

# Rapid spectral cloud screening onboard aircraft and spacecraft

David R. Thompson, *Member, IEEE*, Robert O. Green, Didier Keymeulen, *Member, IEEE*, Sarah K. Lundein, Yasha Mouradi, Daniel Nunes, Rebecca Castaño, Steve A. Chien

**Abstract**—Next-generation orbital imaging spectrometers will generate unprecedented data volumes, demanding new methods to optimize storage and communications resources. Here we demonstrate that onboard analysis can excise cloud-contaminated scenes, reducing data volumes while preserving science return. We calculate optimal cloud screening parameters in advance, exploiting stable radiometric calibration and foreknowledge of illumination and viewing geometry. Channel thresholds expressed in raw instrument values can then be uploaded to the sensor where they execute in real time at Gigabit per second (Gb/s) data rates. We present a decision theoretic method for setting these instrument parameters, and characterize performance using a continuous three-year image archive from the “classic” Airborne Visible / Infrared Imaging Spectrometer (AVIRIS-C). We then simulate the system onboard the International Space Station (ISS), where it provides factor of two improvements in data volume with negligible false positives. Finally, we describe a real-time demonstration onboard the AVIRIS Next Generation (AVIRIS-NG) flight platform during a recent science campaign. In this blind test, cloud screening performed without error while keeping pace with instrument data rates.

**Index Terms**—Imaging Spectroscopy, Lossy Compression, Cloud Screening, Pattern Recognition, Real Time Systems

## I. INTRODUCTION

FUTURE NASA Earth science missions will face unprecedented data volumes. Data product sizes and production rates have increased steadily thanks to improvements in detector, optics, and onboard data handling technology. High resolution spectrometers such as NASA’s OCO-2 mission will yield over one million soundings per day [1]. Proposed imaging spectrometers such as HypSIRI [2] or an International Space Station (ISS) imaging spectrometer would generate data rates on the order of one Gigabit per second (Gb/s). These rates are a consequence of the full spectral measurement at high spatiotemporal resolution required for a range of unique science and application objectives [3]. However, the large data volumes affect mission requirements for the entire data handling chain including onboard digitization, storage, downlink, ground processing and distribution [4]. Bottlenecks along this path can constrain the instrument duty cycle, reducing science and application yield. In particular, bandwidth constraints have motivated new advanced lossless compression techniques such as the FL algorithm [5]–[8] that have achieved compression rates of four or greater. Efforts to optimize lossless methods eventually face theoretical limits, but data rates continue to

increase. The challenge has driven research into other techniques that can further reduce data volumes while preserving science yield.

One promising approach is to avoid storing or transmitting cloud-contaminated data [9], [10]. Historically clouds are estimated to cover 54% or more of the Earth’s land area and 68% or more of the oceans [11]–[13]. Many algorithms to estimate atmospheric or surface properties cannot function in the presence of condensed water/ice clouds. This makes more than half of visible to short wavelength (VSWIR) scenes in remote sensing archives unusable for their intended science and applications purpose [14]. Excising these scenes at the sensor can significantly reduce onboard storage and bandwidth requirements. However, the community lacks a practical algorithm capable of real-time execution in instrument hardware.

This paper addresses the need. We present a real-time cloud screening method that executes on raw sensor data for use onboard aircraft and spacecraft. We report its performance both in simulations and in a deployment on the Airborne Visible / Infrared Imaging Spectrometer - Next Generation (AVIRIS-NG). Our approach is designed for the unique requirements of real-time cloud screening, with rapid Gb/s execution rates and responsiveness to changing terrain and illumination conditions. It amounts to several simple channel thresholds, which are adapted dynamically to account for predicted brightness of clouds and terrain. Some inaccuracy is tolerable since any missed clouds can be excised later on the ground. It is operationally very simple to implement, and conservative settings ensure that good quality science data is preserved at all costs. We will demonstrate that it is possible to achieve data volume reductions near the theoretical maximum without any significant loss of science data.

### A. Prior work

We focus on the VSWIR electromagnetic spectrum from 0.4–2.5  $\mu\text{m}$ . Figure 1 shows an example scene from the “classic” Airborne Visible / Infrared Imaging Spectrometer (AVIRIS-C) with representative spectra of different materials and clouds. There are many studies of cloud detection in these wavelengths, and algorithms vary in their assumptions and complexity. “Classical” cloud screening applies threshold tests to spatial and spectral properties of the image [15]. Pixels whose values fall outside valid ranges are marked as cloudy. For example, the MODIS algorithm compares selected visible and near-infrared (VNIR) and near infrared (NIR) bands to predetermined thresholds, and then aggregates the result in

Authors are with the Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109 USA

Manuscript received 2013;

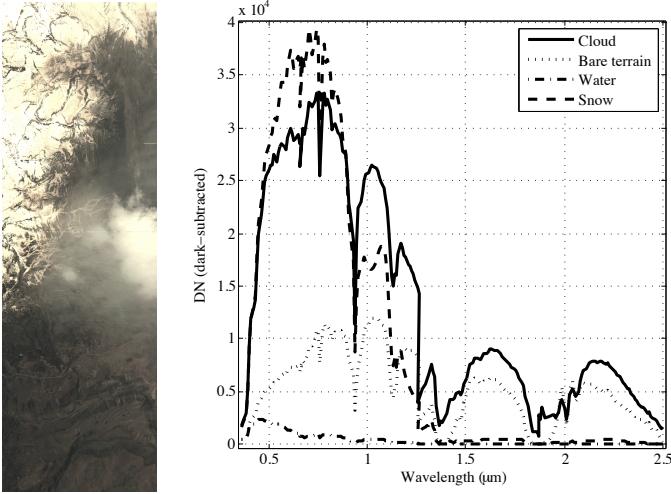


Fig. 1. Left: AVIRIS-C image f100521t02p05, a challenging scene which contains both clouds and snow. Right: Spectra from bare terrain, snow, open water, and clouds, in units of dark-subtracted instrument Digital Numbers (DNs).

different combinations depending on land type [16]–[18]. The algorithm uses a combination of 14 wavelengths and over 40 tests. This underscores the intrinsic difficulty of constructing a universal and complete cloud screening procedure.

Even more complex algorithms are possible. Some state of the art cloud screening techniques estimate the optical path from absorption features like the oxygen A band, as in Gómez-Chova et al. [19] or Taylor et al. [14]. Thermal IR channels can add brightness temperature information. Minnis et al. predict clear sky brightness temperature values using ambient temperature and humidity, and then excise pixels outside these intervals [20]. Texture cues can also be used to recognize clouds by their high spatial heterogeneity [21]. Martins et al. demonstrate that a simple spatial analysis – the standard deviation of VNIR isotropic reflectances in a  $3 \times 3$  pixel window – reliably discriminates clouds from aerosol plumes over ocean scenes [22]. Murtagh et al. represent spatial dependencies using a probabilistic Markov Random Field (MRF) prior [23]. Other efforts use special sensing modalities such as polarization [24].

Of direct relevance to this work, onboard cloud screening has been demonstrated onboard the EO-1 spacecraft [25]. The EO-1 cloud screening uses the solar zenith angle to compute the apparent Top of Atmosphere (TOA) reflectance. Then it applies a branching sequence of threshold tests based on carefully-crafted spectral ratios to distinguish clouds and bright landforms such as snow, ice, and desert sand. The EO-1 cloud detection also acts as a data filtering step prior to onboard flood and cryospheric classification [26], [27]. To our knowledge it is the only previous case of cloud-screening performed on orbit. Due to the limitations of the mission’s 12MHz flight computer, screening a  $1024 \times 256$  image requires about 30 minutes [28] which is three orders of magnitude lower than our desired processing rate. Nevertheless, the work provides an important proof of concept and a foundation for our study.

## B. Algorithm requirements

Previous systems try to screen all clouds to prevent contamination of later retrieval algorithms. In contrast, we aim to reduce the instrument data volume which leads to distinct requirements. Completeness is not critical since the end user can perform a more precise cloud screening later. Our algorithm can be conservative, abstaining from ambiguous classifications to prevent loss of science data. This requires some way to represent classification certainty. There is precedent; for example, Ishida et al. supplement their binary decision with a confidence score [29]. Gómez-Chova et al. use a Gaussian mixture model to produce posterior probabilities [19]. The Bayesian probabilistic model of Merchant et al. combines observational data with prior predictions from atmospheric forecasts, leading to true probabilistic predictions [30]. Rigorous probabilistic approaches are well-suited to an onboard algorithm that abstains from uncertain classifications.

Onboard cloud screening must also satisfy strict computational constraints. The algorithm must process *all* data collected by the spectrometer before it enters the flight recorder. In many cases this requires that the algorithm run *in instrument hardware* such as a Field Programmable Gate Array (FPGA), entailing additional design requirements [9]. For a pushbroom instrument, image lines arrive sequentially. Depending on the buffering strategy it may not be possible to pass more than a handful of values from one line to the next. This limits the use of spatial context. Moreover, classifiers instantiated in hardware logic typically forgo the use of exponentials, transcendental functions and even floating point operations, precluding many nonlinear classifiers and naïve implementations of linear classifiers. Additionally, the cloud screening should operate with Gb/s throughput, using a small fixed number of arithmetic operations on locally-available data, and have a deterministic computational path without recursion or iterative loops. This excludes many classifiers such as nearest-neighbor or decision tree algorithms. Finally, as a consequence of embedded hardware execution, the cloud screening must operate on *raw instrument data values*. This rules out most classical cloud screening algorithms, since it is not realistic to reproduce the ground-side processing which could provide calibrated reflectance as input.

This work presents a technique to satisfy these requirements. We demonstrate a very simple cloud screening algorithm that operates on raw instrument data, significantly reducing its volume while achieving a higher throughput rate than any previously-reported cloud screening system. As with EO-1, the screening decision is a sequence of threshold tests on selected wavelength values. However, these thresholds are recomputed before each observation using foreknowledge of scene parameters: the solar irradiance from orbital ephemeris, and instrument calibration and terrain properties from historical data. These define distributions of raw uncalibrated instrument values for cloud and terrain, which in turn prescribe channel thresholds. Operators update thresholds as often as needed to track changes in imaging conditions and geometry. This partitions the cloud screening calculations into an offline part that benefits from powerful computers and ancillary

meteorological information, and a fast real-time part suited to onboard execution and FPGA logic.

