calibration to produce “Level 1” products, e.g. calibrated radiance measurements at the sensor. These are then geolocalized to produce an image that aligns with specific geographic coordinates for matching against digital elevation models. The “Level 2A” inverts these radiance measurements. It uses physically-motivated surface/atmosphere models to estimate atmospheric properties and surface reflectance. The “Level 2A” products include atmospheric parameters and other ancillary files, but the primary output is the surface reflectance

estimate used for later analysis by mineral detection and mapping algorithms. The mineral detection stage (not shown) performs feature fitting on the reflectance data to estimate mineral occurrence, creating a “Level 2B” map at native instrument resolution. This is aggregated into a coarse “Level 3” product for incorporation into Earth System modeling to evaluate Radiative Forcing (RF) impacts. All stages are instantiated in the EMIT science product generation software operating at the Jet Propulsion Laboratory, California Institute of Technology.

3 Algorithm rationale and prior validation

Atmospheric correction (Thompson et al., 2019) has a multi-decadal history of use for imaging spectrometers viewing the Earth surface. This on airborne precursor instruments such as NASA’s “Classic” Airborne Visible Infrared Imaging Spectrometer (AVIRIS-C, Green et al., 1998) and has been extended to its next generation counterpart (AVIRIS-NG, Thompson et al., 2017). Such analyses have been conducted in dozens of campaigns over decades of successful operations. Many empirical methods based on scene averaging (Kruse 1988), flat fielding (Roberts et al., 1986), QUAC (Bernstein et al., 2005), and cloud shadow methods (Reinersman et al., 1998) are useful but do not scale to global observations with diverse scene content and sparse field data. They rely either on manual intervention, or on specific characteristics of the scene such as a spatially homogeneous atmosphere or known scene content, precluding their use with EMIT. Instead, we favor a physically-motivated correction based on radiative transfer models. These have the dual advantages of superior generalizability across scenes without the need for manual intervention in the analysis, and physical interpretability.

Recent reviews surveying different atmospheric alternatives appear in Thompson et al. (2019), Ientilucci and Adler-Golden (2019), and for aquatic spectra, Frouin et al. (2019). Broadly speaking, physically-based methods themselves fall into two general categories (Frouin et al., 2019). *Sequential methods* first estimate atmospheric properties based on analysis of the radiance spectrum, and then invert the radiance directly via closed-form algebra to estimate the surface reflectance. In other words, atmosphere and surface are estimated in two independent steps. Existing physics-based atmospheric correction codes designed for imaging spectrometers all use this general method. They include ACORN (Kruse et al., 2004), ATCOR (Richter and Shlaepfer 2002), ATREM (Gao, 1993) and the AVIRIS-NG standard approach derived from ATREM (Thompson et al., 2015).

Alternatively, *simultaneous methods* estimate surface and atmosphere simultaneously, as in Bayesian Maximum A Posteriori estimation (Thompson et al., 2018, 2019b). Simultaneous methods carry several advantages that are crucial for the EMIT mission. First, they enable rigorous uncertainty accounting. Uncertainty accounting on the input side means respecting instrument noise in the radiance data which can vary by surface type, observing conditions, and wavelength, as well as incorporating any prior background knowledge available in the form of statistical priors. The ability to seamlessly account for these factors makes the Bayesian inversion a flexible and powerful approach to achieve EMIT’s extreme sensitivity requirements. On the output side, uncertainty accounting lets the algorithm propagate posterior uncertainty estimates downstream, where they can improve the performance of mineral fitting algorithms (Thompson et al., 2020b). A second independent benefit of the simultaneous model inversion approach is the demonstrated ability to use the entire spectral range of acquisition in the atmospheric correction, enabling estimation of subtler broad atmospheric perturbations such as aerosols (an EMIT product, in the form of an AOD mask). The main disadvantage is that the methods use an iterative algorithm, leading to higher computational demands.

The EMIT mission uses a Bayesian model inversion strategy, a formalism known colloquially in the community as Optimal Estimation (OE, e.g. Rodgers 2000), with careful application of geospatial interpolation to glean the benefits of both while minimizing cost. The specific OE-based approach

used in EMIT has been validated by decades of operational use by NASA’s atmospheric remote sounding spectrometers on many missions and millions of acquisitions (Rodgers 2000). The approach recently has been extended to VSWIR imaging spectrometers and validated through peer-reviewed field studies with over 20 in situ validation trials of surface reflectance over synthetic, water, vegetated, and bare terrain (Thompson et al., 2018, Thompson et al., 2019b, Thompson et al., 2019c, Thompson et al., 2020). Outside the imaging spectroscopy community, the OE approach has been In situ measurement protocols have also been vetted by decades of continuing operational use (Thompson et al., 2015). The code used is distributed as open source through the public repository at <https://github.com/isofit/isofit/>. This transparency helps for finding errors, and also for end users who desire details on the implementation specifics (e.g. data layout in memory, command flow, etc.). The code will undergo continuing development by a growing community of users throughout the EMIT mission.

4. Algorithm description

4.1 Input data

The EMIT input and output data products delivered to the DAAC use their formatting conventions, the system operates internally on data products stored as binary data cubes with detached human-readable ASCII header files. The precise formatting convention adheres to the ENVI standard, accessible (Jan 2020) at <https://www.harrisgeospatial.com/docs/ENVIHeaderFiles.html>. The header files all consist of data fields in equals-sign-separated pairs, and describe the layout of the file. The specific input files needed for the L2b stage are:

- I. An **observation metadata file**, typically with the string “obs” in the filename, containing information about the observation geometry for every pixel. The observation file uses the original instrument frame (non-orthorectified) coordinate system with size [rows x cols x 12] in Band-Interleaved by Line (BIL) format and single-precision IEEE little-endian floating point representation. It should overlay the radiance data exactly so that all of the pixels are associated between the two files. The channels contain:
 1. Path length – the direct geometric distance from the sensor to the location on the surface of the Earth, as defined by a Digital elevation model
 2. To-sensor azimuth, in decimal degrees, at the surface
 3. To-sensor zenith, in decimal degrees, at the surface
 4. To-sun azimuth, in decimal degrees, at the surface
 5. To-sun zenith, in decimal degrees, at the surface,
 6. Phase angle in degrees, representing the angular difference between incident and observation rays
 7. Terrain slope in degrees as determined from DEMs,
 8. Terrain aspect in degrees, as determined from DEMs,
 9. The cosine of the solar incidence angle relative to the surface normal
 10. UTC time
- II. A **location file**, typically with the string “loc” in the filename, containing information about the geographic projection of each spectrum. The location file is left in the original non-orthorectified instrument coordinate system, with size [rows x cols x 3] in Band-Interleaved by Line (BIL) format and single-precision IEEE little-endian floating point representation. It should overlay the radiance data exactly. The channels contain:
 1. Latitude of surface, in decimal degrees, with a WGS-84 datum
 2. Longitude of surface, in decimal degrees, in degrees East of zero, with a WGS-84 datum
 3. The average elevation of the surface, as determined from a Digital Elevation Model
- III. A **geographic lookup table file**, typically with the string “glt” in the filename, containing information about the index into the unorthorectified data of each spectrum. It is projected

to a geographic coordinate system, with size [rows x cols x 2] in Band-Interleaved by Line (BIL) format and 32-bit unsigned integer representation. Its columns contain the row and column indices, respectively, of each spectrum in the original unorthorectified data.

IV. **Radiance data at sensor**, typically with the string “rdn” in the filename, in units of $\mu\text{W} / \text{cm}^2 / \text{nm} / \text{sr}$. The data is in the instrument frame (non-orthorectified representation with size [rows x cols x channels] in Band-Interleaved by Line (BIL) format and single-precision IEEE little-endian floating point representation. The precise number of channels is not yet determined at the time of this writing but should be a value close to 300.

