

# Atmospheric inversion in the presence of clouds: an adaptive ELM approach.

Brent Bartlett and John R. Schott

Rochester Institute of Technology, Rochester, NY, USA

## ABSTRACT

Many algorithms exist to invert airborne imagery from units of either radiance or sensor specific digital counts to units of reflectance. These compensation algorithms remove unwanted atmospheric variability allowing objects on the ground to be analyzed. Low error levels in homogenous atmospheric conditions have been demonstrated. In many cases however, clouds are present in the atmosphere which introduce error into the inversion at unacceptable levels. For example, the relationship that is defined between sensor reaching radiance and ground reflectance in a cloud free scene will not be the same as in a scene with clouds. A novel method has been developed which utilizes ground based measurements to modify the empirical line method (ELM) approach on a per-pixel basis. A physics based model of the atmosphere is used to generate a spatial correction for the ELM. Creation of this model is accomplished by analyzing whole-sky imagery to produce a cloud mask which drives input parameters to the radiative transfer (RT) code MODTRAN. The RT code is run for several different azimuth and zenith orientations to create a three-dimensional representation of the hemisphere. The model is then used to achieve a per-pixel correction by adjusting the ELM slope spatially. This method is applied to real data acquired over the atmospheric radiation measurement (ARM) site in Lamont, OK. Performance of the method is evaluated with the Hyperspectral Digital Imagery Collection Experiment (HYDICE) instrument as well as a simulated multi-spectral system.

**Keywords:** Atmospheric inversion, empirical line method, ground truth, MODTRAN, cloud cover, whole-sky image

## 1. INTRODUCTION

Remote sensing offers a unique opportunity to use science to gain an understanding of the world that has only recently been possible. With the advent of multi- and hyper-spectral sensing systems, algorithms have been developed to utilize the information provided by many spectral bands. Due to the nature of electro-optical systems, data is usually collected in the form of a digital number (DN). Many of these algorithms rely on having the acquired imagery converted from DN into specific, material dependent units. This is important for several applications since DN is dependent on many sensor and atmospheric factors. While sensor specific factors tend to be relatively stable, atmospheric factors can change from day to day, or even from minute to minute. Many applications need to compare these airborne measurements to measurements made in the lab. The path length over which light travels in the lab has virtually no atmosphere between the object of interest and the sensor. The usefulness of the airborne image can therefore be increased if it is converted into units that do not depend on the atmosphere. One such unit that is widely used to alleviate these influences is called the surface reflectance factor ( $r_{s\lambda}$ ).<sup>1</sup>

Calibration of airborne imagery into reflectance factor can be accomplished by many different techniques. This research will focus on the empirical line method (ELM) which has been shown to achieve an accuracy of around one reflectance unit under optimal conditions.<sup>2</sup> These conditions occur rarely in a majority of cases and it is therefore beneficial to look for ways in which additional measurements can be made to offset error created by, for example, cloud cover. In order for atmospheric correction to benefit from ground measurement the conditions in which reflectance retrieval exhibits large errors must be explored. In most atmospheric removal algorithms the

---

Further author information: (Send correspondence to J.R.S.)

B.D.B.: E-mail: bdb6919@cis.rit.edu, Telephone: 1 315 212 0245

J.R.S.: E-mail: schott@cis.rit.edu, Telephone: 1 585 475 5170

assumption is made that the atmosphere is spatially homogenous. Since this assumption is violated when cloud cover exists in the scene, the benefits of using ground measurement in the presence of clouds will be explored. The ability of ELM to compensate for spatially changing illumination conditions is poor so a novel approach to inject information from ground measurements into the ELM process is shown. Since the new algorithm will essentially adjust the empirically derived relationship spatially throughout the scene it is termed Adaptive ELM (AELM).