The following sections describe the algorithm's theoretical assumptions. We present a formal Bayesian probabilistic method for selecting thresholds. We then evaluate performance for different operations scenarios using a three year historical image archive of the "classic" Airborne Visible / Infrared Imaging Spectrometer (AVIRIS-C) [31]. A case study quantifies the compression benefits using orbital parameters of the International Space Station (ISS). Finally, we report the results of a field deployment onboard the AVIRIS Next Generation (AVIRIS-NG) airborne imaging spectrometer [32]. A cloud screening testbed was installed in parallel with the regular AVIRIS-NG data system, and operated without error during a recent science campaign.

## II. METHOD

Our cloud screening approach tests specific channels with user-defined thresholds. We will focus on VSWIR imaging spectrometer measurements and will refer to each spectrum as a pixel (i.e. a single image location with all wavelengths). In mathematical terms, a cloud screening algorithm must define an *exclusion region*  $\mathcal{R} \subseteq \mathbb{R}^d$ , a range of instrument data values for which a pixel is judged to be cloudy. The observed spectrum of instrument data forms a vector  $\mathbf{y}$  with multiple spectral channels per pixel. The cloud screening decision maps these pixel brightness values to a binary classification  $c = f(\mathbf{y}) : \mathbb{R}^d \mapsto \{c_1, c_2\}$ , where  $c_1$  represents the event that clear sky is observed and  $c_2$  that there is a cloud present. The corresponding decision rule is simply:

$$f(\mathbf{y}) = \begin{cases} c_1 & \text{if } \mathbf{y} \in \mathcal{R} \\ c_2 & \text{if } \mathbf{y} \notin \mathcal{R} \end{cases} \quad (1)$$

Here we define  $\mathcal{R}$  with a set of channel thresholds  $\phi$ , marking any pixel that exceeds all these thresholds as cloudy. Figure 2 shows the decision for a single channel. The vertical axis indicates probability density. We seek thresholds that best distinguish terrain classes from cloud pixels. Note that there is some overlap between the distributions, so in this one channel the populations are inseparable. Thus there will always be some unavoidable classification error.

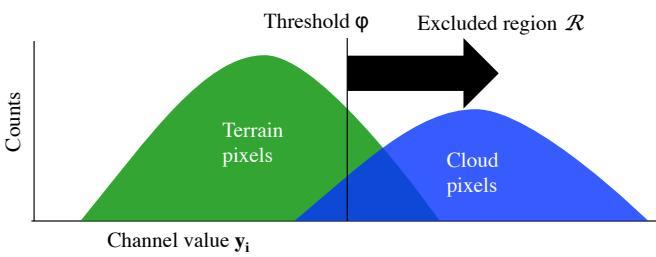


Fig. 2. Thresholds  $\phi$  define an exclusion region to classify pixels as cloudy.

These terrain and cloud brightness distributions depend on scene-specific factors such as land type, seasonal and snow cover effects, and illumination. Fortunately we can predict these factors to first order using historical data and observation geometry. We use the following sequential steps: (1)

in advance, determine the channels that will be used; (2) predict pixel brightness by extrapolation from historical data; (3) optimize channel thresholds to reflect data reduction and false alarm requirements; and (4) in real time, apply these thresholds to excise cloudy data.

### A. Channel selection

Clouds are bright across the ultraviolet (UV), visible wavelengths and Infrared (IR). However, cloud screening algorithms conventionally use only a small subset of the available channels. The MODIS cloud mask uses the  $0.659 \mu\text{m}$  reflectance channel. The  $0.865 \mu\text{m}$  channel, ratioed with  $0.659 \mu\text{m}$ , can identify clouds by their flat spectra [25].  $0.936 \mu\text{m}$  and  $0.940 \mu\text{m}$  channels discriminate low clouds and shadows, respectively. Additional tests on NIR channels at  $1.24 \mu\text{m}$  and  $1.65 \mu\text{m}$  help distinguish snow [33]. The  $1.38 \mu\text{m}$  band indicates cirrus clouds [34]. The channel lies in a  $H_2O$  absorption feature that is typically opaque due to water vapor in the lower troposphere, so large reflectance indicates a reflection from high altitude cirrus.

Postprocessing can often correct translucent clouds like high cirrus, so these can be considered "good data" for our purposes. We will focus exclusively on low, opaque clouds. This simplifies the problem considerably since opaque clouds are easiest to detect. Our approach can use arbitrarily many frequencies, but the following experiments use just two channels for clarity. A blue visible channel discriminates clouds from land and ocean, while a SWIR channel excludes snow and ice. Section III-B demonstrates that this pairing has the highest information content for our dataset.

### B. Estimation of cloud and surface appearances

This section describes models of cloud and terrain appearance that are used on the ground to predict pixel brightness distributions. Our method is similar to that of Merchant et al. [30] which represents explicit distributions of cloud appearances under different imaging conditions, atmospheric status and terrain type. A prior  $P(c_1)$  represents the known probability of observing clouds, which can be a historical average. A state variable  $\mathbf{x}$  represents known background conditions such as the surface type.

To set appropriate thresholds we must ultimately estimate  $P(\mathbf{y} | \mathbf{x}, c)$ , the conditional probabilities of pixel values for clouds and terrain. These uncalibrated instrument values are sensitive to variations in solar input due to observation geometry. We simplify the problem by estimating the related distribution  $P(\mathbf{z} | \mathbf{x}, c)$ , a normalized representation that removes the solar variability. We use *Top Of Atmosphere* (TOA) reflectance values  $\mathbf{z}$  that have been adjusted for the solar zenith angle  $\theta$  using:

$$\mathbf{z} = \frac{\pi d^2}{\cos(\theta)s} \mathbf{g}(\mathbf{y} - \mathbf{b}) \quad (2)$$

Here  $\mathbf{b} \in \mathbb{R}^d$  is a bias and  $\mathbf{g} \in \mathbb{R}^d$  is a gain that translates the measurement to radiance in  $\text{W}/\text{nm}/\text{sr}/\text{m}^2$ . The Earth-Sun distance  $d$  is a function of the Julian day. The value  $s \in \mathbb{R}^d$  the incident solar flux per channel, typically computed by

convolving a solar illumination model with the band spectral response. The resulting values  $\mathbf{z}$  are solar normalized but unitless brightness values decoupled from observation geometry. We accumulate the new values in multidimensional histograms to form  $P(\mathbf{z} \mid \mathbf{x}, c)$ , storing a different histogram for every distinct surface class  $\mathbf{x}$ . Each histogram has one dimension per spectral channel used in the test.

Operationally one may need to analyze an orbital segment spanning a range of solar angles and surface types. In this case  $P(\mathbf{y} \mid \mathbf{x}, c)$  combines each surface type in correct proportion. We perform a weighted summation of the appropriate pixel brightness distributions at each segment time step  $t \in T$ , simultaneously accounting for solar effects and transforming the solar normalized representations back to instrument data values. For a histogram this is a simple operation, with normalized bin coordinates  $\mathbf{z}$  related to new timestep-specific bin coordinates  $\mathbf{y}_t$  using:

$$\mathbf{y}_t \mid \mathbf{x}_t, c = \frac{1}{|T|} \sum_t \left[ (\mathbf{z} \mid \mathbf{x}_t, c) + \frac{\mathbf{b} \cos(\theta_t) \mathbf{s}}{g\pi d^2} \right] \quad (3)$$

Here  $\mathbf{x}_t$  and  $\theta_t$  refer to predictions of land type and solar zenith angle from orbital ephemeris. In the special case of models having two spectral channels, inverting solar normalization is tantamount to a simple affine transformation of a 2D image.

Figure 3 illustrates this process using a typical AVIRIS-C image (f100521t02p05). The top panel shows the original populations of background terrain and cloud pixels when imaged with a solar zenith angle of  $18^\circ$ . The vertical and horizontal axes show brightness values  $y$  for the  $0.45 \mu\text{m}$  and  $1.25 \mu\text{m}$  channels respectively. As an example we transform these data to predict the  $y$  values for a solar zenith angle of  $45^\circ$  (Bottom). Scene dimming is most obvious in the lobe corresponding to snow. Combining such histograms in proportion to the terrain type lets an analyst predict future pixel brightness for *any* anticipated mixture of terrain types and observation geometries.

### C. Threshold Selection

Our approach cannot predict future observations exactly; it only gives probability distributions over the brightness of future clouds and terrain. We account for this uncertainty in our thresholds using Bayesian decision theory [35]. Recall that  $\mathbf{y}$  is a vector of brightness values across several spectral channels, and that  $c_1$  and  $c_2$  are clear and cloudy cases respectively. We define a loss function with a penalty  $\alpha_{FN}$  for false negatives (clouds that pass the filter) and a separate penalty  $\alpha_{FP}$  for false positives (clear scenes that are wrongly excised). The total expected loss  $E[\mathcal{L}]$  accounts for both penalties:

$$E[\mathcal{L}] = \int_{\mathcal{R}} \alpha_{FP} P(c_1 \mid \mathbf{y}, \mathbf{x}) dy + \int_{\mathbb{R}^d \setminus \mathcal{R}} \alpha_{FN} P(c_2 \mid \mathbf{y}, \mathbf{x}) dy \quad (4)$$

This serves as a figure of merit; operators simply choose the channel threshold combination with the lowest expected loss.

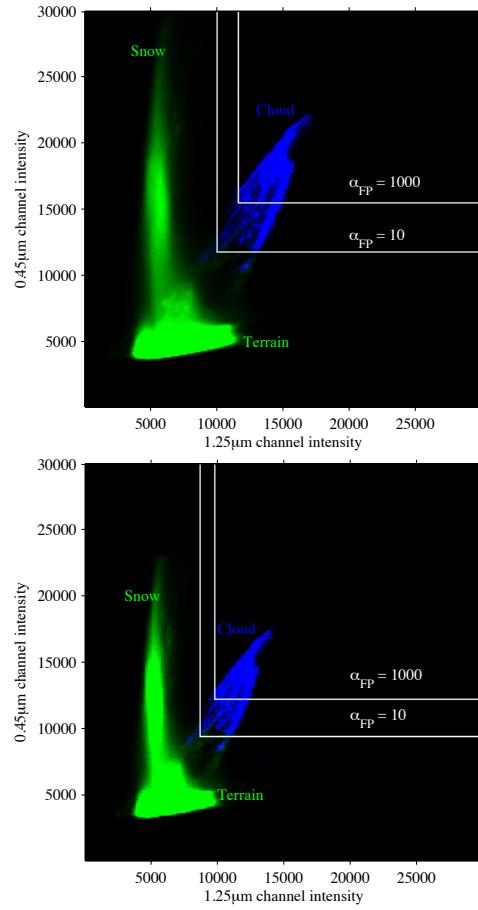


Fig. 3. Brightness distributions in  $0.45 \mu\text{m}$  and  $1.25 \mu\text{m}$  channels for AVIRIS-C image f100521t02p05, a scene which contains both clouds and snow. Top: Original image with a solar zenith angle of  $18^\circ$ . Bottom: synthetic distribution after transforming the image to a solar zenith angle of  $45^\circ$ . Optimal thresholds are shown for an aggressive cloud screening ( $\alpha_{FP} = 10$ ) and a conservative cloud screening ( $\alpha_{FP} = 1000$ ).