“Bad data” at the periphery outside the field of view, or masked as a result of cloud masking or instrument error, is typically assigned the reserved (floating point) value -9999. In addition to these files above, which change on a per acquisition basis, the L2A stage uses a wide range of ancillary files in its configuration. These include configuration files themselves, climatology and physical reference data, surface, atmospheric, and instrument model data, and more. These ancillary files are outside the scope of this ATBD, where we will concern ourselves with the data associated with a particular product and acquisition. We will also disregard internal configurations used by the science data system for managing and running these processes. Table 2 Below enumerates all products.

Input file	Format	Interpretation
Observation Metadata	rows x columns x 12, BIL interleave 32-bit floating point with detached ASCII header	Varies (see text)
Location File	rows x columns x 3, BIL interleave 32-bit floating point with detached ASCII header	Latitude in decimal degrees, Longitude in decimal degrees, elevation of surface in meters
Geographic Lookup Table	rows x columns x 2, BIL interleave 32-bit unsigned integer, detached ASCII header	Row and column index into unorthorectified instrument data
Radiance data	rows x columns x channels, BIL interleave 32-bit floating point with detached ASCII header	Radiance at sensor in $\mu\text{W} / \text{cm}^2 / \text{nm} / \text{sr}$.

Table 2: Input files

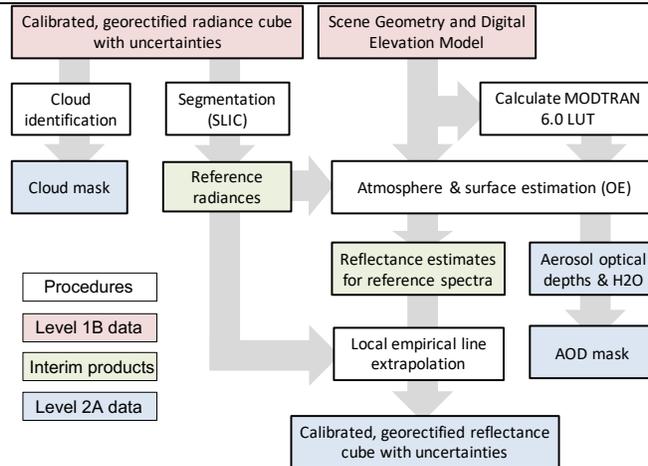
4.2 Theoretical description

Broadly speaking, the EMIT atmospheric correction stage has several goals. All operations involve the atmosphere to some extent, and operate on the calibrated radiance files, so we combine them for organizational convenience into a single product level. The Level 2a output includes:

- Estimates of local aerosol and atmospheric water vapor content of the atmosphere along with the Lambertian-equivalent surface reflectance, all with uncertainty predictions.
- Flags for high-haze conditions (i.e. aerosol optical depths exceeding our working range).
- Mask for cloudy regions of the scene which lets downstream analyses to exclude them.

Figure 3 below illustrates the sequence of operations along with the major input and output products at each stage. All procedures execute sequentially moving from top to bottom. Boxes are colored according to their designation as level 1B, level 2A, or intermediate products. Since cloud masking is a separate operation with minimal dependencies or algorithmic relationship to the surface atmosphere estimation, we treat it separately in a later section.

Figure 3: Sequence of operations in the EMIT level 2A stage. All reflectance and atmosphere estimates also include uncertainty predictions.



4.2.1 Radiative Transfer and Atmospheric Modeling

Physics-based retrieval of atmospheric parameters and surface reflectance typically relies on mathematical models expressing the spectral radiance reflected by the Earth's surface and atmosphere at the top-of-atmosphere (TOA) as a sum of radiative terms from different processes, such as the photons scattered by the atmosphere into the sensor line of sight or those multiply scattered between the atmosphere and the surface (Figure 2). While in general the atmospheric effects are dependent on non-Lambertian properties of surface-atmosphere coupling, the EMIT analyses permit several simplifications. The mineral absorption fits used in later stages are relatively invariant to spectrally-featureless magnitude differences resulting from non-Lambertian behavior. Additionally, surfaces in arid mineral dust forming regions are mostly Lambertian at that instrument's ground sampling, unlike – for example – dense tree canopies or open ocean. Finally, imaging geometry is near to nadir. These circumstances mean that we can report Lambertian-equivalent properties in the general case without significant loss of accuracy to downstream algorithms. This permits the following decomposition (Thompson et al., 2018):

$$\rho_{toa} = \frac{\pi \mathbf{L}_M}{\mathbf{F} \cos(\theta)} = \rho_a + \frac{\mathbf{T} \circ \rho_s}{1 - \mathbf{S} \circ \rho_s} \quad (1)$$

where boldface denotes vector-valued quantities (in this case, spectra) and circles represent element-wise multiplication. The symbol ρ_{toa} refers to the top-of-atmosphere reflectance; it is based on the radiance measurement \mathbf{L}_M , the extra-terrestrial solar flux \mathbf{F} and the solar zenith angle θ . The symbol \mathbf{T} is the direct and diffuse transmission of the mean optical path from sun to ground to sensor, \mathbf{S} is the spectral albedo representing the atmospheric reflectance as seen from the surface, ρ_a is the path reflectance of the atmosphere, and ρ_s is the Lambertian-equivalent surface reflectance.

These terms are related to several physical properties in the atmosphere. Of special interest are the scattering and absorption by molecular gases and aerosols (Figure 4), which all contribute to each of the terms in equation 1. An example of the transmittance contribution from gas absorption appears in Figure 5 below, adapted from (Thompson et al., 2019). We calculate the EMIT atmospheric absorption and scattering property estimates using the MODTRAN 6.0 Radiative Transfer Model (Berk et al., 2016; 2016b). Given a specific atmospheric state, MODTRAN can estimate the optical coefficients \mathbf{S} , \mathbf{T} , and of Equation 1 from physical first principles. We perform this estimation at high spectral resolution, transform the result to the EMIT instrument spectral response.

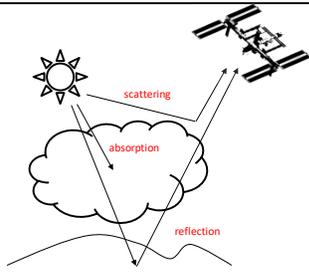
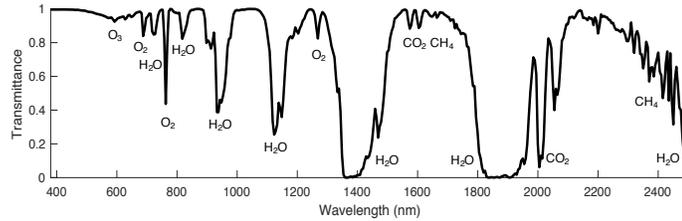


Figure 4: The atmospheric correction process involves jointly estimating the parameters of a model that includes the surface reflectance, the atmospheric constituents, and the instrument. A wide range of different physical effects, including scattering, absorption, and the target surface reflectance signal, all play a role in determining the photon distribution at the sensor.

Figure 5: Atmospheric gas absorption by wavelength across the EMIT spectral interval.



