Accuracy Assessment of Hyperspectral Imagery: Atmospheric Calibration and Image Classification Considerations

Phoebe L. Hauff
Spectral International, Inc
P.O. Box 1027
Arvada, CO 80001, USA

Douglas C. Peters
Peters Geosciences
169 Quaker Street
Golden, CO 80401-5543, USA

William A. Peppin and Eric C. Prosh
Spectral International, Inc
P.O. Box 1027
Arvada, CO 80001, USA

Eric D. Dillenbeck
Dept. of Geology and Geol. Engineering
Colorado School of Mines
Golden, CO 80401, USA

Gary A. Borstad
G.A. Borstad Associates Ltd
9865 W. Saanich Road, Suite 114
Sidney, BC V8L 3S1, Canada

Abstract-**Accuracy Assessment is one of the most important considerations in the evaluation of remotely sensed imagery. Too often, it is not done when imagery is produced. The accuracy of an image is effected by many variables, including the spatial and spectral resolution of the hyperspectral sensor, processing statistics used, types of classifications chosen, limits of detection of different surface materials, suitability of reference spectra used for image analysis training, the type and amount of ground truth data acquisition, and type of atmospheric correction algorithm applied to the imagery.**

This presentation will discuss selected examples generated from work performed under the NASA EOCAP (Earth Observations Commercial Applications Program) project NAS 13-99004. The first example is from the Ray copper Mine in Arizona, USA. It demonstrates the affects of spectral library references vs insitu ground truth, and different processing techniques on the identification and distribution of a target mineral, jarosite, in an image. The second example shows how the choice of processing cutoffs can change the distribution of a target mineral, alunite, in the image. The third example evaluates old and new atmospheric correction algorithms.

I. INTRODUCTION

Accuracy assessment can be defined as a quantitative evaluation of the success and reliability of a classification on a hyperspectral image. An accuracy assessment entails performing direct comparisons between the final classification and information derived from independent ground-truthing. This is affected by many variables, including the spatial and spectral resolution of the hyperspectral sensor, processing statistics used, types of classifications chosen, limits of detection of different surface materials, suitability of reference spectra used for image analysis training, which leads to the method of ground truth acquisition, and type of atmospheric correction algorithm applied to the imagery.

Appropriate accuracy assessment is not routinely reported for hyperspectral image analysis. Rarely are processing statistics and image-derived spectra provided for published images, which would allow evaluation of the degree of accuracy of processing techniques and results. Classifications for different minerals, vegetation, or other materials using the same image usually will have different matching statistics. How this is reported can be misleading.

Differences in the spectral and spatial resolutions of sensors also can produce radically different results in image products because higher reflecting materials may dominate larger (15-30 m) pixel size data, whereas smaller pixels (1-5 m) integrate less area and generally examine more uniform materials on the ground. With higher spectral resolution, finer differences in material also can be discriminated and one component of the ground in a pixel may no longer obscure less dominant materials.

Another important issue is adequate ground truth. For there to be a high degree of confidence in an image classification result, site-specific spectral databases should be used to provide more reliable matches to image data than are possible with generalized spectral libraries or no training reference spectra at all.

Perhaps the most important processing consideration is how do atmospheric correction algorithms affect the accuracy of the image. Most processors rely on some type of program that removes effects of the atmosphere. Usually that is coupled with a field calibration of the image data, with ground spectra collected at or near the time of data acquisition to emulate the concurrent atmospheric conditions. Conventionally this is considered essential to produce the best representation of surface information from the image data. This paper will query the validity of that belief.

II CASE STUDY EXAMPLES

Examples utilizing hyperspectral data from study areas associated with the NASA EOCAP (Earth Observations Commercial Applications Program) project and the U.S. Environmental Protection Agency (EPA) Utah Abandoned Mine Land (AML) Watershed Hyperspectral Analysis Project are used to illustrate accuracy assessment issues and ground truth data acquisition. These examples are particularly useful as they compare three different hyperspectral sensors with several different spectral resolutions (4-15 nm) and ground pixel sizes (4-20 m). These include the full range (400-2500 nm) AVIRIS sensor from the NASA-Jet Propulsion Laboratory; the SWIR-range (1200-2500 nm) SFSI sensor (SWIR Full Spectrum Imager) from Borstad Associates and the CCRS (Canada Centre for Remote Sensing), and the VIS-NIR range (400-1100 nm) CASI sensor (Compact Airborne Spectrographic Imager).

The first example involves establishing the limits of detection for jarosite at the Ray, Arizona, copper mine and shows the results of using different spectral regions and broad-band features in obtaining accurate results. These are image examples. They show that spectral data collected on the ground, at the site under investigation, are essential for accurate image generation.