Following Merchant et al. [30], we rewrite the probability of the cloud case,  $c_1$ , using Bayes' rule:

$$P(c_1 \mid \mathbf{y}, \mathbf{x}) = \frac{P(\mathbf{y} \mid \mathbf{x}, c_1) P(\mathbf{x} \mid c_1) P(c_1)}{P(\mathbf{y} \mid \mathbf{x}) P(\mathbf{x})} \quad (5)$$

Assuming that the background state is independent of the cloud probability, we have  $P(\mathbf{x} \mid c_1) = P(\mathbf{x})$ . We ignore the  $P(\mathbf{y} \mid \mathbf{x})$  term which is the same for both cloudy and clear cases, leaving:

$$P(c_1 \mid \mathbf{y}, \mathbf{x}) \propto P(\mathbf{y} \mid \mathbf{x}, c_1) P(c_1) \quad (6)$$

The two possible cases are clouds  $c_1$  and clear sky  $c_2$ . This permits the following decomposition:

$$E[\mathcal{L}] = \int_{\mathcal{R}} \alpha_{FP} P(\mathbf{y} \mid \mathbf{x}, c_1) P(c_1) dy + \int_{\mathbb{R}^d \setminus \mathcal{R}} \alpha_{FN} P(\mathbf{y} \mid \mathbf{x}, c_2) P(c_2) dy \quad (7)$$

One can minimize this loss using any nonlinear optimization method appropriate for the chosen representation of  $P(\mathbf{y} \mid \mathbf{x}, c)$ . The proposed multidimensional histogram representation of  $P(\mathbf{y} \mid \mathbf{x}, c)$  permits a direct grid search; the integrated

expected loss is a cumulative sum, computable with a fast recursive operation. More generally, gradient descent could be used to find a locally-optimal threshold. Using  $\eta$  to denote the subspace excluding channel  $\nu$ , and  $\psi_\eta$  the set of points on the decision boundary in the subspace excluding channel  $\nu$ , the error gradient with respect to a specific threshold on channel  $\nu$  is:

$$\begin{aligned} \frac{dE[\mathcal{L}]}{d\phi_\nu} = & \int_{\psi_\eta} \alpha_{FP} P(\mathbf{y} | \mathbf{x}, c_1) P(c_1) d\mathbf{y}_\eta \\ & - \int_{\psi_\eta} \alpha_{FN} P(\mathbf{y} | \mathbf{x}, c_2) P(c_2) d\mathbf{y}_\eta \end{aligned} \quad (8)$$

This permits minimization using the Newton method or another gradient-based approach.

Figure 3 shows optimal decision boundaries for the test image with  $\alpha_{FN} = 1$ . The two thresholds correspond to a lenient case where  $\alpha_{FP} = 10$  and a strict case where  $\alpha_{FP} = 1000$ . The scene contains both clouds and snow; lobe of the background distribution corresponding to snow features has a high brightness in the  $0.45\mu\text{m}$  channel but a low brightness at  $1.25\mu\text{m}$ . Consequently the best decision boundary carves out a rectangular exclusion region  $\mathcal{R}$ . The optimal thresholds vary significantly depending on geometry. For the original image with the more lenient false negative penalty  $\alpha_{FN} = 10$ , they are 11800 and 10000 for the  $0.45\mu\text{m}$  and  $1.25\mu\text{m}$  channels respectively. For the stricter case of  $\alpha_{FN} = 1000$  they become 15500 and 11500, focusing on the fraction of cloud pixels that are completely unambiguous. In the dimmer scene the best thresholds are 9400 ( $0.45\mu\text{m}$ ) and 8700 ( $1.25\mu\text{m}$ ) at  $\alpha_{FN} = 10$ , moving to 12200 and 9800 at  $\alpha_{FN} = 1000$ .

After selection of optimal threshold values for a new observation, the flight hardware performs these tests once per pixel, designating any pixel that exceeds all thresholds as “cloudy.”

#### D. Spatial Aggregation

The pixel classification may mislabel isolated bright terrain pixels such as anthropogenic features, sun glint, or other scene clutter. Such localized errors can be addressed by spatial smoothing, with methods like the adjacency tests of the MODIS approach [16], spatial features [19], or even image segmentation [36]. Not all of these remedies are suited for real-time processing since instrument buffers can only store a small portion of the image at one time.

Here we evaluate a simple spatial aggregation method suited for real-time execution in instrument hardware. It operates on a small number of buffered lines simultaneously, and makes an aggregate decision about whether to keep or excise the block. A *spatial coverage threshold* determines the number of cloudy pixels that will cause a vertical block to be excised. This spatial aggregation gives additional resilience to small localized bright patches or single-pixel artifacts. We use a spatial coverage threshold between 25% and 50%. Our rationale is that good quality data can sometimes be recovered from images with less cloud cover than this amount, but images with cloud cover greater than 50% would almost never be used. Optionally, each block can be subdivided horizontally into two or more sub-blocks, with separate keep/reject decisions for each. This finer

```

procedure SETTHRESHOLDS( $\alpha_{FP}$ )
  for all timestep  $t \in \{1, \dots, T\}$  do
    calculate surface type  $\mathbf{x}_t$ 
    calculate solar zenith angle  $\theta_t$ 
    calculate  $P_t(\mathbf{y} | \mathbf{x}_t, c_1)$             $\triangleright$  via Equation 3
    calculate  $P_t(\mathbf{y} | \mathbf{x}_t, c_2)$             $\triangleright$  via Equation 3
  end for
   $P(\mathbf{y} | \mathbf{x}, c) = \frac{1}{|T|} \sum_t P_t(\mathbf{y} | \mathbf{x}_t, c)$ 
   $\phi = \operatorname{argmin}_\phi E[\mathcal{L}]$             $\triangleright$  via Equation 7
end procedure

```

Fig. 4. Algorithm for threshold selection, performed offline. Analysts must specify the false positive penalty  $\alpha_{FP}$

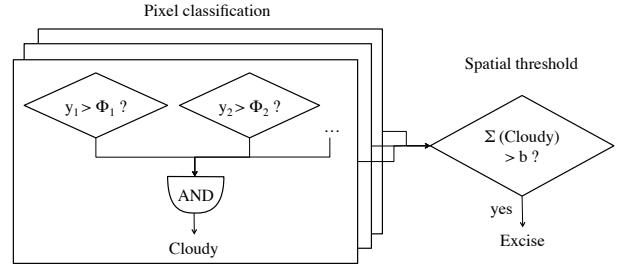


Fig. 5. Real-time cloud excision algorithm (onboard). Here  $b$  represents the minimum number of cloud pixels that trigger an excision.

spatial resolution can potentially preserve more of the good quality data near to clouds.

Figure 4 shows pseudocode for the threshold-setting procedure. Figure 5 shows the real-time portion as a block diagram with pixel-level and spatial aggregation thresholds. Figure 6 then illustrates the result for the image associated with the cloud and snow distributions presented previously in Figure 3. This is AVIRIS-C image f100521t02p00r05, acquired over Senator Beck Basin, Colorado, on May 21 2010. The left panel shows the original scene. The channel threshold operation labels cloudy pixels, shown as darkened areas in the middle panel. The right panel shows spatial aggregation which flags a segment of the image for excision.

### III. EVALUATION ON HISTORICAL AVIRIS-C DATA

We evaluated the method’s performance using the AVIRIS-C instrument’s 2009-2011 data archive [31]. In this period AVIRIS-C flew throughout North America on campaigns related to engineering and calibration, mineralogy, ecology, and disaster response. AVIRIS-C operators prefer to fly on clear days, which reduces the total fraction of cloud cover. However, many images contained clouds and this period provides a rich dataset to test the cloud screening approach. The AVIRIS-C VSWIR imaging spectrometer has  $0.01\mu\text{m}$  resolution in 224 bands from  $0.38\mu\text{m}$  to  $2.5\mu\text{m}$ . All AVIRIS-C images have 680 samples and an along-track dimension that ranges from several thousand to over ten thousand image lines.

#### A. Data Set

Much previous work measures cloud screening performance by correlating observations with dedicated cloud/aerosol sensors [14], [17], or with authoritative standards like the MODIS

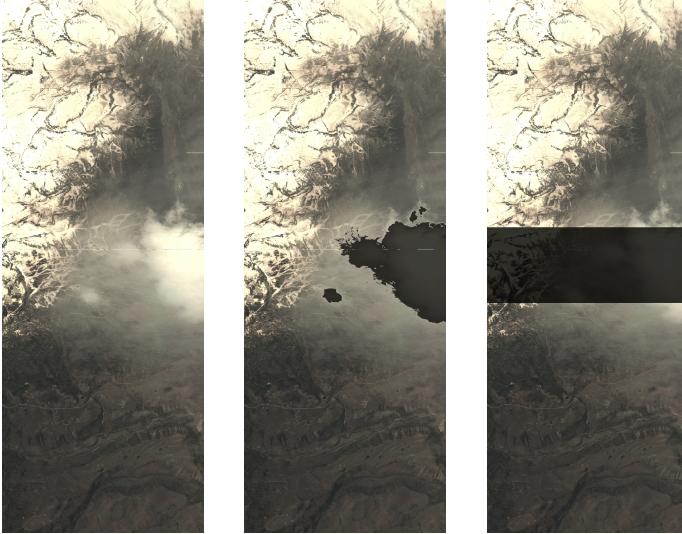


Fig. 6. AVIRIS-C image f100521t02p05, which contains both clouds and snow.

cloud mask [37]. In our case the AVIRIS-C spatiotemporal footprints rarely overlapped with other cloud sensors. Instead we evaluated performance by labeling of every cloud pixel in the archive by hand. We used only flights above 10000m altitude for maximum fidelity to orbital instruments. We inspected this entire catalog of raw uncalibrated instrument data, marking every pixel of each image manually as either “cloudy” or “clear” to form a ground truth cloud classification. We would typically perform this labeling in an image editor by first thresholding each image and then cleaning any misclassified points using a manual paint tool. We only labeled clouds that were opaque, i.e. that completely obscured the terrain color and texture underneath. This was a strict criterion, and occasionally left ambiguous translucent pixels around the edges of labeled clouds. We prevented edge pixels from contaminating the “clear” class during the evaluation by disregarding any terrain spectrum within 10 pixels of a cloud. Our rationale was that either classification would be reasonable for these ambiguous cases so they should not count for either credit or penalty. Such cases constituted a tiny fraction of the dataset, leaving plenty of data points for our evaluation.