The AELM algorithm is validated by using data collected over the atmospheric radiation measurement (ARM) site in Lamont, OK. The Hyperspectral Digital Imagery Collection Experiment (HYDICE) sensor was used to collect several different flight lines. The time difference between flight lines is used to generate a shift in cloud configuration. This is essentially the same as moving spatially in the scene, but with the added benefit of being able to group the validation tarps within a small spatial location. This data was taken in conjunction with the necessary types of ground truth needed to perform an AELM. The algorithmic performance is also evaluated by degrading the spectral resolution to match the multi-spectral sensor: Wildfire Airborne Sensor Program LITE (WASP Lite), which is operated by the Laboratory for Imaging Algorithms and Systems at the Rochester Institute of Technology. In both cases, the ground truth was utilized successfully in the AELM algorithm to decrease the error present in the retrieved reflectance.

## 2. APPROACH AND THEORY

### 2.1 Empirical Line Method

This section will explore how an ELM works as well as when its assumptions fail. ELM uses a simplistic view of how sensor reaching radiance arrives at the aircraft by assuming that the reflectance of the target has no bi-directional effects. It also assumes that there are essentially only three sources of radiation, the direct solar radiation ( $E'_s$ ), the radiance scattered by the atmosphere onto the target called downwelling radiance ( $L_d$ ), and the radiance scattered by the atmosphere upward toward the sensor called upwelling radiance ( $L_u$ ). Equation 1 shows this relationship for two image pixels; note that the wavelength dependence has been dropped for notational convenience.

$$\begin{aligned} L_1 &= \left( \frac{E'_s \cos \theta \tau_1}{\pi} + L_d \right) \tau_2 r_1 + L_u \\ L_2 &= \left( \frac{E'_s \cos \theta \tau_1}{\pi} + L_d \right) \tau_2 r_2 + L_u \end{aligned} \quad (1)$$

The above relationship can be expressed in simple linear terms, and the slope and intercept may be expressed as a function of the radiance and reflectance for two known pixels where  $m = \left( \frac{E'_s \cos \theta \tau_1}{\pi} + L_d \right) \tau_2$  and  $b = L_u$ . Values for  $m$  and  $b$  can then be found by simple linear regression and a reflectance value for each pixel in the image can be found using equation 2.

$$r_\lambda = \frac{L_{pixel} - b}{m} \quad (2)$$

### 2.2 Adaptive ELM

As mentioned previously, one major issue facing the ELM is that of variability created by clouds. It has been suggested that the question in remote sensing is not 'how clear is the sky', but 'how thick are the clouds that I'm looking through?'<sup>3</sup> With this in mind cloud cover effects on the ELM need to be considered. A sky with partial cloudy conditions will at the minimum change the value for the slope spatially across the scene. Conceptually this can be envisioned by viewing a hillside that has several cloud shadows present. Clearly the value for  $E'_s$  and therefore  $m$  will be much different in a shadow region than in the sun. It should be noted that even if two different areas have identical direct illumination from the sun, both  $\tau_1$  and  $L_d$  may also change spatially due to

cloud effects. To combat these changes in the slope term, it is assumed that changes in  $\tau_2$  and  $L_u$  are negligible and can be ignored.

The crux of this research lies in utilizing ground measurements to improve the situation. The measurement used to correct spatial effects on the ground is made with a spectrometer. If it is fitted with a cosine receptor and is up-looking, it will measure all the energy that the ground target is subjected to. This measurement will be referred to as a *SKY* measurement and can be expressed mathematically as:  $E'_s \cos\theta\tau_1 + E_d = SKY$ . We can express the governing equation used in ELM for the sensor reaching radiance in terms of the SKY quantity, shown in equation 3.

$$\begin{aligned} L_{sensor} &= \frac{SKY\tau_2}{\pi}r + L_u \\ &= mr + b \end{aligned} \quad (3)$$

