SOIL MINERALOGY ANOMALY DETECTION IN DIXIE VALLEY, NEVADA USING HYPERSPECTRAL DATA

Gregory D. Nash and Glenn W. Johnson

Energy & Geoscience Institute
University of Utah
423 Wakara Way, Suite 300
Salt Lake city, UT 84108
gnash@egi.utah.edu & gjohnson@egi.utah.edu

ABSTRACT

Anomalous soil mineralogy can indicate buried geologic structures and zones of permeability. Hyperspectral data can be used to map these anomalies in geothermal exploration efforts. This paper describes work done at Dixie Valley, Nevada, in this regard. Dixie Valley, which lies in the Great Basin, west-central Nevada, is host to a structurally controlled deep-circulation geothermal system. Caithness Energy, LLC operates a 65 Megawatt geothermal power plant in the northwest corner of the valley. This study took advantage of archive AVIRIS (Advanced Visible and Infrared Imaging Spectrometer) airborne hyperspectral imagery of the area adjacent to the power plant. Two spectral unmixing methods were employed to separate minerals from other components in a given pixel (ground unit in computer terms). The first was a supervised learning method (which requires the use of a training data set of known mineral spectra). The second method, polytopic vector analysis (PVA), is an unsupervised classification method, which allows spectral fingerprints to be derived based on analysis of ambient data. Both methods produced useful spectral end-members. The supervised methodology facilitated mapping of a strong calcium carbonate anomaly that exists in an area where fumaroles appeared shortly after the image was acquired.

INTRODUCTION

This paper addresses the initial results of work regarding the development of data processing methodologies, which can be transferred to industry, to facilitate the use of hyperspectral data for mapping soil mineralogy anomalies that may be related to hydrothermal convection systems. This effort has been focused on Dixie Valley, Nevada.

The detection of subtle soil anomalies can be useful in locating blind geothermal systems, buried structures, and zones of permeability. Soil mineralogy anomalies can result from many conditions, including: (1) parent material consisting of hydrothermally altered regolith, (2) recent fumarole activity related to relatively short-term reductions in reservoir pressure, caused by steam production or seismic events, which lead to boiling, or (3) minerals from inactive buried fumaroles. Dixie Valley was an excellent study area as reduced reservoir pressures created prerequisite conditions for soil mineralogy anomaly formation in the mid 1990s.

Hyperspectral data has been useful in a number of past mineralogical studies and was, therefore, believed to be a good resource for the detection and mapping of soil mineralogy anomalies. Proper processing of this type of data is critical for accurate interpretation and several considerations must be kept in mind. The spatial resolution of hyperspectral data is such that most, if not all, pixels represent mixtures of surface materials. Thus, spectral unmixing is a necessary part of the processing flow. Both unsupervised and supervised unmixing methods were tested to determine which would be the most effective in terms of cost, ease of use, and results. Additionally, before attempting hyperspectral data processing and analysis, the data must undergo atmospheric correction and conversion to apparent reflectance.

DATA PREPROCESSING

The dataset used in this study was NASA Jet Propulsion Lab's (JPL) Airborne Visible/Infrared Imaging Spectrometer (AVIRIS), which was acquired on May 20, 1995 (Figure 1). This data had 224 channels, with a sampling interval of approximately 10 nm, ranging from 383 nm to 2508 nm. The spatial resolution of the dataset (pixel footprint size) was 20 m and the image was 614 x
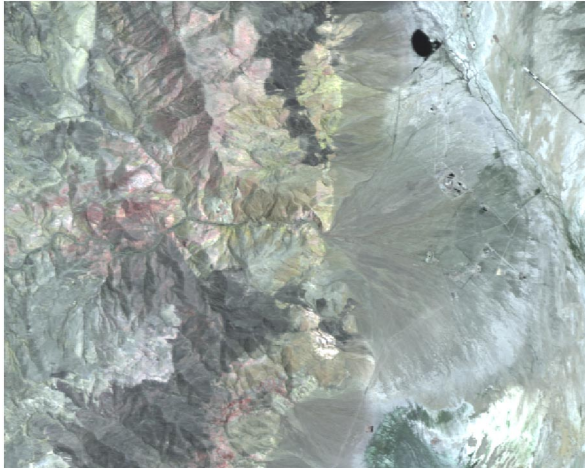
512 pixels.

Figure 1. AVIRIS RGB image of the Dixie Valley study area.