We checked our dataset’s representativeness using typical land cover taxonomies. The 14-class University of Maryland (UMD) system [38] is attuned to biosphere and climate research, but its categories correlate with surface reflectance and can therefore bear on the observed spectral properties. We determined UMD categories for each image using onboard navigation telemetry to find the start location in GPS coordinates. We matched this location to land types recorded in the MODIS Collection 5 global land coverage products [39]. Table I shows the amount of data assigned to each category. Urban and cropland areas were particularly well represented. A large fraction of AVIRIS-C images were acquired over the Gulf Coast ocean due to extensive operations in this area related to the 2010 Gulf Oil spill. The land types differed from the global distributions but included instances from nearly all of the UMD land cover categories.

TABLE I  
LAND AND OCEAN COVER REPRESENTED IN AVIRIS-C DATASETS.

Land cover	Clear pixels	Cloudy pixels	Source
Water	$6.7 \times 10^8$	$2.2 \times 10^7$	UMD
Evergreen needleleaf forest	$3.7 \times 10^7$	$1.4 \times 10^6$	UMD
Evergreen broadleaf forest	$5.8 \times 10^6$	$5.8 \times 10^5$	UMD
Deciduous needleleaf forest	-	-	UMD
Deciduous broadleaf forest	$2.1 \times 10^8$	$4.4 \times 10^6$	UMD
Mixed forest	$1.3 \times 10^8$	$3.7 \times 10^6$	UMD
Closed shrublands	$1.1 \times 10^6$	$6.5 \times 10^5$	UMD
Open shrublands	$1.3 \times 10^8$	$1.8 \times 10^6$	UMD
Woody savannas	$1.3 \times 10^8$	$2.6 \times 10^6$	UMD
Savannas	-	-	UMD
Grasslands	$8.9 \times 10^7$	$3.1 \times 10^6$	UMD
Croplands	$1.0 \times 10^8$	$7.3 \times 10^6$	UMD
Urban and built-up	$3.1 \times 10^7$	$2.9 \times 10^4$	UMD
Snow and ice	$1.3 \times 10^8$	$4.4 \times 10^6$	
Barren	$9.7 \times 10^7$	$1.7 \times 10^2$	
Ocean glint	$3.6 \times 10^8$	$1.5 \times 10^7$	

We removed some specific and exceptional image features. Except where noted, we excluded images containing “sun glint” effects since they would not be used for most applications. We also excluded a set of images of the White Sands National Monument in Southern New Mexico, a feature composed of nearly pure gypsum that is highly reflective across all wavelengths. This feature is an unusual, if not globally unique, phenomenon [40]. Finally, we removed several scenes with opaque smoke from forest fires where the “correct” answer was ambiguous. This left a dataset of 507 images included in the study.

### B. Channel Selection

The AVIRIS-C dataset gave insight into the information provided to the cloud screening decision by different combinations of channels. As apparent from Figure 1, snow was highly reflective in visible wavelengths but dark in the SWIR. Conversely, bare terrain that was bright at SWIR wavelengths was significantly dimmer than clouds in VNIR ranges. This favored using channels in both regions.

*Mutual Information* (MI) indicates the information provided by different channel combinations. MI is a quantity from information theory describing the information value of an observation with respect to another unknown variable [41]. Specifically it quantifies how knowledge of one reduces the Shannon entropy in the other. In our case MI related the knowledge of frequency channels to the binary cloud/clear classification. We computed the MI using the following expression:

$$\begin{aligned} & \sum_{\mathbf{y} \in \mathcal{Y}} P(c_1, \mathbf{y}) \log \left( \frac{P(c_1, \mathbf{y})}{P(c_1)P(\mathbf{y})} \right) \\ & + \sum_{\mathbf{y} \in \mathcal{Y}} p(c_2, \mathbf{y}) \log \left( \frac{P(c_2, \mathbf{y})}{P(c_2)P(\mathbf{y})} \right) \end{aligned} \quad (9)$$

where  $\mathcal{Y}$  was the domain of  $\mathbf{y}$ . Tables II and III show scores for land (368 flightlines) and ocean (139 flightlines) respectively. For these datasets the  $0.45 \mu\text{m}$  channel was the strongest single indicator of clouds. Table IV shows scores for images

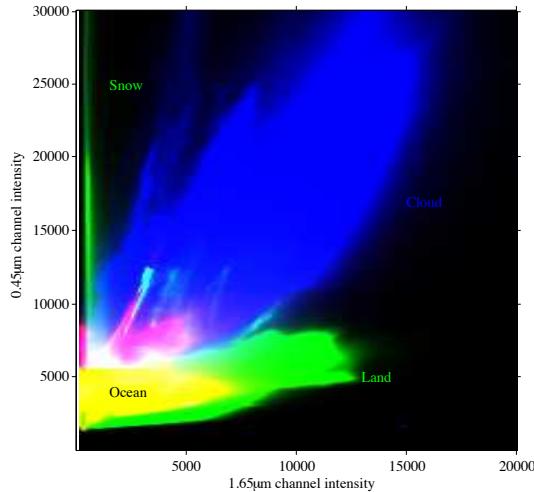


Fig. 7. Dark-subtracted brightness distributions at 0 degree solar zenith angle.

containing snow and ice (43 flightlines). The cirrus channel at  $1.38 \mu\text{m}$  was less valuable for this dataset because high-altitude translucent clouds were not counted. If we also sought to screen cirrus clouds this channel would have been more important.

While the  $0.45 \mu\text{m}$  channel had the highest overall MI, it was not obvious which other channel made its best partner. Combinations with SWIR wavelengths gave the largest net improvement, particularly for snow and ice scenes. The  $1.25 \mu\text{m}$  channel had been used previously for snow/ice tests in production systems such as the EO-1 screening algorithm [25], where it exploited the low SWIR absorption of snow as well as the low brightness of bare terrain in this spectral region. But for our system that considered multiple channels simultaneously, sensitivity to bare terrain provided no advantage. The  $0.45 \mu\text{m}$  channel already discriminated everything but snow, for which either SWIR frequency worked well. Consequently the MIs of the  $[0.45 \mu\text{m}, 1.25 \mu\text{m}]$  pairing and the  $[0.45 \mu\text{m}, 1.65 \mu\text{m}]$  pairing were basically indistinguishable. We return to this question in the following section, where simulations incorporating spatial information show a slight advantage to the  $1.65 \mu\text{m}$  channel.

Figure 7 shows the combined distributions for land, ocean, and cloud pixels, with values given as dark-subtracted Data Numbers translated to  $0^\circ$  solar zenith angle. The terrain distributions had axis-parallel lobes corresponding to barren terrain and snow. There is some mixing between the ocean and land distributions in this image. We labeled each scene with its dominant land cover class, so - for example - ocean scenes were occasionally contaminated by islands and shorelines. There was also some natural overlap due to the ambiguity between opaque cloud, which was labeled, and haze or thin semitransparent cloud, which was not.

Figure 8 compares cloud and surface brightness to predictions from the idealized  $\cos(\theta)$  illumination falloff. We dark-subtracted the raw instrument values and corrected them for earth-sun distance. We then binned the solar zenith angles in ten-degree increments and plotted the mean of the resulting

TABLE II  
MUTUAL INFORMATION OF SELECTED CHANNELS FOR CLOUD SCREENING OVER LAND SURFACES (LARGER IS BETTER). THE FIRST DATA COLUMN SHOWS THE INFORMATION PROVIDED BY A SINGLE CHANNEL TAKEN ALONE. OTHER ENTRIES SHOW THE VALUE OF TWO-CHANNEL COMBINATIONS. BOLD TEXT IDENTIFIES THE TOP-SCORING COMBINATION. REFERENCES: <sup>1</sup>ACKERMAN ET AL. [16]; <sup>2</sup>GAO ET AL. [34]; <sup>3</sup>FRIEDL ET AL. [39]; <sup>4</sup>GÓMEZ-CHOVA ET AL. [19].

Channel	MI	MI when combined with					Ref.
		$0.66 \mu\text{m}$	$0.86 \mu\text{m}$	$1.25 \mu\text{m}$	$1.38 \mu\text{m}$	$1.65 \mu\text{m}$	
$0.45 \mu\text{m}$	0.58	0.61	0.59	0.63	0.59	0.63	<sup>4</sup>
$0.66 \mu\text{m}$	0.48		0.54	0.55	0.51	0.55	<sup>1,3,4</sup>
$0.86 \mu\text{m}$	0.39			0.47	0.46	0.49	<sup>1,3,4</sup>
$1.25 \mu\text{m}$	0.38				0.47	0.45	<sup>1,2,3</sup>
$1.38 \mu\text{m}$	0.20					0.41	<sup>1,2,3</sup>
$1.65 \mu\text{m}$	0.26						<sup>1,3</sup>

TABLE III  
MUTUAL INFORMATION OF SELECTED CHANNELS FOR CLOUD SCREENING OVER OCEAN SURFACES (LARGER IS BETTER). REFERENCES ARE IN THE TABLE II CAPTION.