The second example presents alunite distributions near the Dragon halloysite mine in central Utah. This case addresses matching statistics to show how broad classification windows can imply the presence of a specified mineral when the mineral actually is absent on the ground. Figure 1 compares spectra extracted from a Low Altitude AVIRIS image of the Dragon Mine, using different classifications or SAM (Spectral Angle Mapper) angles from the commercial ENVI™ image processing program. The spectra flagged with an arrow are the reference spectra for alunite. With the best correlations-lowest SAM angle, Figure 1-[3], the reference matches the image extracted spectra almost exactly. As the SAM angle increases, the matches to the reference decrease so that with a SAM angle of <0.09, Figure 1-[1], none of the spectra are an exact match and most of them do not resemble the reference at all. If this cutoff is used, the image is inaccurate.

Therefore, if the matching statistics are not published with the image, it is not possible to evaluate how valid is the distribution of a phase within an image.

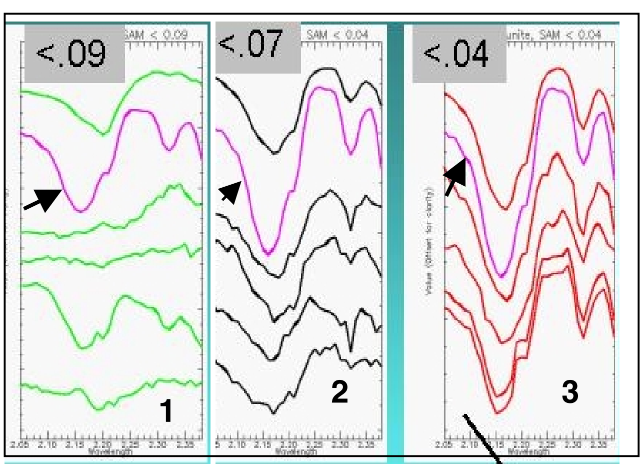

Figure 1 – Spectra extracted from Low Altitude AVIRIS alunite image over the Dragon Mine, UT. The best matches to the reference spectrum (marked with an arrow) are seen in [3], with the lowest SAM angle. The poorest matches are shown in [1].

A third example compares Low Altitude AVIRIS data sets from both the Dragon and Ray Mines and assesses several algorithms (those in common use and a new one) for atmospheric correction. The specific mineral under investigation in these examples is jarosite, a hydroxylated iron sulfate. The algorithms evaluated include: 1). Empirical Line (EL), which is atmospheric correction applied using ground site reflectance calibration data; 2) standard ATREM (ATmospheric REMoval) which applies atmospheric correction factors, without ground calibration; 3) ATREM/EFFORT (Empirical Flat Field Optimal Reflectance) (AE), which combines the EFFORT custom smoothing algorithm with ATREM; and 4) a Modified ATREM (MA) program, written by William Peppin. Spectra extracted from the images created using these four options and compared against ground reference and spectral library spectra are shown in Figure 2.

It should be noted that the "ground truthing" was done in two ways. Not only was reflectance calibration spectra

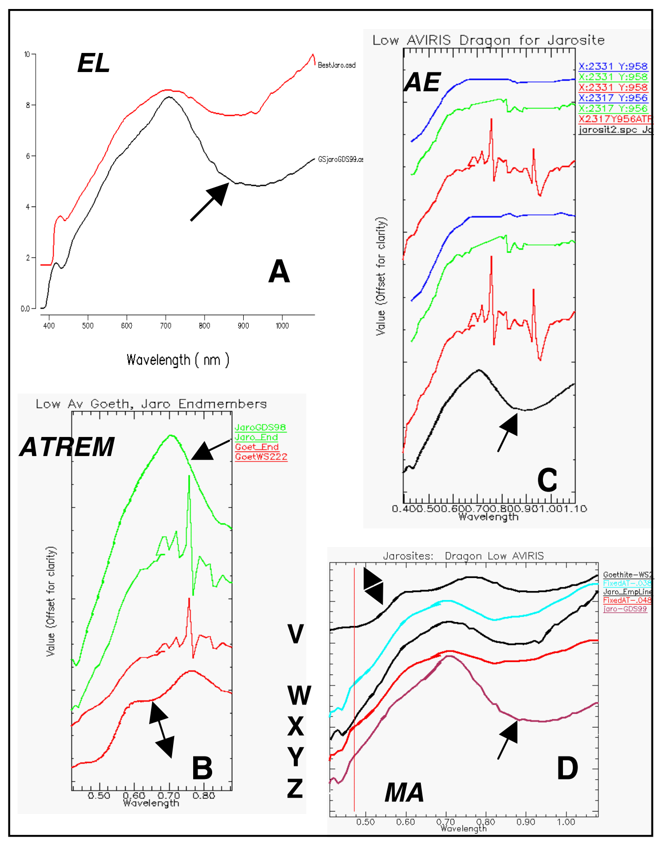

Figure 2– Spectra extracted from the Low Altitude AVIRIS iron mineral image over the Dragon Mine, UT. Reference spectra are flagged with an arrow for jarosite and double arrow for goethite. [A]Empirical Line spectra for jarosite, [B] standard *ATREM corrected spectra for jarosite and goethite, [C] ATREM/EFFORT corrected spectra for jarosite, and [D] Modified ATREM corrected spectra for goethite and jarosite.*