Due to the inherent nature of airborne or satellite hypersepctral data, atmospheric effects are present and must be minimized. Additionally, hyperspectral data are generally not supplied as reflectance data. Therefore, the data need to undergo conversion from radiance to apparent reflectance to better facilitate accurate interpretation. This conversion is generally an integrated part of most atmospheric correction procedures.

In an ideal situation spectral data collected on the ground, during hyperspectral data collection over-flights, can serve to aid in the correction of atmospheric effects. However, as a field effort is not always possible during data collection, techniques have been developed to allow post data-capture processing. Three of these methods were tested in this study including (1) IAR Reflectance (Kruse, 1988) - the conceptually simplest method, which uses the internal average spectrum of a data set to derive relative reflectance and to minimize atmospheric effects, (2) Atmosphere REMoval Program (ATREM) (Bo-Cai et al, 1993; Bo-Cai et al, 1997) which uses radiative transfer modeling, and (3) Atmospheric CORection Now (ACORN) which also uses radiative transfer modeling (more information about ACORN can be found at http://www.aigllc.com/acorn/intro.asp).

Figure 2 shows a single spectrum, from the AVIRIS data, processed using each of the above three techniques. All three spectra clearly show distinctive kaolinite absorptions at 2.165 μm and 2.205 μm, and would be useful for mapping said mineral.

However, visual comparison of the three processed AVIRIS spectra with a laboratory kaolinite spectrum
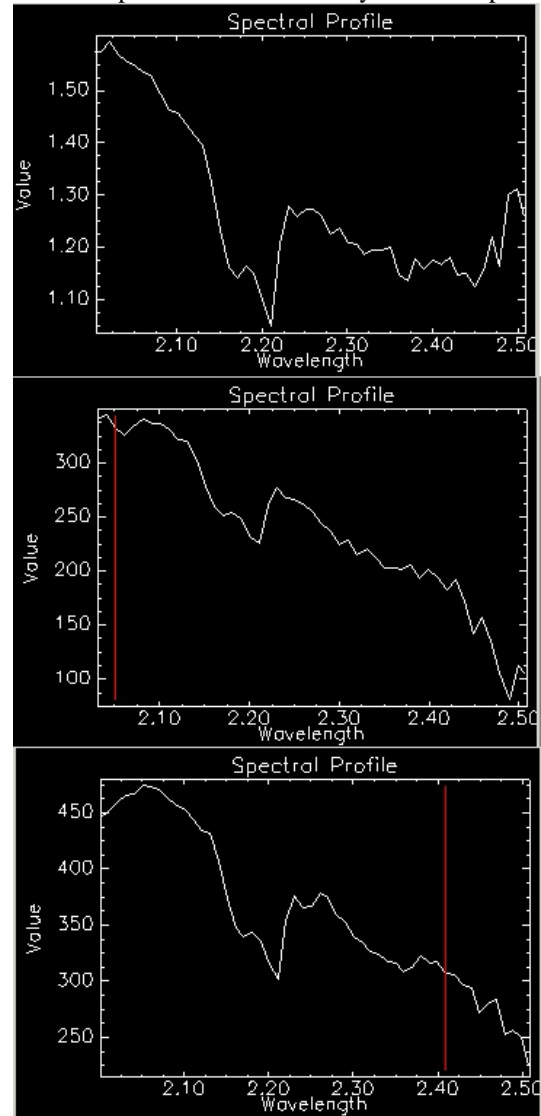

Figure 2. AVIRIS data derived kaolinite spectra, IAR, ATREM, and ACORN processed from top to bottom.

(Figure 3), it becomes evident that the ACORN processed data gives the closest match. The ATREM processed data is also close and could probably be used adequately to accomplish the goals of this project. However, ATREM code is no longer supported or distributed, therefore, its future use is problematic unless the code is in-hand. For general purposes, such as mapping only major mineral, occurrences, the IAR method should work in most cases. This method is built into commonly available software packages, such as Imagine™ and RSI ENVI™. This method is also fast and easy. The user

needs little background in remote sensing to utilize it. However, radiative transfer model based atmospheric
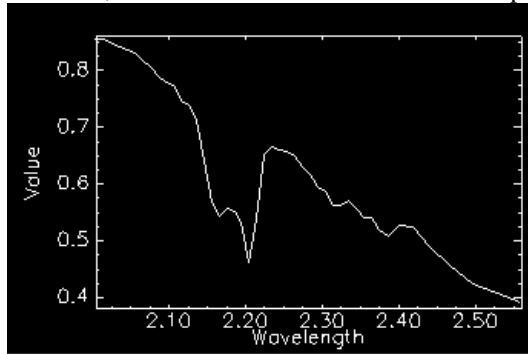

Figure 3. Laboratory kaolinite spectrum.

correction would be recommended for use in mapping detailed mineralogy.