Channel	MI	MI when combined with					Ref.
		$0.66 \mu\text{m}$	$0.86 \mu\text{m}$	$1.25 \mu\text{m}$	$1.38 \mu\text{m}$	$1.65 \mu\text{m}$	
$0.45 \mu\text{m}$	0.61	0.63	0.64	0.63	0.63	0.62	<sup>4</sup>
$0.66 \mu\text{m}$	0.60		0.62	0.62	0.62	0.60	<sup>1,3,4</sup>
$0.86 \mu\text{m}$	0.46			0.52	0.48	0.50	<sup>1,3,4</sup>
$1.25 \mu\text{m}$	0.48				0.48	0.49	<sup>1,2,3</sup>
$1.38 \mu\text{m}$	0.10					0.48	<sup>1,2,3</sup>
$1.65 \mu\text{m}$	0.43						<sup>1,3</sup>

TABLE IV  
MUTUAL INFORMATION OF SELECTED CHANNELS FOR CLOUD SCREENING OVER SCENES CONTAINING SNOW AND ICE (LARGER IS BETTER). REFERENCES ARE IN THE TABLE II CAPTION.

Channel	MI	MI when combined with					Ref.
		$0.66 \mu\text{m}$	$0.86 \mu\text{m}$	$1.25 \mu\text{m}$	$1.38 \mu\text{m}$	$1.65 \mu\text{m}$	
$0.45 \mu\text{m}$	0.45	0.50	0.50	0.63	0.59	0.63	<sup>4</sup>
$0.66 \mu\text{m}$	0.44		0.51	0.62	0.59	0.62	<sup>1,3,4</sup>
$0.86 \mu\text{m}$	0.43			0.61	0.58	0.61	<sup>1,3,4</sup>
$1.25 \mu\text{m}$	0.54				0.60	0.58	<sup>1,2,3</sup>
$1.38 \mu\text{m}$	0.54					0.61	<sup>1,2,3</sup>
$1.65 \mu\text{m}$	0.40						<sup>1,3</sup>

image pixels in each zenith bin. Figure 8 shows the result, with brightness as a function of solar zenith angle, with a solid line indicating the best fitting cosine-proportional curves. There was some natural deviation from the ideal due to the finite data set size, diversity of terrain types, and diversity of zenith angles within each bin. However, the overall result corroborates the cosine solar zenith angle relationship for both cloud and land surfaces.

### C. Predictive thresholding

After forming 2D histograms we calculated thresholds for pairings of the  $0.45 \mu\text{m}$  channel with either the  $1.25 \mu\text{m}$  or  $1.65 \mu\text{m}$  channels. We separated the data into a *training* half used to set thresholds and a *test* half used for evaluation. We then transformed the training half to a TOA reflectance representation to construct surface and cloud brightness distributions. We calculated bias factors using an automatic closed-shutter pre- and post-calibration segment embedded in the raw image. These biases were estimated once on a single calibration image and then applied uniformly across the whole

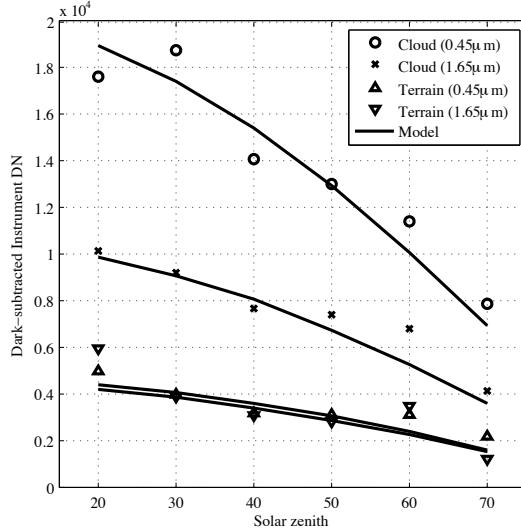


Fig. 8. Cloud and terrain brightness as a function of solar zenith angle, for all images over land. Instrument digital numbers have been corrected for dark current levels.

dataset, relying on the stability of the instrument as would be the case during extended autonomous operations.

We tested a range of operating regimes with different false alarm costs. Our most aggressive setting held false alarm costs equal to the false negative cost. This setting would maximize the compression rate; it would be appropriate for measurements such as large scale atmospheric profiling. At the other extreme, we set false alarm penalties to be a factor of  $10^5$  greater resulting in a very conservative cloud screening that preserved good data at all costs. This would be appropriate for scheduled acquisitions over specific high value targets. For each false alarm cost we used an exhaustive grid search to select the best channel thresholds. This gave a set of channel threshold pairs representing the envelope of optimal parameter settings.

Initially we calculated all thresholds in the normalized TOA representation. We applied them to each new test scene by translating back to the appropriate raw instrument values, transforming brightness according to the test scene's solar zenith angle. We then applied cloud screening with a 25% spatial coverage threshold. We evaluated several granularities of spatial aggregation, subdividing each along-track segment of 32 lines into one, two, or four cross-track blocks.

We also sought to determine the potential for further improvement by using more sophisticated classifiers. To that end, we evaluated the use of three channels in different combinations of VNIR and SWIR wavelengths. The three channel tests used the same solar normalization strategy with an exhaustive grid search to compute Bayes optimal thresholds. This shows the performance gains possible by growing the number of channels further. Still more sophisticated classifiers are possible - a linear decision boundary, suitably stripped of floating point arithmetic, arguably meets the real-time hardware requirements. This classifier calculates a prediction for each pixel independently by first applying the solar zenith correction and then forming a linear combination

of all channels. As before, we smooth the pixel-level classifications using a spatial aggregation test. We trained the linear decision boundary on the held out half of the data, sampling randomly from cloud and background distributions. Training used a stochastic gradient descent algorithm minimizing a sum squared error objective [42].

#### D. Evaluation

The simulation provided two performance values for each false alarm tolerance: a *false alarm rate* giving the fraction of good data accidentally deleted, and a *true positive rate* giving the fraction of cloudy image blocks that were successfully removed. We defined false negatives as any blocks containing greater than 50% cloud that were not excised. An excised block was a false positive if it did not contain any significant clouds (less than 5%). This left a range between 5-50% cloud cover where neither excision nor abstention were penalized. The free range was necessary since cloud edges were often ambiguous making the precise fractional coverage of small image areas somewhat subjective. We evaluated performance first using separate thresholds for land and ocean scenes, and then for a combined dataset which ignored surface type.

Finally, we simulated a hypothetical space mission to evaluate data compression rates. We calculated a year's orbits of the International Space Station at 10 minute time step intervals, recording at each time step the solar zenith angle and the terrain type under observation. Using the Eastman et al. heuristic [11], which was the lowest of recent estimates, we conservatively predicted that 68% of water scenes and 54% of land scenes would be cloudy. Given that instrument operators were able to set thresholds with foreknowledge of the surface type, we applied the appropriate land or ocean performance values in proportion to their appearance in each orbit. This gave the expected fraction of each orbit that could be excised for a desired false positive level. We assumed the instrument would be operating whenever the local solar zenith was less than 75°.

#### E. Results and Discussion

Figures 9-13 show cloud screening performance over land and ocean, as well as for the combined dataset that includes sun glint. We report results using a Receiver Operating Characteristic (ROC) curve plotting false alarm and true positive rates [43]. The ROC curve represents the envelope of performance that can be achieved; designers can move along it by changing the channel thresholds to be more or less aggressive. Desirable performance lies in the upper left, with many excised clouds and few false positives. We show performance for both channel combinations, with and without Solar Zenith Adjustment (SZA). The thick grey line in each plot marks the performance of our reference design, a two-channel cloud screening algorithm operating over land. Performance is consistent with work by Williams et al. on a threshold-based FPGA system combining VSWIR and thermal channels [9]; they report 0.02 – 0.09% missed clouds at a 0 – 0.99% false detection rate, which also lies on this ROC curve.

Our reference design uses the solar zenith correction. We found the correction nearly always improved performance, particularly in the case of the combined dataset including diverse land cover, ocean, and glint scenes. For the  $1.65\text{ }\mu\text{m}$  and  $0.45\text{ }\mu\text{m}$  channel combination at 0.1% false positive rates, the solar correction halved the false positive rate. However, performance was surprisingly good even without solar correction. We concluded that illumination was benign for most AVIRIS-C images in our catalog, and effects of solar variability were less important than large intrinsic differences in surface and cloud reflectance.

The experiments revealed minor differences between  $1.25\text{ }\mu\text{m}$  and  $1.65\text{ }\mu\text{m}$  channels. Many cloud screening algorithms favor the  $1.25\text{ }\mu\text{m}$  channel as an exclusion test for snow and ice, and we also found it effective for this purpose. The  $1.25\text{ }\mu\text{m}$  channel can also be used to exclude other dark terrain types. However, when combined with a  $0.45\text{ }\mu\text{m}$  channel this second role was completely redundant, so that either SWIR channel was an effective pairing. In fact, snow absorption was even stronger at  $1.65\text{ }\mu\text{m}$  consistent with studies by Painter et al. [44]. Consequently, the  $1.65\text{ }\mu\text{m}$  channel outperformed  $1.25\text{ }\mu\text{m}$  over land. All methods performed equally well over ocean. We also considered using three channels simultaneously and found that the combination of  $0.45\mu\text{m}$ ,  $1.25\mu\text{m}$  and  $1.65\mu\text{m}$  provided a slight additional benefit. Other triples failed to match the performance of the  $0.45\mu\text{m}$  and  $1.65\mu\text{m}$  pairing and are excluded from the plot. Adding a fourth channel might improve performance slightly more, though the histogram representation would be nearly a Terabyte in size.

A more practical approach is to incorporate more channels using the more traditional linear classifier. Figure 12 compares its performance to our reference design. The linear classifier performs best of all in the extreme low false positive regime, excising up to 97% of clouds with negligible loss of science data. This suggests that still better performance is possible if the instrument hardware can support the required multi-channel operations.

Finally, Figure 13 shows the performance for different spatial aggregation strategies. Subdividing the swath horizontally into smaller spatial blocks harmed classification performance since small, isolated bright terrain areas more easily triggered false alarms. As some compensation, each block was smaller and screening decisions were made at a fine spatial resolution. Therefore an excised cloudy block was less likely to contain clear pixels.