The MODTRAN 6.0 atmospheric gas absorption model uses a “correlated k” approach with absorption coefficients from the HITRAN 2012 line list (Rothman et al., 2012). Following on prior work, we augment the basic configuration with three basic aerosol signatures (Thompson et al., 2019b) representing small, medium, and large particles such as soot, sulfates, and dust. The soot and sulfate-derived signatures are spherical, while dust particles are nonspherical. All three are described by spectral absorption, extinction, and asymmetry profiles in prior work (See Figure 6, adapted from Thompson et al., 2019c). The complete aerosol is a contribution of all three aerosol optical depths, each specified independently (typically at the reference wavelength 550 nm). These three signatures are used structured error terms in the inversion process to improve atmospheric correction. This also permits an aggregate AOD estimate for scene flagging. In Figure 6, type A is a strongly absorbing aerosol signature derived from soot. Type B is a separate signature based on continental dust absorption and scattering coefficients. Type C is a small scattering particle based on a sulfate signature. We advise against interpreting the individual retrieved as physical properties of the particles.

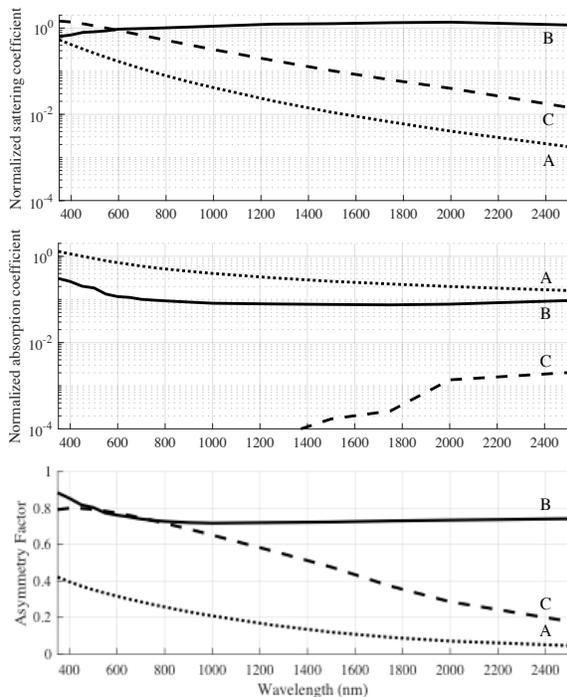


Figure 6: Aerosol profiles (image and approach adapted from Thompson et al., 2019c). We use three signatures as structured error terms in the inversion process to improve atmospheric correction and to permit an aggregate AOD estimate for scene flagging. Type A is a strongly absorbing aerosol signature derived from soot. Type B is a separate signature based on continental dust absorption and scattering coefficients. Type C is a small scattering particle based on a sulfate signature. While the inversion process estimates the AOD of each signature independently, we advise against interpreting the individual retrieved as physical properties of the particles due to uncertainties in vertical distribution and ambiguity in optical properties.

4.2.2 Model Inversion

Our retrieval algorithm is based on Bayesian Maximum A Posteriori (MAP) inversion of equation 1, using a strategy known colloquially as Optimal Estimation (OE, Rodgers, 2000). This approach has been demonstrated previously in multiple imaging spectrometer field studies (Thompson et al., 2018, 2019b, 2019c). Its advantages include rigorous uncertainty propagation and the ability to estimate atmospheric aerosol constituents in high AOD conditions. The main disadvantage is a high computational cost due to the iterative inversion algorithm, which must run independently on every spectrum. Here, we address this by running the full algorithm on a representative subset of several thousand spectra per scene. These results enable a highly accurate, spatially-local empirical line estimate for the remainder, allowing millions of spectra to be corrected and capturing the benefits of the iterative approach at a feasible computational cost.

The Bayesian Model inversion acts as a local ascent of the posterior probability density for a state vector \mathbf{x} consisting of surface and atmosphere parameters (Figure 7). As in Thompson et al. (2018) we initialize the result to a heuristic estimate using a band ratio across water vapor absorption features, and an algebraic inversion of equation (1). Then, an iterative gradient-based Levenberg Marquardt follows the (negative) derivative of the following cost function until converging to a local minimum:

$$\chi^2(\mathbf{x}_r) = \frac{1}{2}(\hat{\mathbf{x}}_L - \mathbf{F}(\mathbf{x}_r))^T \Psi_L^{-1}(\hat{\mathbf{x}}_L - \mathbf{F}(\mathbf{x}_r)) + \frac{1}{2}(\mathbf{x}_r - \mu_r)^T \Sigma_r^{-1}(\mathbf{x}_r - \mu_r) \quad (2)$$

The first term is related to the logarithm of the multivariate data likelihood at the current state vector; the second term penalizes departures from the prior in similar fashion. All probability distributions are multivariate Gaussians. Here Ψ_L is the observation noise that incorporates measurement noise in the radiance measurement $\hat{\mathbf{x}}_L$ as well as any unknowns in the surface atmosphere system that are treated here as random variables. The forward model $\mathbf{F}(\mathbf{x}_r)$ maps the reflectance and atmosphere state vector, \mathbf{x}_r , to the measurement space using Lookup table interpolation of optical coefficients in Equation 1. The multivariate Gaussian prior over surface and atmosphere is defined by Covariance matrix Σ_r and mean μ_r . These priors are intentionally set to be extremely broad in order to avoid estimation bias in atmospheric parameters. Similarly, we use a very loose and heavily regularized surface prior. It is based on a collection of multivariate Gaussians, as suggested in Thompson et al., (2018, 2019a, 2019b), using the Euclidean-nearest component of the initial state calculated in reflectance space as the prior. All spectra are L2-normalized for the purposes of calculating these distances and prior distributions so that the distribution affects the shape but not the magnitude of spectra. The only difference with the formulation in these previous studies is that all wavelengths outside critical atmospheric windows are left *entirely decorrelated*. This allows instrument noise to enter the reflectance estimate unmodified, and permits highly accurate retrieval of absorption features in mineral bands.

Upon convergence, the linearization of the forward model produces an estimate of the posterior probability density. For \mathbf{K}_r representing Jacobian matrices of partial derivatives, i.e. the instantaneous change in the state vector from a change in the calibrated radiance, the posterior covariance takes the form:

$$\Psi_r = (\mathbf{K}_r^T \Psi_L^{-1} \mathbf{K}_r + \Sigma_r^{-1})^{-1}$$

This yields a reflectance, atmosphere, and uncertainty estimate for each reference spectrum. The final step is an Empirical Line operation (Thompson et al., 2016) that uses the k nearest solutions to extrapolate an exact solution for the high-resolution data.

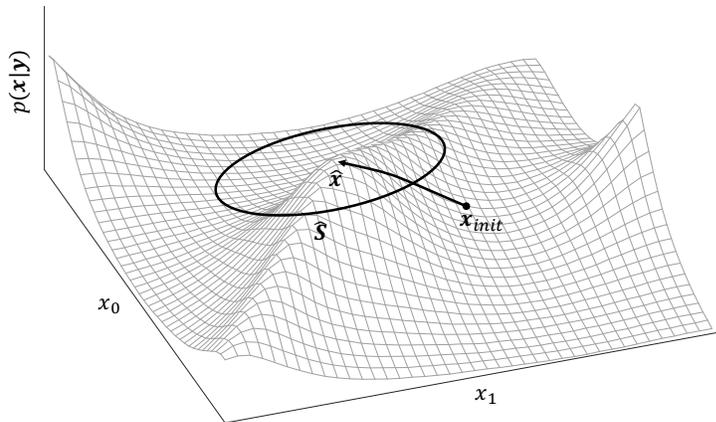
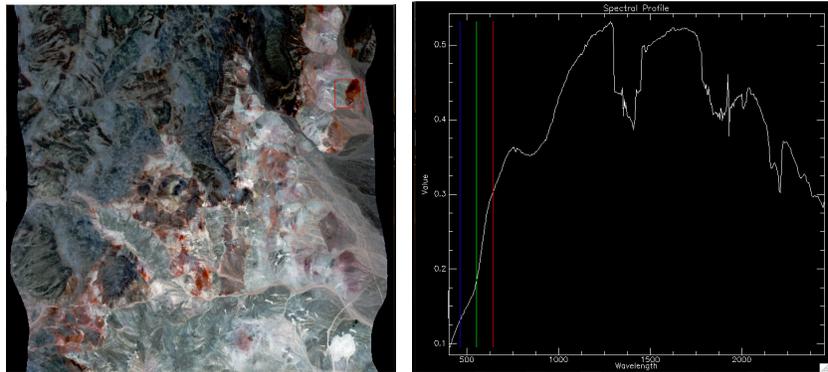


Figure 7: The Bayesian model inversion begins at an initial guess, and climbs the local gradient of the posterior probability density (equivalently, minimizing the cost function in equation 2). At the time of convergence, this produces a linearized estimate of posterior uncertainty, portrayed here as an ellipsoid.