To further illustrate this, Figure 4 shows another example of a spectrum, taken from the same pixel, processed with IAR, ATREM, and ACORN, from top to bottom respectively. Again, all three spectra show distinctive kaolinite absorption features. However, in the ATREM processed spectrum, an absorption feature at 2.350 μm (which is clearly shown in the ACORN processed spectrum and to a lesser degree in the IAR processed spectrum) is minimized. This absorption is likely to be indicative of another mineral, such as muscovite, which is mixed within the pixel with kaolinite, although not necessarily within the same rock. This being the case, the ACORN processed data would be more useful for detailed mineralogy mapping. The only drawback to ACORN is that the user needs to have a reasonable understanding of hyperspectral data processing.

DATA PROCESSING

The AVIRIS data, after preprocessing for atmospheric correction and conversion to apparent reflection, were processed to (1) extract mineral end-member spectra and (2) to map the relative abundance of these minerals. This was done in two ways. The first was an unsupervised method – polytopic vector analysis (PVA). This technique was tested because it could be easily adopted by the geothermal industry, in exploration efforts, for use by personnel with minimal training in image analysis. The second method tested, supervised unmixing and classification, on the other hand, is more complicated and difficult to use.

Unsupervised Classification – Polytopic Vector Analysis

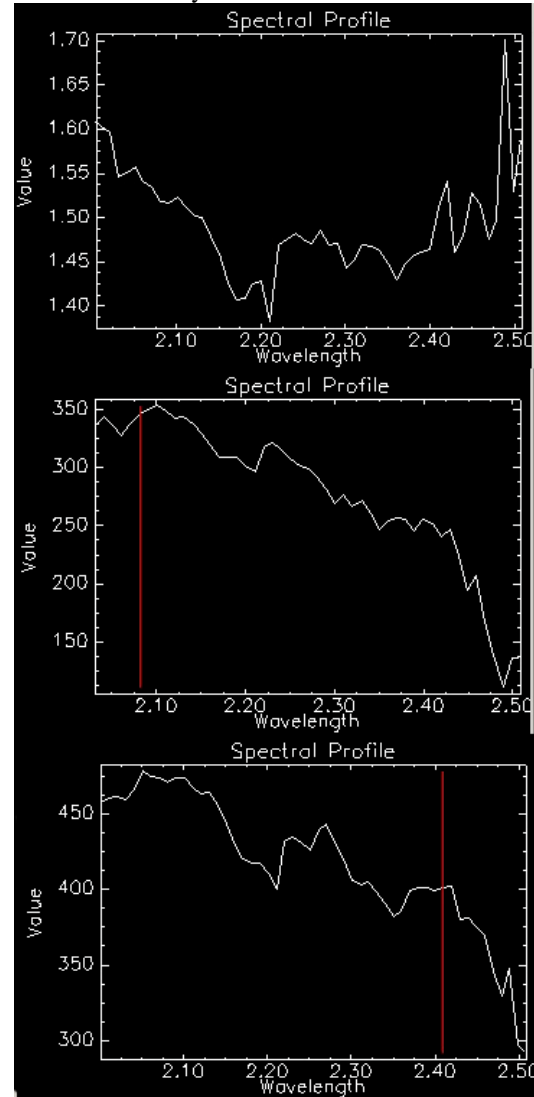PVA is a multivariate method developed in the earth sciences for analysis of mixtures. PVA is an

Figure 4. Kaolinite and unknown mineral mixed spectrum.

unsupervised classification method and also referred to as a "self-training classifier." In other words, with PVA, the analyst does not need not to know, or assume, the composition of contributing end-member spectra. This is advantageous because it allows feasible end-members to be inferred in areas where little ground-truth data exists and with little user input. Given samples from a mixed system, PVA is used to resolve three parameters of interest: (1) the number of end-members contributing to the mixture; (2) the spectral composition of each end-member (i.e. the spectral "fingerprint"); and (3) the relative proportion of each end-member in each sample.