We used this result to estimate compression potential for future Earth-orbiting missions. Figure 14 shows compression performance based on land and ocean fractions observed during a year of International Space Station (ISS) orbits. We assumed cloud screening used the  $0.45\text{ }\mu\text{m}$  and  $1.65\text{ }\mu\text{m}$  channels. The horizontal axis shows the fraction of good data that was accidentally removed, which depends on the false alarm tolerance chosen by the designer. We considered false positive rates ranging from 0.001% (one such error every hundred thousand blocks, essentially negligible) to 0.01%. The data reduction achieved on each orbit then depends on the proportion of land and ocean encountered. We calculated the reduction using the performance estimates from our AVIRIS-

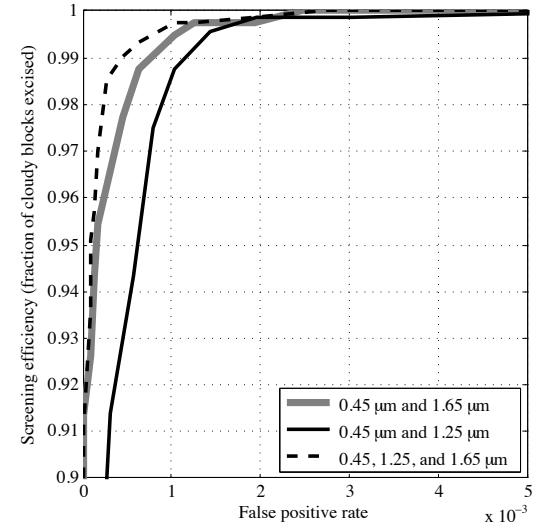


Fig. 9. Performance on AVIRIS-C 2009-2011 test scenes. We evaluated two channel combinations, with and without a Solar Zenith Adjustment (SZA).

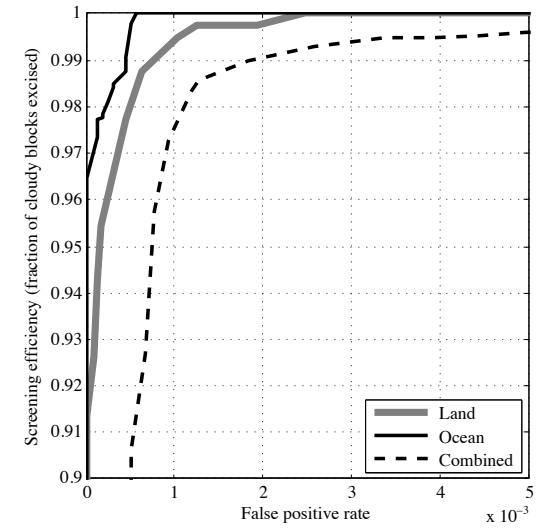


Fig. 10. Performance on AVIRIS-C 2009-2011 test scenes.

C study, presuming that the ground control could apply land or ocean threshold sets over the appropriate surface types.

The vertical axis shows the fraction of pixels removed. Boxes show the median and quartiles, with dashed whiskers indicating the extent of the most extreme orbits. At a false alarm rate of 0.001% cloud screening reduced data rates by approximately 60%. Performance never rose very far above this level, even for very lenient instrument settings, and the overall span of system performance was less than the natural cross-orbit variability. The strictest threshold achieved a rate reduction better than 90% of the theoretical optimum. This analysis excluded ocean scenes containing glint, so performance could be different for the special case of cloud screening during glint-mode atmospheric retrievals. Our analysis also ignored latitudinal differences in cloud cover and land type [13]. Northern latitudes pose the special challenge of the

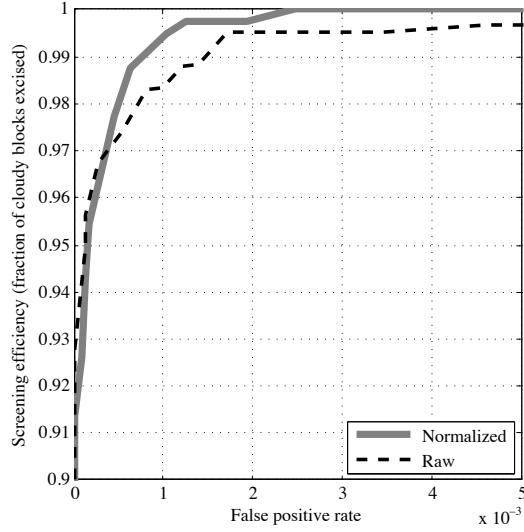


Fig. 11. Performance on AVIRIS-C 2009-2011 test scenes.

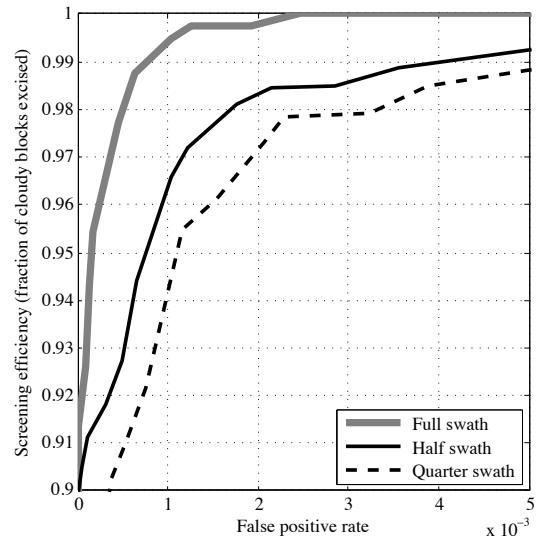


Fig. 13. Performance on AVIRIS-C 2009-2011 test scenes.

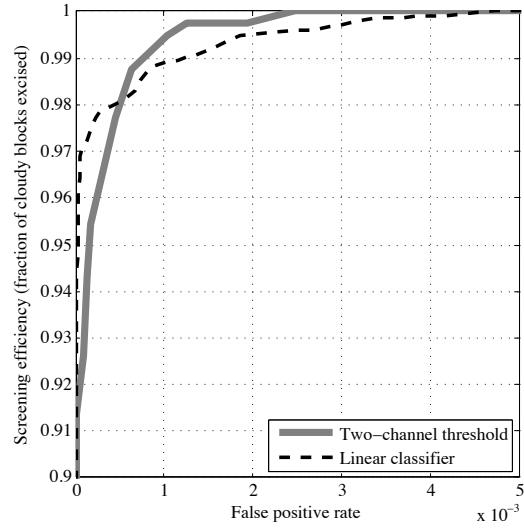


Fig. 12. Performance on AVIRIS-C 2009-2011 test scenes.

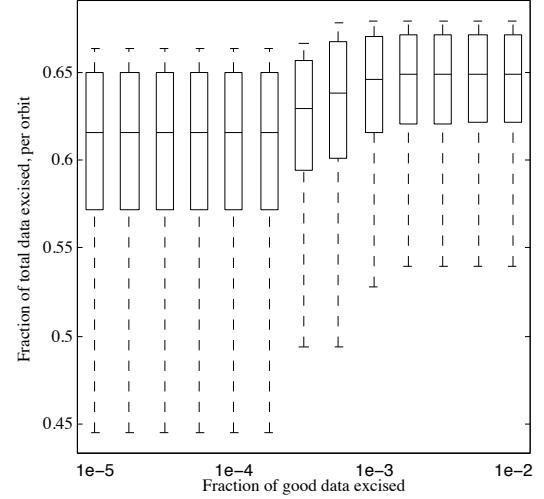


Fig. 14. Performance on ISS orbits, disregarding scenes with significant ocean glint.

finest grained snow, which is highly reflective [44], [45] and somewhat rare in this test dataset. If needed, missions could mitigate these effects by defining a specific new land cover type devoted to snow which would naturally lead to more conservative thresholds over such regions.

#### IV. AIRBORNE DEMONSTRATION

This section describes a deployment onboard the AVIRIS-NG airborne imaging spectrometer [32]. AVIRIS-NG is a next generation pushbroom instrument that measures the 0.38 – 2.5 micron region with 480 bands at 5nm spectral resolution. Its 640 cross track samples provide spatial resolutions of 1 meter or better, depending on altitude. Minor differences in spectral sampling should not affect the cloud screening since both clouds and bright terrain features are spectrally smooth in the wavelengths of interest.

The onboard cloud screening system used a commensal data path that ran in parallel alongside the science data acquisition and synchronization process. The primary data handling path used a Virtex-5 FPGA connected to the instrument through a Camera Link interface and to the IMU/GPS device through a serial and digital input interface. The commensal path and science data acquisition and synchronization were housed on a controller board based on a 2.3GHz quad-core Intel i7 processor from National Instruments. This processor was connected through a 64Gb/s PCI Express bus to a 1TB Solid State Drive (SSD) configured as RAID0 to achieve a read/write bandwidth of 6.4Gb/s throughput. The high throughput allowed cloud screening to access the SSD simultaneously with the processor's own acquisition and synchronization processes. The processor board, the FPGA board, and the SSD were integrated into a PXI chassis from National Instruments. A

GUI was designed to allow the user to control the data acquisition and synchronization process as well as the onboard cloud screening system.

### A. Procedure

We designed the system to implement the cloud screening algorithm, computing solar-normalized cloud thresholds in advance using the archive of historical AVIRIS-NG test flights. We used these data to calculate optimal thresholds for the  $0.45\text{ }\mu\text{m}$  and  $1.65\text{ }\mu\text{m}$  channels. When a flight began the real-time system monitored the filesystem for new data and immediately applied cloud screening. It started by first reading the current aircraft location from the Inertial Measurement Unit / Global Positioning System (IMU/GPS) information synchronized with the instrument data and stored with the images on the SSD. It calculated the solar ephemeris and adjusted cloud thresholds to the new observation geometry. The system then analyzed the image as it was recorded, producing two products: a real time operator notification displaying the average cloud fraction of the scene, and a *cloud mask* image recording pixel-level classifications for later analysis.

We evaluated the system during a twelve day science campaign over Casper, Wyoming, USA in June 2013. This campaign collected 23 flight segments comprising over a Terabyte of raw data. The images showed a mixture of bare terrain, industrial facilities, and vegetation. While the terrain content was diverse, science operations required clear skies so the data was not representative of global cloud and clear-sky distributions. However, in one flight when the ground observation sortie was cut short by clouds, the aircraft intentionally climbed from 2290m above mean-sea level (MSL) to 4420m MSL in order to fly above the rapidly growing cumuli.

### B. Results

Soon after reaching the desired altitude, the software's cloud-fraction display abruptly departed from 0.0, where it had consistently remained during the earlier portion of the flight. The cloud-fraction rapidly climbed to 1.0 while flying over the denser portions of the clouds. Post-analysis revealed that the system had successfully identified the opaque clouds and, when compared with a human interpretation, had labeled the correct segments of the flight line for excision. Figure 15 shows the result: the left panel is the original image, the center panel shows the pixels exceeding both thresholds, and the right panel shows the excised image blocks.