This formulation assumes the measurement of Lambertian surfaces since we are simply dividing by  $\pi$  to convert from irradiance to radiance. If it is assumed that the transmission from the ground to the aircraft does not change,  $\tau_2$  can be solved for. The location where  $\tau_2$  is determined will be called the calibration site and it will be used to adjust the slope at remote locations. Now for each location away from the calibration site that has a SKY measurement, the slope can be adjusted as shown in equation 4. In this way the ELM derived at a calibration site can be adapted to work spatially throughout the scene, hence the name ALEM.

$$\begin{aligned} \tau_2 &= \frac{m_{calSite}\pi}{SKY_{calSite}} \\ m_{remote} &= \frac{SKY_{remote}\tau_2}{\pi} = \frac{SKY_{remote}}{SKY_{calSite}}m_{calSite} \end{aligned} \quad (4)$$

This theory can also be extended to use imagery that is not in absolute radiometric units. Equation 3 can be re-written into equation 5 which shows how the governing equation for sensor reaching radiance can be written in terms of digital count (DC).

$$\begin{aligned} DC &= \frac{SKY\tau_2c}{\pi}r + b \\ &= mr + b \end{aligned} \quad (5)$$

where  $c$  is a calibration constant and  $b$  is a sensor bias. It is also important to note that as long as each SKY measurement is cross-calibrated to respond identically to the same irradiance, the requirement for absolute radiometric calibration is relaxed. The above process can then be repeated as shown in 6 where  $k = c\tau_2$ .

$$\begin{aligned} k &= \frac{m_{calSite}\pi}{SKY_{calSite}} \\ m_{remote} &= \frac{SKY_{remote}k}{\pi} = \frac{SKY_{remote}}{SKY_{calSite}}m_{calSite} \end{aligned} \quad (6)$$

### 2.3 Radiometric Modeling

While this process will work very well for each region which has a SKY measurement available, it is not logistically possible to measure each pixel in a scene. For this reason, the SKY term must be modeled so that the ELM relationship can be corrected for each pixel.

In order to model the SKY term accurately at least one more source of ground truth is needed, bringing the total requirement up to two measurements. This new source of data is a whole-sky image which can be

acquired through the use of a fisheye lens mounted to a camera. It can be used to generate a cloud mask which contains locations for all clouds present in the sky during an over-flight. Once the cloud locations are known, the radiative transfer code MODTRAN is run to produce a three-dimensional model of the hemisphere. This is accomplished by breaking the hemisphere into many discrete parts called quads. MODTRAN calculates the radiance produced along a path running through each quad. The whole-sky image is used to classify each quad as one of three different types of sky. The first type is blue sky which is found by comparing the blue and red channels of the image. The second type is a dark cloud which is transmitting incident sunlight. Both of these first two types of quad can be generated readily with MODTRAN. The third type of quad is a bright cloud in which sunlight is being scattered and reflected from the cloud itself. Since MODTRAN operates by assuming the atmosphere is broken into many infinite parallel planes, it will always generate by default a solid overcast sky when the cloud model is activated. In order to get the correct geometry for the bright cloud runs, a single run can be done with MODTRAN looking down at the cloud from just above the cloud deck. This process is shown in Figure 1.

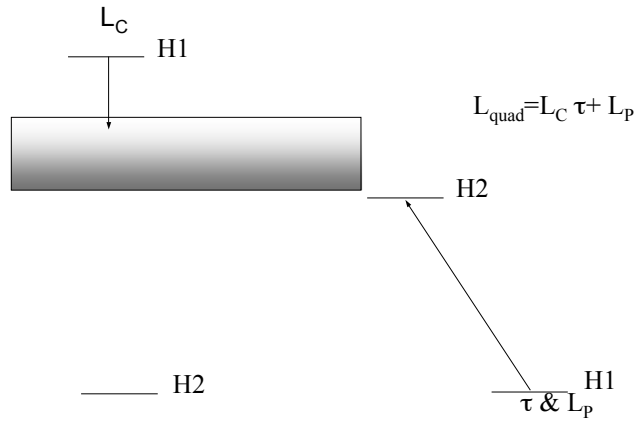


Figure 1. Two runs used to produce a bright cloud quad. H1 and H2 are the positions of the observer and target in MODTRAN.