PVA has been used extensively for analysis of multivariate mixtures in the earth and environmental sciences. Johnson et al. (2002) outline its development and formalism. It is a relatively new tool in remote sensing applications, but it has been applied to the analysis of Dixie Valley AVIRIS data (Johnson & Nash, 1998) in a vegetation study focused within a small subset of the study area.

Prior to PVA implementation, the AVIRIS image data were preprocessed by (1) atmospheric correction using ATREM; and (2) reduction to 44 bands in the 2.0309 μm to 2.4592 μm range. This range was chosen as it eliminates strong water and vegetation signatures but includes important hydroxyl absorption features found in many hydrothermal alteration minerals.

PVA was implemented using MATLAB code (The Mathworks, Inc., Natick, MA). A commercial version of the PVA software is available through Tramontane, Inc. (http://www.etramontane.com). One modification to the PVA algorithm was made for this study: After the principal components analysis step and determination of the number of end-members, the DENEG algorithm of Full et al., (1981) was applied only to those pixels located on the convex hull of the reduced dimensional data cloud. This increased the speed of the DENEG algorithm because it operates on a very small subset of the original number of pixels (161 of > 300,000 pixels) while yielding an identical solution as the analysis of the full data set.

For this exercise, library reference spectra were used to help identify end-members produced by PVA processing. This was done initially with no ground-truth information to simulate conditions where the analyst has minimal *a priori* knowledge of site geology. In such a situation, a user would likely be forced to rely on published spectra. These spectra were from the USGS Digital Spectral Library (Clark et al., 2000). Derived end-member compositions were compared to the library spectra by visual inspection/comparison, and by calculation of the cosine theta (cos θ) similarity metric (Davis, 1986). The values of cos θ range from 0 (for two completely dissimilar spectra) to 1.0 (for two identical samples).

Supervised Classification

Unlike PVA, the supervised unmixing and classification requires a number of different processing steps. These included the following:

1. Minimum noise fraction (MNF) transformation;

2. Pixel purity index (PPI) generation;
3. Selection of mineral spectra end-members from the PPI;
4. Mixture tuned matched filtering (MTMF); and
5. Color enhancement (optional).

ATREM and ACORN processed data were classified in the project with the results of the ACORN processed data being reported here.

RSI ENVI™ image analysis software (http://www.rsinc.com/envi/index.cfm) was used for data processing and interpretation. It is not known if other image analysis packages allow the specific processing flow described here, however, other packages do have specific hyperspectral data processing and analysis capabilities.

The first step, the MNF transformation is, in practice, a pair of cascaded principal components (PCA) transforms, with the first using an estimated noise covariance matrix with the second being a standard PCA transform (Green et al., 1988; Boardman and Kruse, 1994). The most significant difference in the result from a standard PCA is that the data output bands are ordered by noise. When using a standard principal components transform this is generally not the case. The noise bands resulting from the MNF transform are easily recognized both visually and by observing the resultant eigenvalues. Noise bands can be removed from further processing. A way to further reduce the dataset is to use only the bands necessary to achieve the desired results, as input to the MNF transform. As hydroxyl bearing minerals and calcium carbonate were the primary targets of interest in this study, the input data set was reduced to only those bands in the 2.0 μm to 2.5 μm range.

The second step used the MNF transformed data, minus the noise bands, as input to create an ENVI™ PPI image. The PPI image indicates the most spectrally pure pixels in the MNF transformed image. The PPI image and the atmospherically corrected apparent reflectance data set can be linked on the computer screen so that, as each potential end-member pixel on the PPI image is checked, the spectrum can also be viewed. Figure 5 shows Dixie Valley apparent reflectance image on top of a PPI image along with a chlorite end-member that was found using the PPI image. It must be noted, however, that not every pixel identified in a PPI image is going to produce a useful end-member. Each potential end-member must be validated in step three.