Table V summarizes the entire campaign. Columns show the time and the number of images in each batch, the total number of image lines, the fraction of pixels that were flagged as cloudy prior to spatial aggregation, and the final cloud screening result - whether any lines were excised and whether clouds were actually present. In cloudless scenes, isolated structures and bright objects occasionally exceeded both channel thresholds. This occurred for a fraction of pixels from 0.02% to 0.007%. However, these small pixel areas were successfully ignored during the spatial aggregation step, so the onboard system passed all clear-sky images without modification and committed no false excisions. The overall

result was that cloud screening performed without error and kept pace with the sensor's 0.5 Gb/s data production rate throughout the campaign. It will continue to be deployed on future flights.

TABLE V  
AVIRIS-NG FLIGHT CAMPAIGN RESULTS. THE *lines* COLUMN INDICATES THE NUMBER OF IMAGE LINES IN EACH GROUP. THE *pixel fraction* COLUMN INDICATES THE PROPORTION OF INDEPENDENT PIXELS THAT WERE MARKED AS CLOUDY DURING THE INITIAL THRESHOLDING. NONE OF THESE RESULTED IN ANY FALSE POSITIVE EXCISIONS.

Date	Time	images	lines	pixel fraction	excised	clouds
14 June 2013	14:43:44	1	960	0	N	N
14 June 2013	22:52:28	4	186560	$7.38 \times 10^{-5}$	N	N
18 June 2013	20:25:42	2	4480	0	N	N
18 June 2013	21:59:05	2	2880	0	N	N
18 June 2013	22:59:41	3	70720	$1.75 \times 10^{-3}$	N	N
18 June 2013	23:18:56	1	32000	$1.45 \times 10^{-3}$	N	N
18 June 2013	23:26:57	2	36480	$1.27 \times 10^{-3}$	N	N
18 June 2013	23:36:12	1	24000	$1.12 \times 10^{-3}$	N	N
19 June 2013	15:56:23	1	4800	0	N	N
19 June 2013	18:23:42	5	60480	$1.55 \times 10^{-4}$	N	N
19 June 2013	18:57:54	8	257920	$6.58 \times 10^{-4}$	N	N
20 June 2013	15:45:15	13	508160	$3.01 \times 10^{-4}$	N	N
20 June 2013	18:34:10	4	177280	$3.36 \times 10^{-4}$	N	N
21 June 2013	17:32:09	7	224000	$4.96 \times 10^{-4}$	N	N
21 June 2013	18:44:38	1	960	0	N	N
21 June 2013	18:46:57	9	333440	$3.14 \times 10^{-4}$	N	N
23 June 2013	17:01:46	3	108800	$1.76 \times 10^{-4}$	N	N
23 June 2013	17:45:53	13	448320	$2.09 \times 10^{-4}$	N	N
24 June 2013	16:00:40	5	147840	$3.00 \times 10^{-4}$	N	N
24 June 2013	17:01:04	10	221120	$2.06 \times 10^{-3}$	N	N
24 June 2013	17:56:11	5	158080	$1.41 \times 10^{-3}$	N	N
25 June 2013	15:57:28	5	153600	$3.94 \times 10^{-4}$	N	N
25 June 2013	16:49:28	7	174720	$6.32 \times 10^{-3}$	Y	Y
Total		142	3847266			

### V. CONCLUSIONS

This work describes a novel method for cloud screening onboard spacecraft at Gb/s data rates. We perform the most challenging computations on the ground, exploiting foreknowledge of observation geometry and surface type to predict the brightness of terrain pixels and cloud pixels. We calculate optimal thresholds in uncalibrated instrument values that can be uploaded for real-time execution by the flight system. Analysis on a three-year archive of AVIRIS-C images demonstrates that this simple approach can reduce instrument data volumes by a factor of two with insignificant loss of science data. A deployment onboard the AVIRIS-NG platform corroborates this performance in blind real time demonstrations. The testbed system will continue to accumulate additional operational experience in future AVIRIS-NG campaigns.

The excised data is, by definition, unrecoverable. This is a problem for investigations that are wholly intolerant of data loss, or that study the clouds themselves. However, having implemented the screening option, designers would be free to choose the threshold level most appropriate to the science needs of each observation. This contrasts with the status quo, where there is no such option and missions cannot escape the resource cost of clouds. Our tests suggest that high rates of

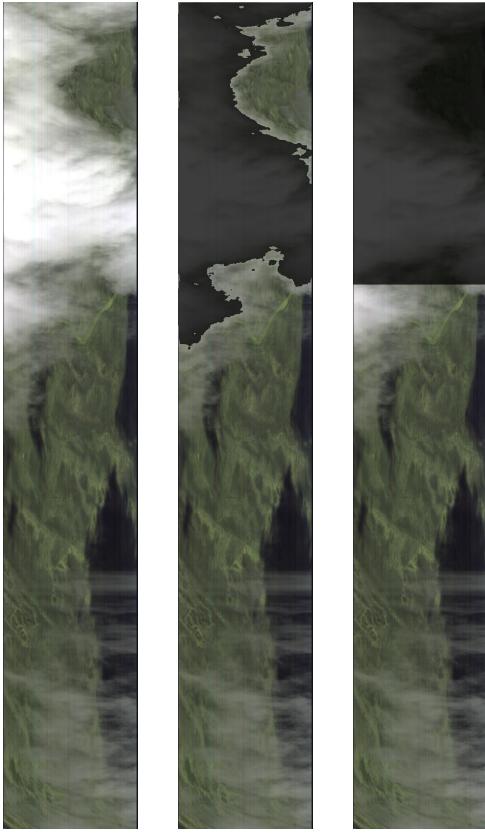


Fig. 15. AVIRIS-NG image ang20130625t174216. The images at center and right show the result of onboard real time cloud screening.

data reduction are achievable for very conservative settings. For an Earth-orbiting spectrometer, trivial cases such as open ocean yield most of the benefit, and very strict thresholds can excise most clouds over land terrain. We anticipate the highest risk of false alarms for scenes with bright snow or sun glint. As compensation, snow can be segregated as a separate terrain type and tracked over time using snow cover products from other missions. As an extreme measure, one could simply turn cloud screening off to avoid seasonal areas where snow could appear. Sun glint can be anticipated from imaging geometry. However, many investigations consider sun glint to be a contaminant similar to clouds.

The specific operational concept may vary depending on mission needs. Our prototype consists of a realtime component requiring streaming Gb/s processing, and an offline component having effectively no restrictions on computation or communication. While this is an idealized view, it is appropriate for real space operations scenarios. The offline computing needs would not be a significant bottleneck, since calculations take just a few minutes on a modern laptop computer. High Performance Computing (HPC) resources would make it a real-time operation. Nor is irregular communications an obstacle, since threshold settings could be computed long in advance based on the known observation geometry. Finally, one could always compromise computing or communication by limiting the number of threshold updates and applying the same thresholds over longer time segments spanning multiple

terrain types and solar zenith angles. More exotic operational concepts are possible. If the spacecraft or instrument can be pointed, the spectrometer could select targets based on the cloud screening result - perhaps scanning across its field of regard until a suitable clear scene is found.

There are several straightforward ways to improve accuracy. One could use three or more channels in the threshold decision. However, it is likely that more channels will offer diminishing returns. They would also require additional histogram dimensions, which quickly becomes intractable as the number of channels increases. Future work could seek alternative representations that scale better with dimensionality. A more promising approach would be to incorporate additional domain knowledge into the state vector. One could condition thresholds on very specific land types or on real-time cloud products like the GOES cloud mask. Preprocessing and feature extraction could also improve performance. For example, one could compute spectral derivatives; sums, differences, or ratios of channels; or continuum-relative absorption band depths [46]. Such spectral features could potentially improve results at a low computational cost.

One could always design a more complex cloud classifier to disambiguate the most difficult pixels and consequently achieve a slight improvement in data volume reduction. However, our simple algorithm already achieves better than 90% of the theoretical maximum making it sufficient for many applications and a useful point on the design trade space. At a time when communications and storage subsystems struggle under increasing data rates, the potential for onboard cloud screening has remained relatively unstudied. Mission designers should bear in mind that a few simple design considerations - the introduction of channel and aggregation thresholds - can enable factor-of-two reductions in data volume.

#### ACKNOWLEDGMENT

We thank Charles Sarture and the AVIRIS team for vital assistance in performing the airborne demonstrations. We thank Drs. Pantazis Mouralis, Jason Hyon and the JPL Earth Science Directorate for their support. The Airborne Visible / Infrared Imaging Spectrometer is operated by JPL with support from the National Aeronautics and Space Administration. The MODIS land cover data were obtained through the online Data Pool at the NASA Land Processes Distributed Active Archive Center (LP DAAC), USGS/Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota ([https://lpdaac.usgs.gov/get\\_data](https://lpdaac.usgs.gov/get_data)). Copyright 2013 California Institute of Technology. All Rights Reserved. U.S. Government support acknowledged.

#### REFERENCES

- [1] D. Crisp, R. Atlas, F.-M. Breon, L. Brown, J. Burrows, P. Ciais, B. Connor, S. Doney, I. Fung, D. Jacob et al., "The Orbiting Carbon Observatory (OCO) Mission," *Advances in Space Research*, vol. 34, no. 4, pp. 700–709, 2004.
- [2] R. O. Green, G. P. Asner, S. G. Ungar, and R. Knox, "Results of the decadal survey Hyperspectral imaging spectrometer concept study: a high signal-to-noise ratio and high uniformity global mission to measure plant physiology and functional type." in *Proceedings of IEEE Geoscience and Remote Sensing Symposium (IGARSS)*, 2008.