This radiance can then be modified using the transmission and path radiance present along a path from the ground to the bottom of the cloud deck, as shown in equation 7.

$$L_{bright} = L_{cloud}\tau_{path} + L_{path} \quad (7)$$

Once the settings for each quad have been determined, MODTRAN is run producing a database of quad radiance. The SKY term is found by numerical integration of each quad using equation 8 where n and m are the total number of quads in the zenith and azimuth.

$$SKY = \frac{\left[ E_{direct} + \sum_{i=0}^N \sum_{j=0}^M L_{d\lambda}(i, j) \cos\sigma_{[i,j]} \sin\sigma_{[i,j]} \Delta\sigma \Delta\phi \right]}{\pi} \quad (8)$$

For locations on the ground away from where the whole-sky image was taken, the cloud positions can be predicted if the cloud height is known. The cloud height can be found in many cases by data taken at airport locations. The cloud base height measurement is accurate over a fairly large spatial range since cloud formation occurs at the lifting condensation level which changes on a relatively slow spatial scale.<sup>4</sup>

## 2.4 Validation Data Set

The validation data comes from a site located in Lamount, OK which is an ideal location for studying atmospheric effects. The site has a permanent infrastructure installed for collection of ground truth and is situated such that there are minimal background effects. Figure 2 shows the ground location of the tarps and the sky camera.

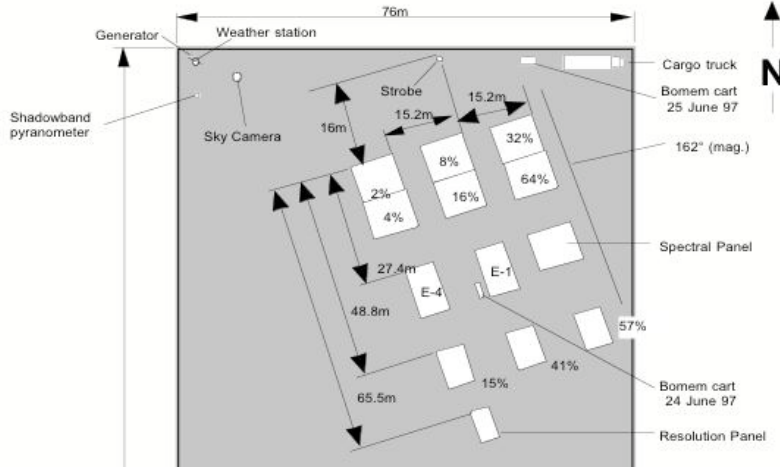


Figure 2. Layout of ground truth used in validation.<sup>5</sup>

The data collection was performed from June 24<sup>th</sup> through the 26<sup>th</sup> in 1997. The sensing system used to collect data was the Hyperspectral Digital Imagery Collection Experiment (HYDICE) which is a hyperspectral sensing system which has a spectral range of approximately 400 - 2500nm.<sup>6</sup> Ground truth was also collected of the area shown in Figure 2 by MTL Systems, Inc. (MTL)

## 2.5 Multi-Spectral System

Many systems used today do not have the hyper-spectral characteristics of a system like HYDICE. The spectral response of the multi-spectral system WASP LT was used to degrade the spectral resolution of the data from the ARM site. Figure 3(a) shows a diagram of the layout of the front of WASP LT which shows that it contains seven cameras. The response functions for the lower five cameras are used. Each camera has a different spectral filter, the set of which is comprised of a Dichroic red, green, blue, a 695nm longpass, and a 715nm longpass. The normalized spectral response functions are shown in Figure 3(b).