collected for atmospheric corrections, but numerous mineral samples were also analyzed from ground sites at both the Dragon and Ray mines to verify the image contents. As can be seen by visual comparison between the four different combinations of spectra in Figure 2, the Empirical Line [A] method gives the best results and this is to be expected. Although the target mineral is jarosite, there is usually a mixture with goethite, an iron oxide. Note that both ATREM and ATREM/EFFORT do not remove all the interference and noise contributed from the atmosphere. Although this allows these algorithms to produce data sets that can distinguish the target minerals, there is much higher SNR (Signal to Noise) which results in less accurate imagery. The distribution and concentrations of target materials are not as well defined the images producing the spectra shown in Figure 2 - [B] and [C].

The spectra shown in Figure 2-[D] are reasonably close to those extracted from the Empirical Line image [A]. These spectra were created using a new algorithm that modifies the ATREM algorithm. In this figure, spectra [V] goethite and [Z] jarosite, are spectral library references. Compare spectra [W] which is a jarosite dominated mix with goethite, derived using the modified ATREM algorithm, against spectrum [X], which is a similar composition, but derived using the Empirical Line method. The comparison is quite good and shows most atmospheric affects removed using the Modified ATREM program. A similar result is shown in spectrum [Y], which is a jarosite dominated spectrum generated using the Modified ATREM algorithm.

This new algorithm shows great potential for reducing some of the atmospheric calibration routines required for hyperspectral image processing. It is in the process of undergoing additional intensive tests. A potential issue is slightly reduced SNR. This is not a problem with high quality data sets from bright targets; however, with noisy data and dark targets, the signal extracted from the data will fall off, especially in higher wavelengths. However, the advantage of not having to do field reflectance calibration, especially in inaccessible terrains, is major.

III. SUMMARY

1) Spectral data collected on the ground, at the site under investigation, with a field spectrometer, are essential for accurate image generation and verification of image end members.

2) Great care must be taken in the choice of processing SAM angle cutoffs in order to generate accurate image classifications. The cutoffs used for different image end members must be reported for the image to appropriately represent the ground surface.

3) Atmospheric correction is not always necessary. A New algorithm (Modified ATREM) can provide comparable image results to the standard Empirical Line techniques.

IV. FOR FURTHER INFORMATION

Additional information on the Utah study areas can be found by clicking on the NASA logo at web site *http://www.SpecMin.com*. Further information on the NASA/JPL AVIRIS sensor can be found at *http://makalu.jpl.nasa.gov/aviris.html*. Information for the SFSI and CASI sensors is on the G.A. Borstad Associates website: *http://www.borstad.com*. Information on the NASA EOCAP program, including commercial partners, can be obtained from the NASA web site at *http://www.crsp.ssc.nasa.gov/hyperspectral/hypermain.htm*. We also have prepared a CD of selected PowerPoint™ presentations and papers on our EOCAP work. This CD is available upon request from Phoebe L. Hauff by writing to her at the address noted above or through her e-mail address (*pusa@rmi.net*). CDs of reports and data prepared for our EOCAP project also are available upon request, but have an associated nominal cost for reproduction, handling, and shipping.

ACKNOWLEDGMENTS

The following collaborators and cooperators also have been involved in our NASA EOCAP project to various extents. F.B. Henderson III, HENDCO Services, L.G. Closs and K. Lee, Colorado School of Mines, J.L. Thiros, formerly with the Utah Dept. of Environmental Quality, D.W. Coulter, Newmont Exploration Ltd., D.A. Robbins, ASARCO Inc., Paul Spoor, Tintic Mines and K.W. Wangerud, USEPA Region 8. We also acknowledge EOCAP project NAS 13-99004 and the Project Scientists and Administrators at NASA Stennis Space Center without whom this work would not have been possible.

SELECTED REFERENCES

P.L. Hauff, D.C. Peters, W.A. Peppin, E.C. Prosh, E. Dillenbeck, G.A. Borstad R. Diaz, and D.A. Robbins, "The Ray copper mine, Arizona--Accuracy Assessment of Hyperspectral Imagery," in *Proceedings of the 10th Annual JPL AVIRIS Workshop*, February 28-March 3, 2001, Pasadena, California. Jet Propulsion Laboratory, in press.

William A. Peppin, Phoebe L. Hauff, Douglas C. Peters, Eric C. Prosh, and Gary Borstad**,** "Different Correction-to-Reflectance Methods and Their Impacts on Mineral Classification Using SFSI and Low- and High-Altitude AVIRIS," in *Proceedings of the 14th International Conference and Workshops on Applied Geologic Remote Sensing*, November 6-8, 2000, Las Vegas, Nevada. Ann Arbor, Environmental Research Institute of Michigan, 2000, pp. 78-79.