Figure 8: (Left) Cuprite, NV scene. (Right) Interpolated OE estimation of a single reflectance spectrum, via the local empirical line solution. Sharp, spectrally-diagnostic Kaolinite features are visible in the 2-2.5 micron range.



4.2.3 Superpixel Segmentation

Since complete model inversion of every spectrum is computationally intractable, we use a segmentation to identify representative spectra in the flightline where we apply our model inversions. After performing the atmosphere/surface estimation on the representative subset of, we assign the atmospheric estimates to each location associated with that segment. We then use the representative spectra to calculate local “Empirical line” solutions (Moran et al., 2001, Thompson et al., 2016). The empirical line performs the exact atmospheric correction for all independent (non-aggregated) spectra at maximum spatial resolution.

The initial segmentation uses a superpixel aggregation approach based on the SLIC algorithm (Achanta et al., 2012). We reduce all spectra in the file to a basis of five orthogonal dimensions with principal components analysis, and segment the result into regions that are (a) contiguous and (b) contain several hundred pixels of similar radiance properties. Figure 9 illustrates the superpixel segmentation of a scene from NASA’s Next Generation Airborne Visible Infrared Imaging Spectrometer (AVIRIS-NG). It results in a reduced subset of locally-representative radiances and associated regions. This dataset is typically 2-3 orders of magnitude faster to analyze. Additionally, it significantly reduces noise variance to assist with accurate atmosphere estimation. Similarly, we take the mean radiance and location of each segment as the input to the following atmospheric correction.

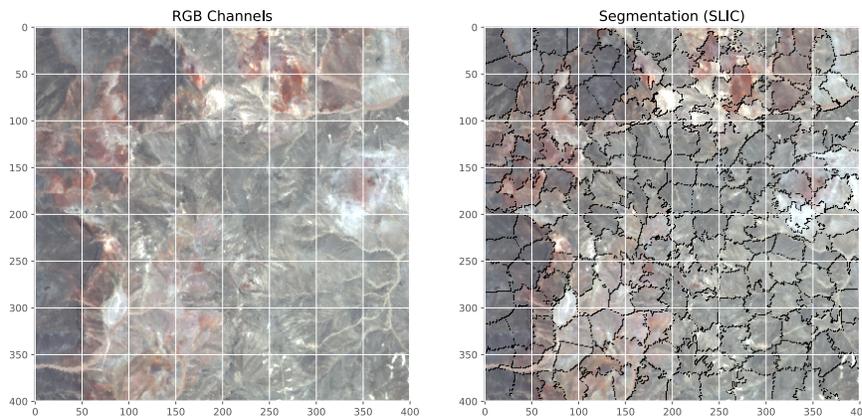
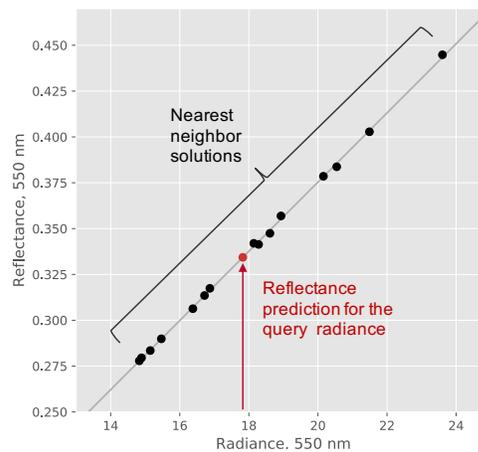


Figure 9: SLIC segmentation combines contiguous pixels of similar radiance properties into a single local reference area and associated radiance spectrum.

4.2.4 Empirical Line extrapolation

The empirical line (Moran et al., 2001) first identifies the nearest 15 representative spectra for the current segment. Optionally, one can weight vertical distance on the DEM differently from horizontal distance. It then calculates an independent linear least-squares regression solution for every wavelength, finding the offset and coefficient that maps radiance onto reflectance for that specific atmospheric condition. Coefficients are calculated once per segment and cached for application to the fine-scale spectra within. Because the empirical line estimate is spatially localized the linear solution for each wavelength respects local variability of the atmosphere while providing both AOD and H₂O estimates over the entire scene. It accurately identifies signatures that appear in small single- or sub-pixel locations that are not apparent in the spatially-aggregated estimate. Figure 10 shows an example of the empirical line prediction for the Cuprite, NV scene at 550 nm, showing that the representative superpixels' radiances and reflectances do indeed exhibit a locally-linear relationship. This permits a highly accurate estimate for the fine-scale spectrum shown in red.

Figure 10: Empirical line estimate for an example segment of the Cuprite flightline portrayed in Figure 6. Black points show the mean spectra of the 15 nearest reference spectra to a fine-scale radiance we aim to invert. Since the relationship is locally linear, it is easy to quickly determine an accurate solution via linear regression.



4.2.5 Cloud Masking

The radiance data analysis begins with a cloud masking operation following on the prior work of Thompson et al (2014). This procedure places prior distributions on the distributions of top of atmosphere reflectances at three representative wavelengths: 420 nm, 1250 nm, and 1650 nm. The distributions of intensities over clouds and Earth's surface are well-separated, permitting a single trivariate threshold to flag clouds.

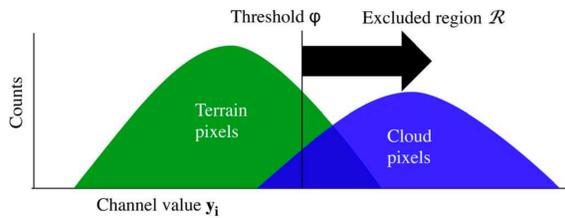


Figure 11: Thresholds in three channels combine to define a hypercubic “excluded region.” Any pixel exceeding the threshold in all three channels is considered cloudy.