## 3. VALIDATION

The data which this validation uses has limited spatial extent in its ground truth. In other words, the validation panels are close spatially to each other. However, the AELM can be also be applied to the same scene in which just the time has changed. Since the cloud field is in motion, each scene will therefore have a unique cloud configuration. As long as the sun movement is accounted for, the observed variability will be that which is generated by the cloud movement.

Extensive ground truth was collected which took many different forms.<sup>5</sup> This research uses three measurements from the collection. The first is standard panel reflectance which is measured for the northmost grouping of six panels. The 2%,4%,32%, and 64% panels are used to generate the ELM while the 8% and 16% panels are used for validation. The second measurement used is the SKY term which was obtained by measurement of the reflective standard Spectralon© using a Geophysical Environmental Research Mark IV spectroradiometer (GER). The final measurement used was collected using the sky camera. This camera produced whole sky images showing cloud location / orientation. These were captured on film and then scanned to produce digital images.

The airborne asset used in this collect was the system HYDICE. It flew over the site several times at various altitudes and orientations. Two runs in particular were chosen for use as seen in table 1. The runs were chosen

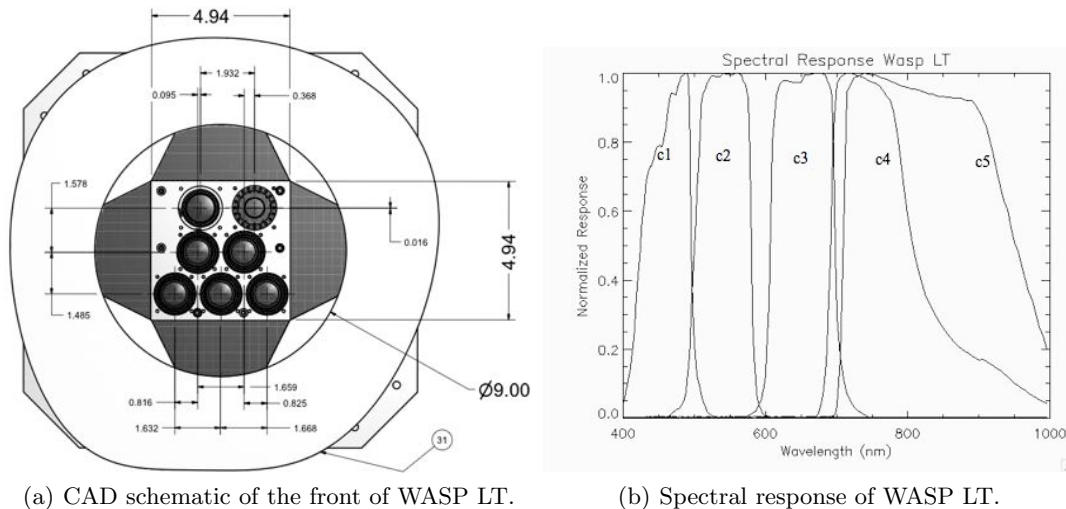
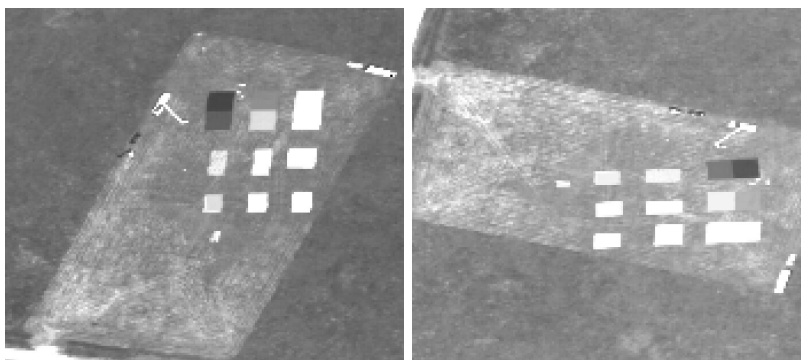