The third step is to identify the end-members

produced in the PPI image. To aid in the identification of spectral end-members, mineral
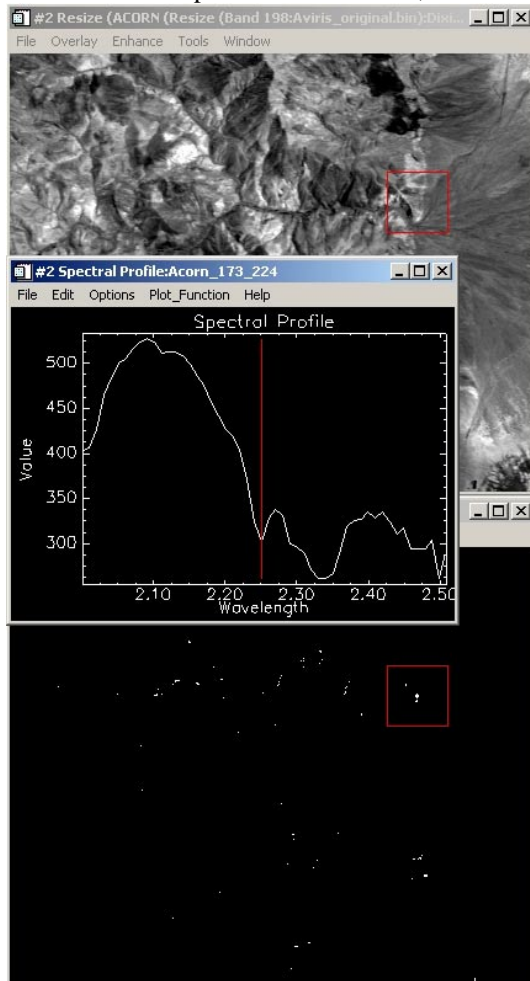

Figure 5. Dixie Valley apparent reflectance image on top of a PPI image and chlorite spectral end-member.

spectra libraries can be used for comparisons. Additionally, statistical correlations can be done to help identify library spectra that is similar to that identified as an end-member by the PPI image. The software used in this study facilitated this through several algorithms found in its "Spectral Analyst" feature.

The fourth step is to use the ENVI™ MTMF routine to create end-member distribution maps. To accomplish this, the end-members that have been identified are extracted from the MNF image and used as input along with the MNF image. This results in a dataset that contains double the number of bands as input end-members. These consist of a matched filter score band and an infeasibility band for each

end-member. Using these bands can help eliminate false positives. Figure 6 shows a matched filter score
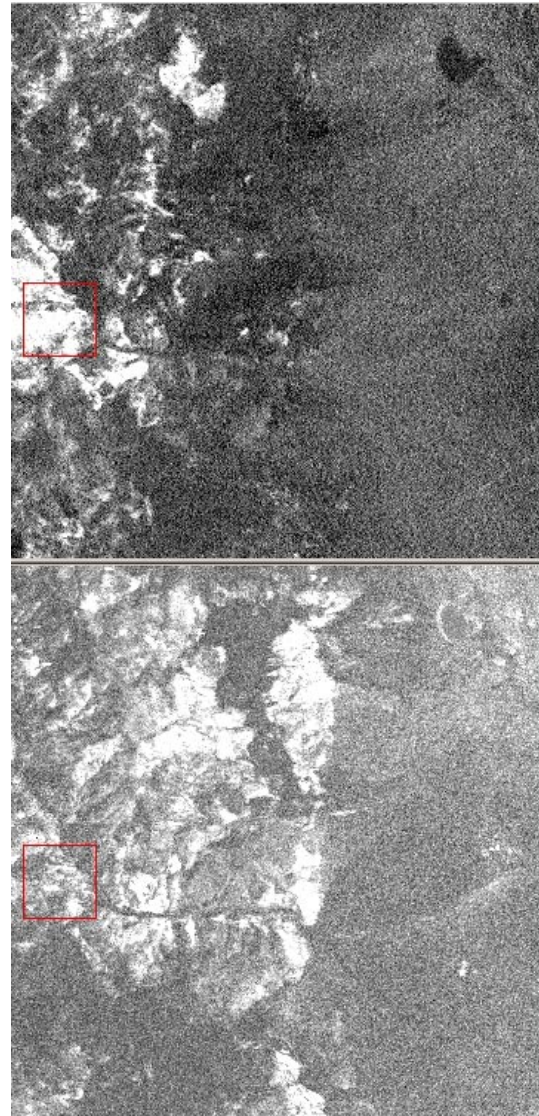

Figure 6. MTMF results. The matched filter score results are on top and the infeasibility scores on the bottom. Bright pixels indicate high scores.

band (top) and an infeasibility band (bottom) for kaolinite. Figure 7 shows a spectrum from a pixel with a matched filter score of 0.233 out of a possible 1.0 (highest score in this example) and an infeasibility score of 7.88, where the highest score was 214.1. This suggests that there is a high probability of the pixel containing a component of kaolinite. The displayed spectrum indicates that kaolinite is indeed present in area of the selected pixel, although not pure.