Any pixel exceeding the threshold in all three channels is flagged as cloudy, and the result written to a binary cloud mask for use in later analysis. The thresholds define an exclusion region \mathcal{R} (Figure 11) that defines the boundary between cloud and clear locations. Consequently, for a top of atmosphere reflectance spectrum \mathbf{y} the decision rule $f(\mathbf{y})$ classifying the pixel as either cloudy, written c_1 , or clear, written c_2 , is simply:

$$f(\mathbf{y}) = \begin{cases} c_1, & \text{if } \mathbf{y} \in \mathcal{R} \\ c_2, & \text{if } \mathbf{y} \notin \mathcal{R}. \end{cases}$$

The Bayesian formalism of Thompson et al (2014) allows the designer to set a threshold on all three channels that optimally balances the risk of data loss from false positives against the risk of passing cloud pixels into subsequent analysis. For the EMIT mission, most surfaces of interest are far darker than clouds, so we set a conservative threshold during cloud masking. The expected loss is a function of the prior class probabilities (taken here as uninformed or uniform), the probability density of a given observation for cloudy classes, for a given state variable \mathbf{x} representing the local surface “type” and season. For simplicity we ignore this last variable, leading to a uniform decision rule applying equally to all different surfaces. The expected loss for a relative weighting of false positive and false negative errors, respectively written α_{FP} and α_{FN} , is:

$$E[\mathcal{L}] = \int_{\mathcal{R}} \alpha_{FP} P(\mathbf{y} | \mathbf{x}, c_1) P(c_1) d\mathbf{y} + \int_{\mathbb{R}^d \setminus \mathcal{R}} \alpha_{FN} P(\mathbf{y} | \mathbf{x}, c_2) P(c_2) d\mathbf{y}.$$

A false positive weighting of 10-1000 is a conservative threshold that is very unlikely to exclude bright surface spectra. Figure 12 below shows that bare terrain is much darker than typical cloud spectra, particularly in the ultraviolet and blue channels. Water is uniformly dark. Snow or ice can be bright in the visible channels, but exhibits high absorption in the near and shortwave infrared which permits any thresholding method of two or more channels to separate them effectively. The panel at right shows the bivariate exclusion regions for two representative false positive weight values, illustrating that the two are sufficient to discriminate snow and terrain from cloud pixels in a diverse historical dataset. Figure 13 below, taken from Thompson et al. (2014), shows an example application of this approach to a scene from NASA’s “Classic” Airborne Visible Infrared Imaging Spectrometer (AVIRIS-C). The left panel shows the original scene in red, green, and blue optical wavelengths. The scene contains both bright snow and dispersed clouds. The middle panel shows the result of the cloud masking operation using the channelwise threshold. We note that an onboard cloud masking operation performs a similar operation, excising the most obvious clouds to reduce data volumes. This operation uses an even more conservative threshold, and excises entire vertical segments of a scene with too many cloud pixels.

We dilate the detected cloud masks in order to avoid cast shadows and disruption of the incident light field adjacent to clouds. Figure 14 shows a graphical illustration of the method. The maximum cloud height and solar zenith angle define a geographic exclusion area around each cloud where shadows may be found. We apply an efficient image-space distance transformation to the cloud mask, producing an array specifying every image locations’ distance to the nearest cloud pixel. We flag any pixel whose distance lies within the exclusion interval. The maximum cloud height is 3000

m, which is low enough to preserve data but high enough to capture the majority of optically-thick cumulus clouds. Higher altitude stratus clouds are less important for this purpose because they are more likely to be large, spatially-continuous cloud masses. Such clouds subtend a large enough fraction of the image to be filtered by the onboard cloud masking system. Additionally, their altitudes are high enough that they would require an infeasibly large exclusion region. This cloud height may be reassessed during the mission.

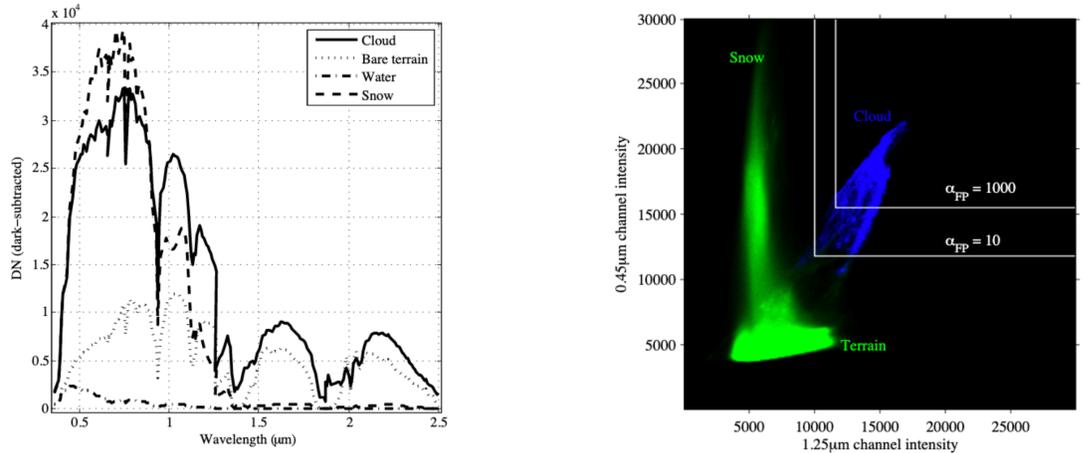


Figure 12: Cloud screening illustration, from Thompson et al. (2014). Left: Brightness distributions for cloud, bare terrain, and snow have very different spectral shapes. Right: Two of three threshold channels showing “exclusion regions” defined by different tolerances for false positives.

Figure 13: L2A cloud screening compared to onboard (real-time) excision, adapted from Thompson et al (2013). The right panel shows an acquisition by NASA’s “Classic” Airborne Visible Infrared Imaging Spectrometer (AVIRIS-C). We show visible channels of scene content including bare dark terrain, bright snow-covered terrain, and cloud. The middle panel shows the L2A pixel-wise cloud masking. The onboard excision performs a pre-screening using an even more conservative threshold to reduce transmitted data volume; this excises vertical segments of each scene that contain more than an acceptable number of cloud pixels.

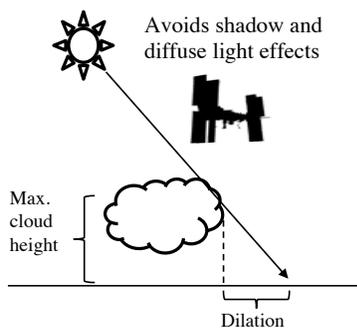
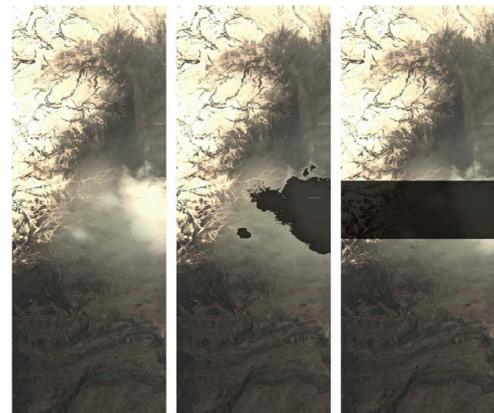


Figure 14: The cloud mask dilation excludes pixels that are likely to contain contamination by cloud diffuse illumination or cloud cast shadows. We excise pixels in a conservative exclusion region defined by the solar zenith and a *maximum cloud height* parameter.

4.3 Practical Considerations

Due to the computationally-demanding nature of the EMIT L2A stage, operators must attend to the balance between accuracy and speed in their settings for approximations like the lookup table grid

spacing (which affects the number of MODTRAN runs) and the number of superpixels (which affects the accuracy of empirical line extrapolation). Currently, a three- or four-point Aerosol AOD model is used, with linear interpolation between. The H₂O model uses a 0.2 g/cm² spacing. As computational resources permit, these numbers will be relaxed. As of the writing of this document, a typical airborne flightline requires 1-2 days to complete for a single CPU; given a cluster with many CPUs, keeping up with the EMIT datastream is feasible. However, we anticipate further accuracy improvements as additional CPUs come online.