- [3] National Research Council, *Earth science and applications from space: National imperatives for the next decade and beyond*. Washington D.C.: National Academies Press, 2007.
- [4] L. Mandrake, C. Frankenberg, C. W. O'Dell, G. Osterman, P. Wennberg, and D. Wunch, "Semi-autonomous sounding selection for oco-2," *Atmospheric Measurement Techniques Discussions*, vol. 6, no. 3, pp. 5881–5922, 2013.
- [5] M. Klimesh, "Low-complexity lossless compression of hyperspectral imagery via adaptive filtering," *IPN Progress Report*, vol. 42–163, pp. 1–10, 2005.
- [6] M. Klimesh, A. Kiely, and P. Yeh, "Fast lossless compression of multispectral and hyperspectral imagery," *Proc. 2nd International Workshop on Onboard Payload Data Compression*, pp. 1–8, 2010.
- [7] N. Aranki, A. Bakhshi, D. Keymeulen, and M. Klimesh, "Fast and adaptive lossless on-board hyperspectral data compression system for space applications," *IEEE Aerospace Conference*, pp. 1–8, 2009.
- [8] D. Keymeulen, N. Aranki, B. Hobson, A. Kiely, M. Klimesh, and K. Benkrid, "Gpu lossless hyperspectral data compression system for space applications," *IEEE Aerospace Conference*, 2012.
- [9] J. A. Williams, A. S. Dawood, and S. J. Visser, "Fpga-based cloud detection for real-time onboard remote sensing," *Proceedings of the IEEE International Conference on Field-Programmable pp. 110–116, 2002.*
- [10] E. El-Araby, M. Taher, T. El-Ghazawi, and J. Le Moigne, "Prototyping automatic cloud cover assessment (ACCA) algorithm for remote sensing on-board processing on a reconfigurable computer," pp. 207–214, 2005.
- [11] R. Eastman, S. Warren, and C. Hahn, "Variations in cloud cover and cloud types over the ocean from surface observations, 1954–2008," *Journal of Climate*, vol. 24, no. 22, pp. 5914–5934, 2011.
- [12] M. Mercury, R. Green, S. Hook, B. Oaida, W. Wu, A. Gunderson, and M. Chodas", "Global cloud cover for assessment of optical satellite observation opportunities: A HypIRI case study," *Remote Sensing of Environment*, vol. 126, no. 0, pp. 62–71, 2012.
- [13] M. King, S. Platnick, W. Menzel, S. Ackerman, and P. Hubanks, "Spatial and temporal distribution of clouds observed by MODIS onboard the Terra and Aqua satellites," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 7, pp. 3826–3852, 2013.
- [14] T. Taylor, C. O'Dell, D. O'Brien, N. Kikuchi, T. Yokota, T. Nakajima, H. Ishida, D. Crisp, and T. Nakajima, "Comparison of cloud-screening methods applied to GOSAT near-infrared spectra," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 1, pp. 295–309, 2012.
- [15] M. Wang and W. Shi, "Cloud masking for ocean color data processing in the coastal regions," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 11, pp. 3196–3105, Nov 2006.
- [16] S. A. Ackerman, K. I. Strabala, W. P. Menzel, R. A. Frey, C. C. Moeller, and L. E. Gumley, "Discriminating clear sky from clouds with MODIS," *Journal of Geophysical Research: Atmospheres*, vol. 103, no. D24, pp. 32 141–32 157, 1998.
- [17] R. A. Frey, S. A. Ackerman, Y. Liu, K. I. Strabala, H. Zhang, J. R. Key, and X. Wang, "Cloud detection with MODIS. Part I: Improvements in the MODIS cloud mask for collection 5," *Journal of Atmospheric and Oceanic Technology*, vol. 25, no. 7, pp. 1057–1072, 2008.
- [18] S. Ackerman, R. Holz, R. Frey, E. Eloranta, B. Maddux, and M. McGill, "Cloud detection with MODIS. Part II: validation," *Journal of Atmospheric and Oceanic Technology*, vol. 25, no. 7, pp. 1073–1086, 2008.
- [19] L. Gómez-Chova, G. Camps-Valls, J. Calpe-Maravilla, L. Guanter, and J. Moreno, "Cloud-screening algorithm for ENVISAT/MERIS multispectral images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 12, pp. 4105–4118, 2007.
- [20] P. Minnis, Q. Z. Trepte, S. Sun-Mack, Y. Chen, D. R. Doelling, D. F. Young, D. A. Spangenberg, W. F. Miller, B. A. Wielicki, R. R. Brown et al., "Cloud detection in nonpolar regions for CERES using TRMM VIRS and terra and aqua MODIS data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 11, pp. 3857–3884, 2008.
- [21] J. Lee, R. Weger, S. Sengupta, and R. Welch, "A neural network approach to cloud classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 28, no. 5, pp. 846–855, sep 1990.
- [22] J. V. Martins, D. Tanré, L. Remer, Y. Kaufman, S. Mattoo, and R. Levy, "Modis cloud screening for remote sensing of aerosols over oceans using spatial variability," *Geophysical Research Letters*, vol. 29, no. 12, p. 8009, 2002.
- [23] F. Murtagh, D. Barreto, and J. Marcello, "Decision boundaries using Bayes factors: the case of cloud masks," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 12, pp. 2952–2958, 2003.
- [24] F.-M. Bréon and S. Colzy, "Cloud detection from the spaceborne POLDER instrument and validation against surface synoptic observations," *Journal of Applied Meteorology*, vol. 38, no. 6, pp. 777–785, 1999.
- [25] M. K. Griffin, H. K. Burke, D. Mandl, and J. Miller, "Cloud cover detection algorithm for EO-1 hyperion imagery," *17th SPIE AeroSense 2003, Orlando FL, Conference on Algorithms and Technologies for Multispectral, Hyperspectral and Ultraspectral Imagery IX*, 2003.
- [26] F. Ip, J. Dohm, V. Baker, T. Doggett, A. Davies, R. Castao, S. Chien, B. Cichy, R. Greeley, R. Sherwood, D. Tran, and G. Rabideau, "Flood detection and monitoring with the Autonomous Sciencecraft Experiment onboard EO-1," *Remote Sensing of Environment*, vol. 101, Issue 4, pp. 463–481, 2006.
- [27] T. Doggett, R. Greeley, S. Chien, B. Cichy, A. Davies, G. Rabideau, R. Sherwood, D. Tran, V. Baker, J. Dohm, and F. Ip, "Autonomous on-board detection of cryospheric change," *Remote Sensing of Environment*, vol. 101, pp. 447–462, 2006.
- [28] D. Mandl, M. Griffin, H. Burke, and J. Miller, "Validation of on-board cloud cover assessment using EO-1," in *ESTO Technology Conference, University of Maryland, College Park, MD*, 2003.
- [29] H. Ishida and T. Y. Nakajima, "Development of an unbiased cloud detection algorithm for a spaceborne multispectral imager," *Journal of Geophysical Research: Atmospheres*, vol. 114, no. D7, 2009.
- [30] C. Merchant, A. Harris, E. Maturi, and S. MacCallum, "Probabilistic physically based cloud screening of satellite infrared imagery for operational sea surface temperature retrieval," *Quarterly Journal of the Royal Meteorological Society*, vol. 131, no. 611, pp. 2735–2755, 2005.
- [31] R. O. Green, M. L. Eastwood, C. M. Sarture, T. G. Chrien, M. Aronsson, B. J. Chippendale, J. A. Faust, B. E. Pavri, C. J. Chovit, M. Solis et al., "Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (AVIRIS)," *Remote Sensing of Environment*, vol. 65, no. 3, pp. 227–248, 1998.
- [32] L. Hamlin, R. Green, P. Mouroulis, M. Eastwood, I. McCubbin, D. Wilson, D. Randall, M. Dudik, and C. Paine, "Imaging Spectrometer Science Measurements for Terrestrial Ecology: AVIRIS and the Next General AVIRIS Characteristics and Development Status," *NASA Earth Science Technology Forum (ESTF2010)*, 2010.
- [33] V. V. Salomonson and I. Appel, "Estimating fractional snow cover from modis using the normalized difference snow index," *Remote Sensing of Environment*, vol. 89, no. 3, pp. 351–360, 2004.
- [34] B. Gao, P. Yang, W. Han, R. Li, and W. Wiscombe, "An algorithm using visible and 1.38- $\mu\text{m}$  channels to retrieve cirrus cloud reflectances from aircraft and satellite data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 8, pp. 1659–1668, 2002.
- [35] C. M. Bishop and N. M. Nasrabadi, *Pattern Recognition and Machine Learning*. New York: Springer, 2006, vol. 1.
- [36] D. Thompson, B. Bornstein, S. Chien, S. Schaffer, D. Tran, B. Bue, R. Castano, D. Gleeson, and A. Noell, "Autonomous spectral discovery and mapping onboard the EO-1 spacecraft," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 6, pp. 3567–3579, 2013.
- [37] T. Y. Nakajima, T. Tsuchiya, H. Ishida, T. N. Matsui, and H. Shimoda, "Cloud detection performance of spaceborne visible-to-infrared multispectral imagers," *Appl. Opt.*, vol. 50, no. 17, pp. 2601–2616, Jun 2011.
- [38] M. C. Hansen, R. S. Defries, J. R. G. Townshend, and R. Sohlberg, "Global land cover classification at 1 km spatial resolution using a classification tree approach," *International Journal of Remote Sensing*, vol. 21, pp. 1331–1364, 2000.
- [39] M. A. Friedl, D. Sulla-Menashe, B. Tan, A. Schneider, N. Ramankutty, A. Sibley, and X. Huang, "MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets," *Remote Sensing of Environment*, vol. 114, no. 1, pp. 168–182, 2010.
- [40] J. D. Lindberg and M. S. Smith, "Reflectance spectra of gypsum sanda from the White Sands National Monument and basalt from a nearby lava flow," *American Mineralogist*, vol. 58, pp. 1062–1064, 1973.
- [41] D. J. MacKay, *Information Theory, Inference and Learning Algorithms*. Cambridge University press, 2003.
- [42] J. Langford, L. Li, and A. Strehl, "Vowpal wabbit open source project," *Technical report, Yahoo!*, 2007.
- [43] A. P. Bradley, "The use of the area under the ROC curve in the evaluation of machine learning algorithms," *Pattern Recognition*, vol. 30, no. 7, pp. 1145–1159, 1997.
- [44] T. H. Painter, J. Dozier, D. A. Roberts, R. E. Davis, and R. O. Green, "Retrieval of subpixel snow-covered area and grain size from imaging spectrometer data," *Remote Sensing of Environment*, vol. 85, no. 1, pp. 64–77, 2003.

- [45] T. H. Painter, F. C. Seidel, A. C. Bryant, S. M. Skiles, and K. Rittger, “Imaging spectroscopy of albedo and radiative forcing by light absorbing impurities in mountain snow,” *Journal of Geophysical Research: Atmospheres*, vol. doi:10.1002/jgrd.50520, 2013.
- [46] V. Carrere and J. E. Conel, “Recovery of atmospheric water vapor total column abundance from imaging spectrometer data around 940 nm—sensitivity analysis and application to airborne visible/infrared imaging spectrometer (aviris) data,” *Remote Sensing of Environment*, vol. 44, no. 2, pp. 179–204, 1993.