Step 5 is simply using a gradational color scheme to
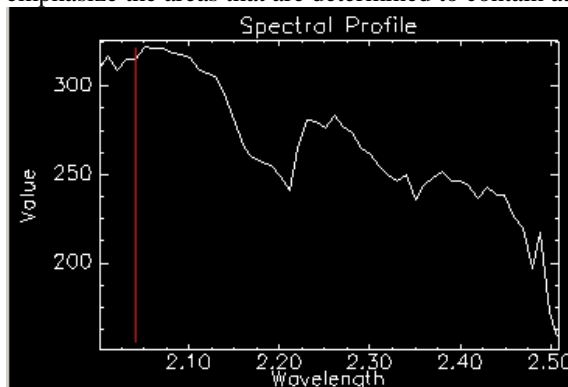
emphasize the areas that are determined to contain at

Figure 7. Kaolinite spectrum derived using the matched filter score image from Figure 6.

least some proportion of the spectral end-member in question and to indicate the relative the purity.

RESULTS

PVA Unsupervised Method

Initial analyses of data processed using PVA showed that ATREM resulted in abhorrent data values in many surface water pixels, particularly in pond pixels. This was due to residual effects of the atmospheric correction as water absorption bands are used in the process. Thus surface water pixels were also removed from the image. This resulted in removal of 551 of the original 314368 pixels. The image analyzed by PVA was 313817 pixels and 44 bands.

The number of end-members in the model were determined using method outlined by Johnson et al., 2002, in particular, the CD scatter plot method (Miesch, 1976; Johnson et al., 2000). Based on these goodness of fit diagnostics, a five end-member model was chosen. The model was resolved using the DENEG algorithm of Full et al., (1981), modified as described in the methods section, above. Resolved end-member spectra were consistent with published mineral spectra (Clark et al., 2000). The PVA model resolved end-member patterns fairly consistent with those chosen in the supervised method (kaolinite, chlorite, and muscovite). However, in addition spectra were also resolved that were consistent with water absorption and several mafic minerals (olivine, hypersthene). The full model will be presented in a subsequent paper. For the purposes of this paper, we will focus on just one of the five end-members (End-Member 2), as it illustrates an important caveat for application of unsupervised classification methods.

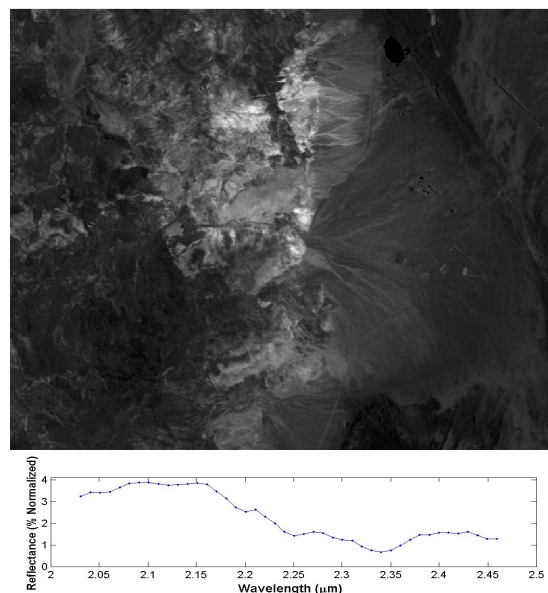The spectrum and geographic distribution of EM-2 is illustrated on Figure 8.

Figure 8. PVA End-Member 2 gray-scale abundance map, and spectra. Black pixels on map indicate absence of this spectrum. White pixels indicate maximum observed in the image (59%).

The EM-2 spectral pattern shows high reflectance in the shorter wavelengths, and an absorption feature at 2.33 - 2.34 μm. Unfortunately, there are several minerals that exhibit this characteristic (Figure 9).
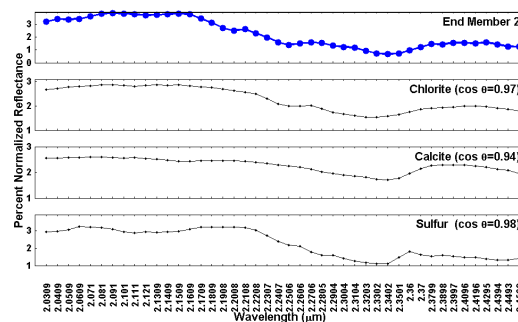

Figure 9. Comparison of End-Member 2 to mineral spectra in the USGS library. Spectra PVA EM-2 shows high similarity to three mineral spectra known to be present in the Dixie Valley study area.