Figure 3. Wasp LT system used to degrade spectral resolution of HYDICE data.

for several reasons. First, the amount of cloud cover in the sky was sufficient to produce significant variability in the radiance field. Second, the flying height of the aircraft was low enough to give several fully resolved pixels of the tarps at a good signal to noise ratio. Third, both the sky camera and the SKY measurements were made at a time which coincides very closely to the time in which the HYDICE imagery was collected. Finally, MODTRAN model parameters were developed which work well with atmospheric conditions present during the collect.<sup>7</sup> Each run shows a slightly different view of the calibration site which can be seen in Figure 4.

Table 1. Flight and Ground Truth Information

Flight Information			Time		
Run	Date	Altitude (ft)	Hyperspectral Imagery	Whole Sky Image	SKY Radiance
07	6/24/97	6075	12:26	12:30	12:27
10	6/24/97	6030	12:47	12:50	12:47



(a) run 07 (b) run 10

Figure 4. Calibration site from each flight-line. Atmospheric effects have not yet been accounted for.

A radiometric model was also produced of the hemisphere as described in section 2. The cloud locations were determined by digitally processing the whole-sky image to produce a cloud map. The cloud map was broken into 1224 quads of equal angular segments. MODTRAN was run for each quad to produce the radiance for each direction. The modeled SKY term was then calculated from the model data. This process was performed for both run 07, which is shown in Figure 5, and for run 10.

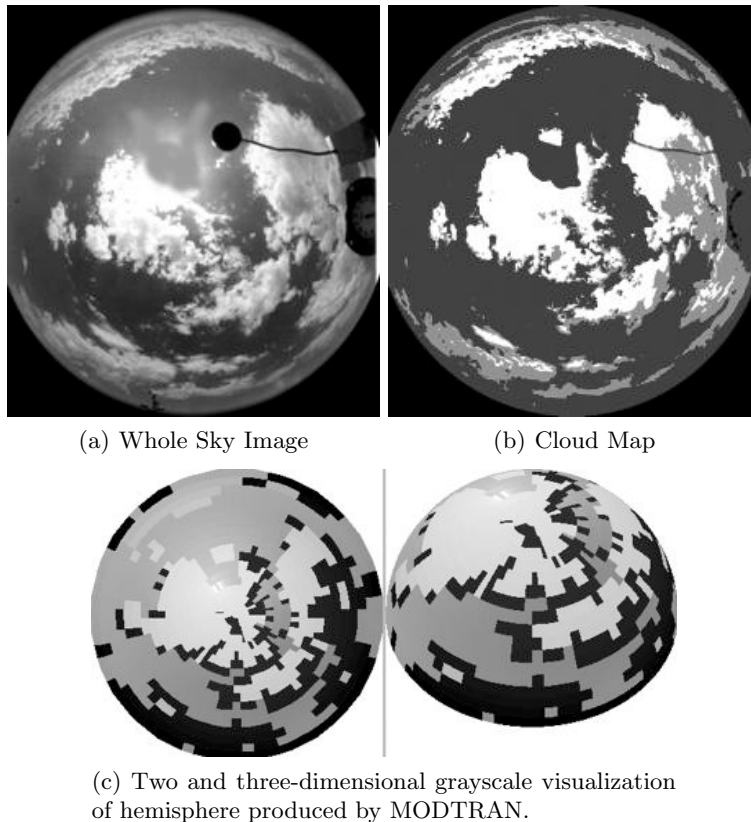


Figure 5. Process to produce hemispherical cloud model for run 07. The whole sky image in 5(a) is used to produce a map containing blue sky, dark cloud, and bright cloud in 5(b). The map is broken into 1224 quads which represent the location of each MODTRAN run seen in 5(c).

An important consideration is that of the sun location. Since each run is taken at a different time, the sun will move. The change in solar energy, assuming that  $\tau_1$  does not change significantly, will be the ratio of the cosine of the solar zenith angle. Table 2 shows the solar angles and percent change expected for each of the runs.