5. Output Data

The EMIT output data products delivered to the DAAC use their formatting conventions, the system operates internally on data products stored as binary data cubes with detached human-readable ASCII header files. The precise formatting convention adheres to the ENVI standard, accessible (Jan 2020) at <https://www.harrisgeospatial.com/docs/ENVIHeaderFiles.html>. The header files all consist of data fields in equals-sign-separated pairs, and describe the layout of the file. The specific output files from the L2b stage are:

- I. A **surface reflectance file**, typically with the string “rfl” in the filename, containing the estimated spectral surface reflectance for every pixel. It is provided in the non-orthorectified instrument coordinate system with size [rows x cols x channels] in Band-Interleaved by Line (BIL) format and single-precision IEEE little-endian floating point representation. It should overlay the orthorectified radiance data exactly so that all of the pixels are associated between the two files.
- II. A **reflectance uncertainty file**, typically with the string “uncert” in the filename, containing predicted uncertainty in the reflectance measurement for each channel, in units of standard deviations (presuming a Gaussian distribution). Covariance is ignored. It is provide in the non-orthorectified instrument coordinate system with size [rows x cols x channels] in Band-Interleaved by Line (BIL) format and single-precision IEEE little-endian floating point representation. It should overlay the reflectance and radiance data exactly.
- III. A **mask file**, typically with the string “mask” in the filename, containing channels with the following information:
 1. Probability this pixel is cloud
 2. Probability this pixel is standing water
 3. Dilated cloud mask
 4. Aerosol Optical Depth (550 nm)
 5. Estimated Columnar Water Vapor (g cm⁻²)
 6. Aggregate bad data flag

The fourth channel applies EMIT’s masking rules to the otherchannels in order to determine whether that pixel will be used in subsequent aggregation to the Level 3 product. The file is projected into a geographic coordinate system with size [rows x cols x channels] in Band-Interleaved by Line (BIL) format and single-precision IEEE little-endian floating point representation. It should overlay the reflectance and radiance data exactly.

Any file can contain “bad data” as a result of cloud masking or instrument error. These pixels are typically assigned the reserved (floating point) value -9999. Table 2 Below enumerates all products.

Output file	Format	Interpretation
Reflectance	rows x columns x channels, BIL interleave 32-bit floating point with detached ASCII header	Lambertian-equivalent surface reflectance
Uncertainty	rows x columns x channels, BIL interleave 32-bit floating point with detached ASCII header	Reflectance uncertainty (one standard deviation)

Mask	rows x columns x 5, BIL interleave 32-bit unsigned integer, detached ASCII header	Varies by channel (see above).
------	---	--------------------------------

Table 3: Output files

6. Calibration, Validation, and Field Measurement

Level 2 reflectances will be validated using standard field protocols used in prior field studies (Thompson et al., 2018, 2019a, 2019b, 2020a). We will measure surface reflectance of a large uniform bright surface, such as a playa, using field spectroradiometers, with coincident in-situ AEROSOL optical depth estimation by sun extinction measurements from the ground, during the EMIT overflight. Instrument measurement and spatial variability, combined with uncertainties in the atmospheric model and retrieval, can demonstrate closed uncertainty budgets as in Thompson et al. (2020a) or simply good agreement between the estimate and reality, as in Thompson et al. (2018). Figure 15 below shows examples of a calibration/validation experiment at Stonewall Playa, Ivanpah, with the spectroradiometer field unit (left panel), the playa itself (center panel), and the comparison of reflectances (right panel). Our calibration and validation plan includes several locations that we will use opportunistically in response to ISS overpasses.

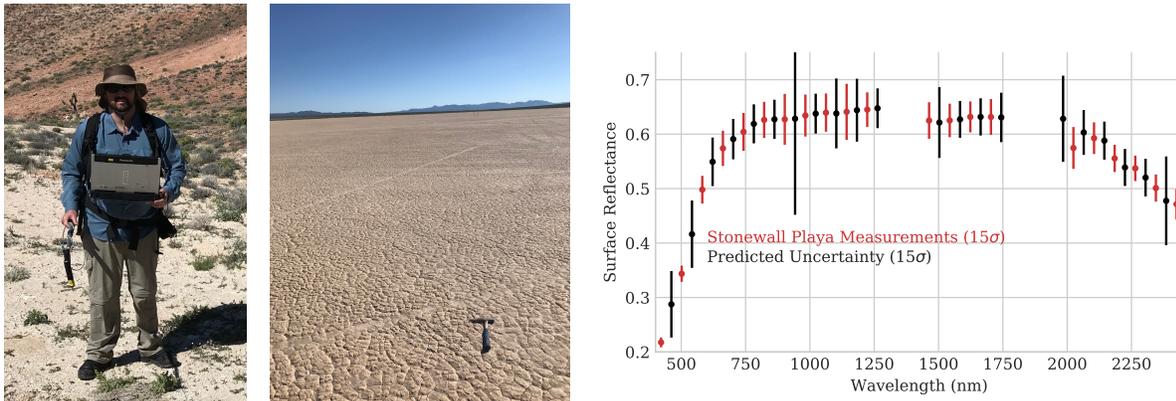


Figure 15: Left: Field spectroradiometer for validation. Center: Stonewall Playa validation site. Nimrod Carmon demonstrating. Right: Remote and in-situ retrievals with 15σ uncertainty predictions (Thompson et al., 2020).

Prior verification and validation for the Level 2 algorithm takes several approaches. The codebase is available as open source (ISOFIT, 2019) and has a growing community of users in the research community. The method draws from decades of atmospheric sounding research (Rogers 2000) and its specific application to imaging spectroscopy has been vetted for multiple instruments and campaigns across continents, compared with in situ data and published in peer reviewed literature. Publications referencing the results of this code on airborne precursor data include work by Thompson et al. (2018, 2019b, 2019c), Frouin et al. (2019), and Bue et al. (2019). Field trials demonstrate good alignment with in-situ reflectance data, and residuals consistent with posterior error predictions. Figure 16 shows one example from Ivanpah Playa, conducted in 2018.

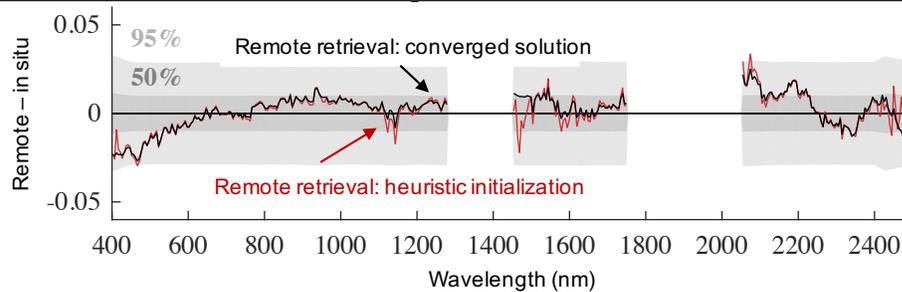
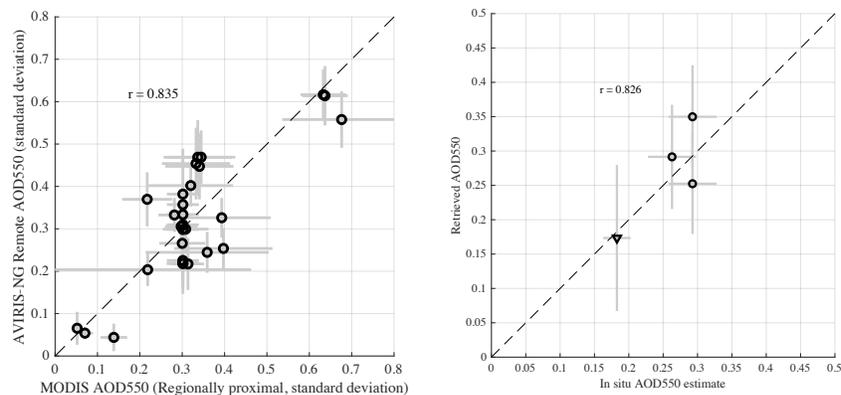


Figure 16: In situ validation of reflectance estimation algorithm. (Above) In situ and remote measurements align to within posterior error predictions. Adapted from Thompson et al. (2018). (Right) Field validation at Ivanpah Playa, from Thompson et al (2019).