Three of those (chlorite, calcite and sulfur) are shown in comparison to EM-2. All three minerals are known to be present in the study area, and when compared to EM-2, all three exhibit a cos θ value greater than 0.94. Retention of more than five end-

members did not result in differentiation of chlorite, calcite or sulfur. Therefore, it is likely that this "end-member" itself is a mixture that represents the contribution of more than one mineral spectrum. The problem illustrated here is known as "colinearity." That is, multiple minerals exhibit very similar spectral patterns within in the chosen wavelength range. Colinearity is a well-known problem in application of mixing models. Given collinear contributing end-members, it is difficult for any multivariate method to discriminate between those patterns.

Colinearity is a concern for both supervised and unsupervised methods. However, in supervised methods, assuming that the analyst has correctly chosen the contributing spectra, the colinearity problem is not as severe, because the analyst is forcing the method to differentiate between the those similar patterns (in this case: chlorite and calcite). Furthermore, if minerals with similar spectra are known to be important constituents of a study, the colinearity problem may be minimized by *a priori* selection of spectral bands that maximize the contrast between those spectra (elimination of variables that are nearly collinear). However, both examples assume the analyst has considerable a priori knowledge.

In PVA, the colinearity problem may be dealt with additional data preprocessing, or by selection of spectral bands after review of results of the initial unsupervised model. This is the approach that will be taken in subsequent PVA analyses of the Dixie Valley image. Now that we know that calcite and chlorite (and to a lesser extent, sulfur) are important spectral components of the image, subsequent PVA models may be improved by picking a wavelength range that extends down into shorter wavelengths. Figure 10 shows the full 224 band spectra for chlorite, calcite and sulfur. There is significant contrast between these spectra at lower wavelengths. However, subsequent analyses will likely not extend into the visible spectrum, because of the increased
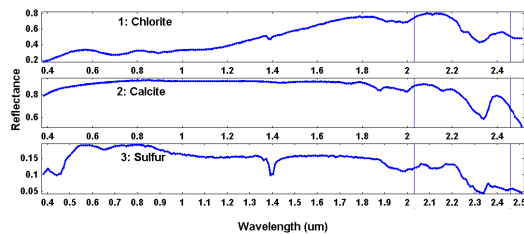

Figure 10. Full 224-band spectra for minerals similar to EM-2. Vertical lines at 2.03 and 2.46 um show spectral range of present PVA model.

potential for interference from vegetation spectra. Subsequent PVA models will, however, extend down into the 1.75 μm range*,* where there is a notable contrast between these minerals (Figure 10).

This study highlights the advantage of unsupervised methods in situations where the analyst has little or no *a priori* information, but it also points out a caveat; that we must recognize and be able to take into account colinearity of contributing spectra.

Supervised Method
Following the steps outlined in the Data Processing section above, four excellent end-member spectra were derived and used in the MTMF process. This allowed the creation of four matched filter and infeasibility score images for kaolinite, muscovite, chlorite, and calcite.

As for soil mineralogy anomalies, there were no significant anomalies for kaolinite, muscovite, or chlorite. However, a significant calcium carbonate anomaly (Figure 11) exists adjacent to an area where a hydrothermal convection related vegetation die-off occurred and along a path where fumarolic activity began a short time after the image was acquired, clearly indicating a buried structure and a zone of permeability. Fumaroles may have been active in this area sometime prior to production. The anomaly trends NNE (the image is not georeferenced).
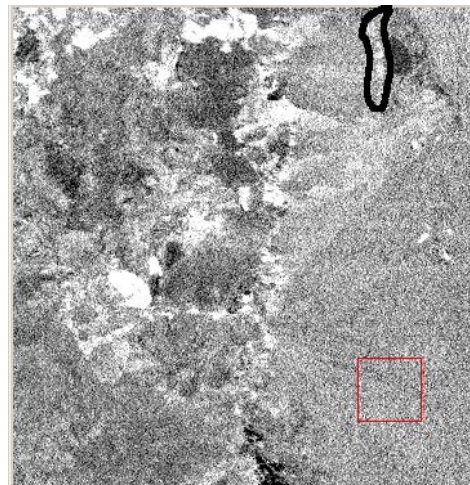

Figure 11. Soil calcium carbonate anomaly bounded by a dark black outline.