Table 2. Change in Solar Energy

Run	Time	Solar Declination ( $\sigma$ )	$\cos\sigma$	Percent Change
07	12:26	21.11°	0.93	
10	12:47	19.10°	0.95	1.23

An ELM was performed using the data from run 07. The slope was then adjusted by forming the SKY ratio seen in equation 4 so that it will produce accurate results when applied to run 10 data. The ratio was formed using both the measured and modeled SKY values. Each source of data was also spectrally degraded by the WASP LT spectral response. Figure 6 shows the ratio formed for both the hyper- and multi-spectral systems. It can be seen that in the case of a clear sky model that the ratio is boosting the slope by an appropriate amount to account for solar illumination changes.

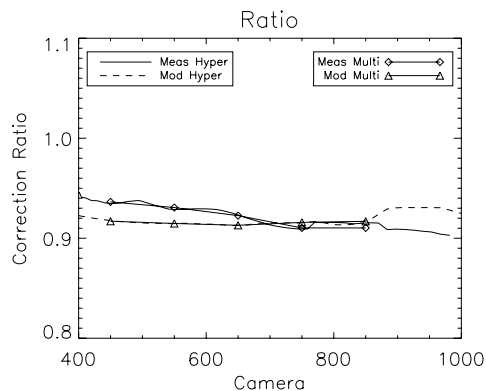
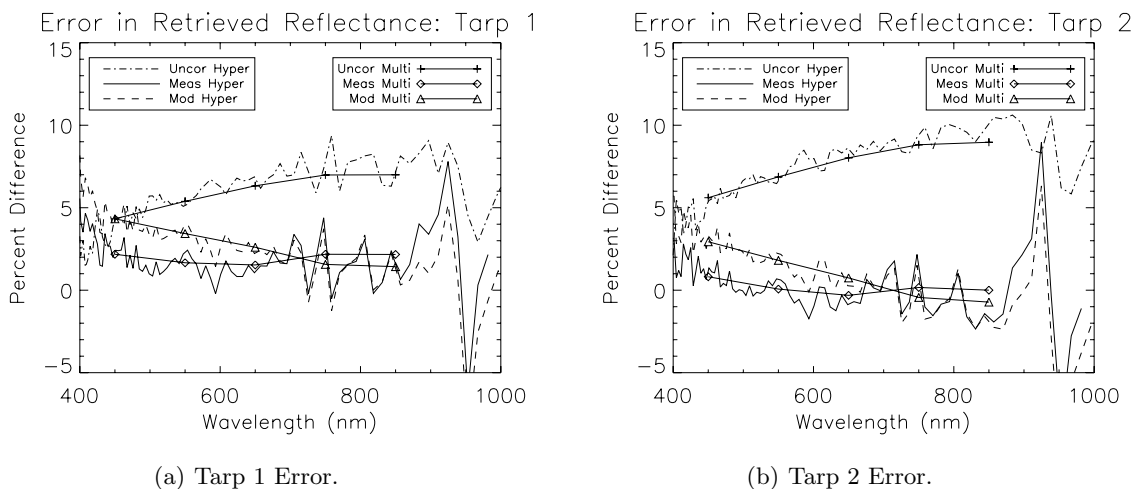


Figure 6. Slope correction ratio which adjusts for changes in illumination. The five bands of WASP LT are shown overlaid with the hyper-spectral data.

In order to assess the performance of the corrections, the 8% and 16% tarps in Figure 2 were used. The reflectance from each tarp was retrieved in run 10 and compared to the retrieved reflectance obtained using the run 07 ELM. The percent difference for both tarp 1 (8% tarp) and tarp 2 (16%) is shown in Figure 7.



(a) Tarp 1 Error.

(b) Tarp 2 Error.