AOD estimates show good alignment with spatiotemporally-proximal MODIS retrievals over difficult hazy conditions, and with in-situ estimates by handheld sunphotometers (Figure 17).

Figure 17: (Left) MODIS AOD550 estimates align with remote airborne retrievals acquired on the same day within a latitude/longitude degree. (Right) Airborne retrievals align with in-situ sunphotometry. Both images are from Thompson et al. (2019c).



For the EMIT mission we performed a separate sensitivity study to determine the degree to which aerosol type mismatch during atmospheric correction could impact surface mineralogy estimates. Specifically, we examine a mismatch between the template aerosol profiles in the EMIT surface/atmosphere retrieval process and the “true” optical properties of aerosols in the atmosphere. It is likely that the optical properties in the retrieval and atmosphere never match exactly; templates are intended as generic flavors of distortion that the inversion can mix in proportions to achieve good quality inversions. It is reasonable to ask whether an unforeseen optical type, not captured by the combinations of palette options, could induce an erroneous residual shape in the surface reflectance. Most damaging would be an absorbing aerosol that bears its own minerals inducing some hallucinatory mineral-like change in the surface reflectance. Such situations would not be common in practice, though mineral absorption profiles are occasionally visible in dust plumes imaged historically by spectrometers under extreme conditions (Chudnovsky et al., 2009).

Our experiment uses an atmosphere based on the iron-oxide-bearing dust mineral profile in the CAM earth system model. This is a strongly absorbing aerosol with shapes distinctly different from the profile palette in our inversion. Notably, the shapes of optical absorptions by atmospheric dust also differ significantly from the surface minerals. They are also somewhat muted in their airborne dust form due to embedding within larger particles. As a consequence, we hypothesize that a band depth estimate of hematite absorption surface signatures should not be significantly affected by any surface reflectance error from this mismatch. To test this, we simulate a stressing case in which the instrument observes a hematite absorption feature, with and without an additional perturbation at 2% relative band depth. This level of sensitivity is the detection limit targeted by EMIT. Our

reference atmosphere presumes typical ISS viewing geometry, but under very hazy conditions. The AOD550 is fixed at our mission-level acceptance threshold 0.4, beyond which a “bad data” flag would be triggered. Figure 18 shows the optical properties of the ESM iron oxide aerosol, given as spectral absorption and single scattering albedo efficiencies for a total extinction normalized to unity at 550 nm.

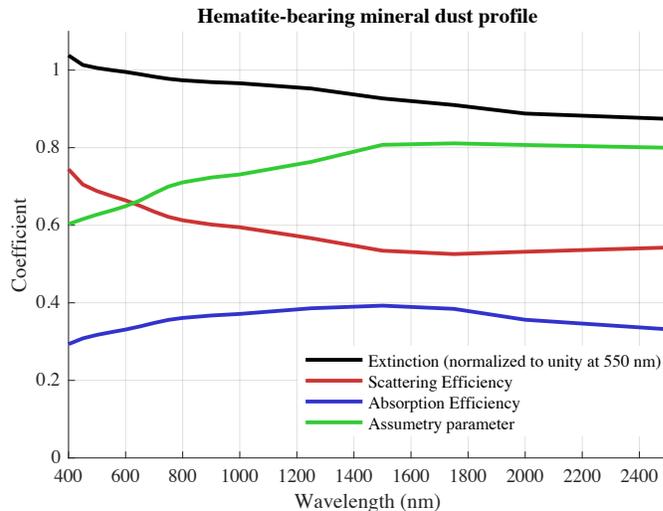
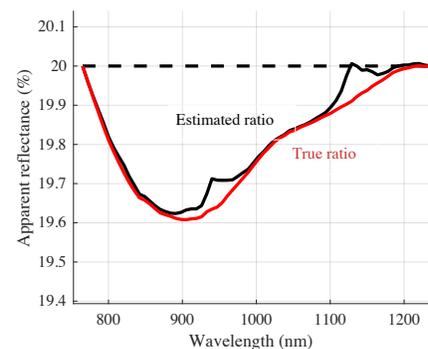


Figure 18: Dust optical properties for the iron oxide profiles used in the CAM ESM model. The aerosol absorption profile, in blue, contains subtle iron oxide absorption features but these atmospheric signatures are muted relative to the (stark) mineralogical features visible at the surface.

We calculate the continuum-relative surface reflectance absorption from the USGS spectral library version 7.0. We performed forward prediction of TOA radiances via the MODTRAN RTM, added relevant instrument noise calculated via Current Best Estimate (CBE) instrument models, and finally inverted the result with our standard atmospheric correction algorithm. For each result, we estimated the reflectance of the unperturbed and perturbed case, with the ratio of the two showing the estimated relative difference in hematite. The resulting relative absorptions appear in Figure 19 below. Note that the unexpected distortion of H₂O vapor features induces some structured error near those windows at 940 and 1140 nm. However, the overall depth and shape of the critical hematite absorption is not significantly affected. This is also apparent in the resulting band depth estimate vis a vis the interpolated continuum, which differs by a small percent of the surface reflectance - 19.626% vs 19.608% for the undistorted case. In other words, small perturbations of the background, recognized using spectral shapes based on relative radiometry (the EMIT strategy), are not significantly distorted. The relative difference of <0.1% would not endanger the ability to detect the addition of hematite to the surface at 2% areal fractional occurrence, and would not affect the mineralogy estimates to a level that would endanger mission success.

Figure 19. The result of atmospheric simulation and inversion under mismatched aerosol optical types, retrieving an iron oxide mineral signature under an iron-oxide-bearing aerosol. The unexpected distortion of H₂O vapor induces some structured error near those absorptions at 940 and 1140 nm. However, the overall depth and shape of the critical hematite absorption feature is not significantly affected.



7. Constraints and limitations

Two main caveats on the atmospheric correction bear emphasis. First is the challenge of generalizing performance guarantees past the nominal range of observing conditions. The EMIT mission uses conservative values to create masks and acquisition plans to exclude poor observing conditions that would spoil atmospheric correction model assumptions and/or accuracy. These include masking on:

- Total aerosol optical depth at 550 nm (AOD550)
- Solar zenith angle
- Distance to screened clouds.

These thresholds to control the pixels that appear in the mask and enters the level 3 stage. They were designed conservatively to ensure good data quality downstream. However, since all spectra will be made available at the L2A stage alongside the masks, investigators may choose not to apply them and use the “bad” data anyway. We caution the investigators that such atmospheric and observing regimes are outside the bounds of our modeling and analysis, and our performance assessments cannot apply in those cases.

A second important caveat, noted above, is that the main purpose of the level 2A stage is to determine the surface reflectance. Because the values of Aerosol and water vapor retrievals may not be validated against physical standards, we caution against their interpretation as physical parameters of the atmosphere. In particular, it is likely that neither quantity would exactly match direct in-situ observations of similar quantities due to differences in the optical absorption path. But even along the same path, we intend these parameters as a means to correct atmosphere-like distortions in surface reflectance rather than measurement targets in themselves.