Work is now being done to refine this methodology to allow identification of additional minerals. Samples, collected from the study area, are undergoing X-ray diffraction to determine their bulk compositions. The X-ray results will act as ground

truth to determine what minerals may have been missed in the first pass of analysis.

CONCLUSIONS

Both supervised and unsupervised unmixing and classification techniques can be useful for detecting soil mineralogy anomalies in geothermal exploration efforts. PVA provides a set of interpretable end-members very quickly and can be used with no prior knowledge of the geology of a study area. In this study five readily interpretable end-members were generated, but the contribution of more than one mineral per end-member precluded discrimination between calcite and chlorite. Several options are now being considered as solutions for the colinearity problem. The obvious first step in solving such a problem is recognizing that it exists. Comprehensive spectral libraries, such as that of Clark et al. (2000), are a significant aid in deciphering problem results. In this case modification of the wavelength range used (provided it does not extend down into the visible range) may help.

Supervised unmixing and classification techniques, although more time consuming and difficult to use, may prove more useful where detailed information is needed as it is likely that additional end-members can be identified in this manner. However, this amount of detail may not always be necessary for exploration efforts. The supervised methodology led to the mapping of a soil calcium carbonate anomaly that is directly related to the hydrothermal convection system and that indicates a zone of permeability and a buried structure.

ACKNOWLEDGEMENTS

REFERENCES

Boardman, J. W., and F. A. Kruse (1994), "Automated Spectral Analysis: A Geologic Example Using Aviris Data, North Grapevine Mountains, Nevada" in *Proceedings of the Tenth Thematic Conference on Geologic Remote Sensing*, May 9-12, San Antonio, TX., I-407 to I-418.

Clark, R.N, G.A. Swayze, A. Gallagher, T.V.V. King, and W. M. Calvin (2000), "USGS Digital Spectral Library", File downloaded contains AVIRIS 1995, 224 channel wavelengths & resolution. http://speclab.cr.usgs.gov/spectral-lib.html

Full, W.E., R. Ehrlich, J.E. Klovan (1981), "EXTENDED QMODEL - Objective Definition Of External End Members In The Analysis Of Mixtures" in *Math. Geol.* **13**, 331-344.

Gao, B-C, K. B. Heidebrecht, and A.F.H. Goetz (1993), "Derivation Of Scaled Surface Reflectances From AVIRIS Data" in *Remote Sensing of the Environment"*, **44: 2**, 165-178.

Gao, B-C, K. B. Heidebrecht, and A.F.H. Goetz (1997), "Atmosphere Removal Program (ATREM) Version 3.0 Users Guide", *Center for the Study of Earth from Space, University of Colorado, Boulder*, 27.

Green, A. A., M. Berman, P. Switzer, and M. D. Craig (1988), "A Transformation For Ordering Multispectral Data In Terms Of Image Quality With Implications For Noise Removal", in *IEEE Transactions on Geoscience and Remote Sensing*, **26: 1**, 65-74.

Johnson, G.W., Ehrlich, R., and Full, W. (2002), "Principal Components Analysis And Receptor Models In Environmental Forensics", in *An Introduction to Environmental Forensics* (Murphy and Morrison, Editors), Academic Press, San Diego, 461-515.

Johnson, G.W., Jarman, W.M., Bacon, C.E., Davis, J.A., Ehrlich, R., and Risebrough, R. (2000), "Resolving Polychlorinated Biphenyl Source Fingerprints In Suspended Particulate Matter Of San Francisco Bay", *Environ. Sci. Technol.,* **34**, 552-559.

Johnson, G. W. and G. D. Nash (1998), "Unmixing of AVIRIS Hyperspectral Data from Dixie Valley, Nevada", *Proceedings: Twenty-Third Workshop on Geothermal Reservoir Engineering*, Stanford University, Stanford, California, **23**, 240-245.

Miesch, A.T. (1976), "Q-Mode Factor Analysis Of Geochemical And Petrologic Data Matrices With Constant Row Sums", *U.S. Geological Survey Professional Paper 574-G*, 47.