Figure 7. The percent difference between the retrieved reflectance obtained with AELM for each validation tarp for run 10 and the retrieved reflectance obtained with ELM for run 07.

Performing no correction to the slope results in approximately six to ten percent error over most wavelengths, or approximately two reflectance units in the tarp 2 case. This error level decreases drastically in both cases when the SKY ratio is used. The measured SKY ratio and the modeled SKY ratio perform comparably at longer wavelengths, with the measured values performing better at shorter wavelengths. The model starts to show higher error at low wavelengths due to an increase in Rayleigh scattering. Since the exact atmospheric constituents are not known, the model will not be able to completely characterize the scattering process. The effects seen in the shorter wavelengths also show that the correction, which is modeled as multiplicative, also has an additive term which should be considered in future work.

The multi-spectral system also has the interesting effect of improving the correction for longer wavelengths. This is because the spectral resampling essentially averages the error which oscillates about the x-axis. In this case, this resampling produces a net difference which is very close to zero.

## 4. CONCLUSIONS

A novel approach for improvement of retrieved reflectance in the presence of clouds was presented. Two methods for adjusting an ELM derived slope were developed. The first used a simple ground based measurement of the total sky illumination. The second method used a model of this measurement derived with the additional input of a whole-sky image.

The approach was validated through the use of real data collected over an ARM site. This data was collected in hyper-spectral resolution in the air by the system HYDICE and on the ground by a GER spectro-radiometer. The spectral resolution was also degraded by the application of the spectral response function of WASP LT. An ELM was performed using data from run 07, and the slope was adjusted to match the illumination conditions of run 10. The retrieved reflectance from each validation tarp for run 10 was compared to that of run 07. There was significant improvement (on the order of five to seven percent of the measured value is seen in most wavelengths) with the measured correction. Further work on the SKY modeling process is needed to achieve the results seen by measurement in spectral regions below 650 nm.

In an operational sense, model inputs could be derived from data sources such as satellite products eliminating the need for the additional ground truth used in this validation study. For example, if fisheye imagery is not available, an estimate of cloud location and height can be derived from a data source such as the Geostationary Operational Environmental Satellite (GOES). Combined with a digital elevation map of the scene, the angular location of each cloud as seen from each pixel can then be approximated. This data can then be used to model the SKY term and implement the AELM over much larger spatial areas. This process will be limited spatially as the  $\tau_2$  and  $L_u$  terms begin to change. While a study of the spatial effects on this process is beyond the scope of this paper, it should be considered for real world applications.

## REFERENCES

1. K. Anderson, E. J. Milton, and E. M. Rollin, "Sources of uncertainty in vicarious calibration," in *Geoscience and Remote Sensing Symposium*, **IEEE**, 2003.
2. R. Price, C. D. Anger, and S. Mah, "Preliminary evaluation of casi preprocessing techniques.," in *Proceedings of the 17th Canadian Symposium on Remote Sensing*, pp. 694–697, Canadian Remote Sensing Society and Canadian Aeronautics and Space Institute, 1995.
3. J. Shanks and B. Shetler, "Confronting clouds: detection, remediation and simulation approaches for hyper-spectral remote sensing systems," in *29th Applied Imagery Pattern Recognition Workshop*, pp. 25–31, IEEE, 2000.
4. J. M. Wallace and P. V. Hobbs, *Atmospheric Science: An Introductory Survey*, Academic Press, 1977.
5. MTL, "Atmospheric compensation investigation: Arm site," tech. rep., Lockheed-Martin Missiles and Space Company, August 1997.
6. R. W. Basedow, D. C. Carmer, and M. E. Anderson, "Hydice system: implementation and performance," in *Imaging Spectrometry*, **2480**, pp. 258–267, 1995.
7. L. C. Sanders, *An Atmospheric Correction Algorithm for Hyperspectral Imagery*. PhD thesis, Rochester Institute of Technology, 1999.