8. Code repository and references

8.1 Repository

The EMIT L2a code is based on the ISOFIT codebase, open source under the Apache 2.0 license and available at the following URL:

<https://github.com/isofit/isofit>

Tutorial materials on the atmospheric correction process and code examples are located at:

<https://github.com/davidraythompson/istutor>

8.2 References

- Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., & Süsstrunk, S. (2012). SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(11), 2274-2282.
- Chapman, J., Thompson, D. R., Helmlinger, M. C., Eastwood, M. L., Bue, B. D., Geier, S., Green, R. O., Lundeen, S. R., Olson-Duvall, W. (2019). Spectral and Radiometric Calibration of the Next Generation Airborne Visible Infrared Spectrometer (AVIRIS-NG). *Remote Sensing*, 11(18), 2129.
- Chudnovsky, A., E. Ben-Dor, A. B. Kostinski, and I. Koren. "Mineral content analysis of atmospheric dust using hyperspectral information from space." *Geophysical Research Letters* 36, no. 15 (2009).
- Berk, A., et al. (2016). Algorithm Theoretic Basis Document (ATBD) for Next Generation MODTRAN®. Spectral Sciences, Inc.: Burlington, MA, USA.
- Berk, A., J. van den Bosch, F. Hawes, T. Perkins, P.F. Conforti, G.P. Anderson, R.G. Kennett, P.K. Acharya (2016b). MODTRAN®6.0.0 User's Manual (revision 5). Spectral Sciences, Inc.: Burlington, MA, USA. SSI-TR-685.
- Bernstein, L.S., Adler-Golden, S.M., Sundberg, R.L., Levine, R.Y., Perkins, T.C., Berk, A., Ratkowski, A.J., Felde, G. and Hoke, M.L., (2005). Validation of the QUick Atmospheric Correction (QUAC) algorithm for VNIR-SWIR multi-and hyperspectral imagery. In Defense and Security (pp. 668-678). International Society for Optics and Photonics.
- Bue, B. D., Thompson, D. R., Deshpande, S., Eastwood, M., Green, R. O., Natraj, V., ... & Parente, M. (2019). Neural network radiative transfer for imaging spectroscopy. *Atmospheric Measurement Techniques*, 12(4), 2567-2578.
- Frouin, R. J., Franz, B. A., Ibrahim, A., Knobelspiesse, K., Ahmad, Z., Cairns, B., ... & Huang, X. (2019). Atmospheric correction of satellite ocean-color imagery during the PACE era. *Frontiers in Earth Science*, 7, 145.
- Gao, B.-C., Heidebrecht, K. B., & Goetz, A. F. H. (1993). Derivation of scaled surface reflectances from AVIRIS data, *Remote Sensing of Environment*, 44, 165-178.
- Gao, B.-C., Montes, M. J., Davis, C. O., and Goetz, A. F. (2009). Atmospheric correction algorithms for hyperspectral remote sensing data of land and ocean. *Remote Sensing of Environment*, 113, S17-S24.
- ISOFIT: Imaging Spectrometer Optimal FITting. (2019) Public, open source repository available for view, download, comment and contributions. <http://github.com/isofit/isofit>

-
- Kruse, F. A. (1988). Use of airborne imaging spectrometer data to map minerals associated with hydrothermally altered rocks in the northern Grapevine Mountains, Nevada and California. *Remote Sensing of Environment*, 24, pp. 31–51.
- Kruse, F. A. (2004). Comparison of ATREM, ACORN, and FLAASH atmospheric corrections using low-altitude AVIRIS data of Boulder, CO. In *Summaries of 13th JPL Airborne Geoscience Workshop, Jet Propulsion Laboratory, Pasadena, CA*.
- Moran, M.S., Bryant, R., Thome, K., Ni, W., Nouvellon, Y., Gonzalez-Dugo, M.P., Qi, J. and Clarke, T.R., (2001). A refined empirical line approach for reflectance factor retrieval from Landsat-5 TM and Landsat-7 ETM+. *Remote Sensing of Environment*, 78(1-2), pp.71-82.
- Perkins, T., Adler-Golden, S., Matthew, M. W., Berk, A., Bernstein, L. S. Lee, J. and Fox, M., (2012). “Speed and accuracy improvements in FLAASH atmospheric correction of hyperspectral imagery“, *Opt. Engineering*, Vol. 51, 111707-1 -111707-7.
- Reinersman, P. N., K.L. Carder, R.F. Chen, (1998). Satellite-sensor calibration verification with the cloud-shadow method. *Applied Optics*, 37, pp. 5541–5549
- Richter, R., & Schläepfer, D. (2002). Geo-atmospheric processing of airborne imaging spectrometry data, Part 2: atmospheric/topographic correction, *International Journal of Remote Sensing*, 23(13), 2631-2649.
- Roberts, D. A., Yamaguchi, Y., & Lyon, R. (1986). Comparison of various techniques for calibration of AIS data, in Proceedings of the 2nd Airborne Imaging Spectrometer Data Analysis Workshop (G. Vane and A. F. H. Goetz, Eds.), *JPL Publication 86-35*, 21-30, Jet Propulsion Lab, Pasadena, CA.
- Rodgers, C. D. (2000). *Inverse Methods for Atmospheric Sounding: Theory and Practice*. World Scientific.
- Rothman, L. S. (2010). The evolution and impact of the HITRAN molecular spectroscopic database. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 111(11), 1565-1567.
- Thompson, D. R., Green, R. O., Keymeulen, D., Lundeen, S. K., Mouradi, Y., Nunes, D. C., Castaño, R. & Chien, S. A. (2014). Rapid spectral cloud screening onboard aircraft and spacecraft. *IEEE Transactions on Geoscience and Remote Sensing*, 52(11), 6779-6792.
- Thompson, D. R., Gao, B. C., Green, R. O., Roberts, D. A., Dennison, P. E., & Lundeen, S. R. (2015). Atmospheric correction for global mapping spectroscopy: ATREM advances for the HypIRI preparatory campaign. *Remote Sensing of Environment*, 167, 64-77.
- Thompson, D. R., Natraj, V., Green, R. O., Helmlinger, M. C., Gao, B. C., & Eastwood, M. L. (2018). Optimal estimation for imaging spectrometer atmospheric correction. *Remote Sensing of Environment*, 216, 355-373.
- Thompson, D. R., Guanter, L., Berk, A., Gao, B. C., Richter, R., Schläpfer, D., & Thome, K. J. (2019). Retrieval of atmospheric parameters and surface reflectance from visible and shortwave infrared imaging spectroscopy data. *Surveys in Geophysics*, 40(3), 333-360.
- Thompson, D. R., Cawse-Nicholson, K., Erickson, Z., Fichot, C. G., Frankenberg, C., Gao, B. C., ... & Thompson, A. (2019b). A unified approach to estimate land and water reflectances with uncertainties for coastal imaging spectroscopy. *Remote Sensing of Environment*, 231, 111198.
- Thompson, D. R., Babu, K. N., Braverman, A. J., Eastwood, M. L., Green, R. O., Hobbs, J. M., ... & Mathur, A. (2019c). Optimal estimation of spectral surface reflectance in challenging atmospheres. *Remote Sensing of Environment*, 232, 111258.
- Thompson, D. R., Babu, K. N., Braverman, A. J., Eastwood, M. L., Green, R. O., Hobbs, J. M., ... & Mathur, A. (2019c). Optimal estimation of spectral surface reflectance in challenging atmospheres. *Remote Sensing of Environment*, 232, 111258.
- Thompson, D. R., Braverman, A., ... & Wettergreen, D. S. (2020) Remote Spectroscopy of Surface Composition with Quantified Uncertainties. *Manuscript in Review*.

Thompson, D. R., D. Blaney, N. Bowles, B. H. Ehlmann, A. Fraeman, R. O. Green, R. Greenberger, R. Klima, Pantazis Mouralis, M. Sandford, C. Pieters, W. Williamson (2020b). On the information content of remote imaging spectroscopy for quantifying lunar water. *Manuscript in preparation*.