

GEOLOGICAL MAPPING BY COMBINING SPECTRAL UNMIXING AND CLUSTER ANALYSIS FOR HYPERSPECTRAL DATA

N. Ishidoshiro^a, Y. Yamaguchi^{a,*}, S. Noda^a, Y. Asano^a, T. Kondo^b, Y. Kawakami^b, M. Mitsuishi^b, H. Nakamura^b

^aGraduate School of Environmental Studies, Nagoya University, Chikusa-ku, Nagoya, 464-8601 Japan - yasushi@nagoya-u.jp

^bExploration Technology Division, Japan Oil, Gas and Metals National Corporation, 2-10-1 Toranomon, Minato-ku, Tokyo, 105-0001 Japan – mitsuishi-mayumi@jogmec.go.jp

Commission VIII, WG VIII/5

KEY WORDS: Geology, Mineral, Unmixing, Cluster Analysis, Hyperspectral, Cuprite

ABSTRACT:

Spectral unmixing of hyperspectral data often fails to select some minerals and rocks having flat spectra but no diagnostic absorption features as endmembers, even if they are actually important endmembers. To avoid this problem, we propose a novel approach that combined two methods: spectral unmixing and full-pixel classification. First, all pixels were divided into two categories, hydrothermally altered areas and unaltered rocks based on the absorption depth of 2.0 to 2.5 μm . For the hydrothermally altered areas, endmembers were extracted by the Improved Causal Random Pixel Purity Index (ICRPPI) method, which was improved from the existing Pixel Purity Index (PPI) and Causal Random Pixel Purity Index (CRPPI) methods. Endmember abundance in each pixel was calculated by linear spectral unmixing. In a separate operation, k-means clustering was applied to the unaltered rock areas. Finally, the results of these two methods were combined to generate a single distribution map of rocks and minerals. This approach was applied to the airborne hyperspectral HyMap data of Cuprite, Nevada, U.S.A. We confirmed that our mapping result was consistent with the existing geological map as well as our field survey result.

1. INTRODUCTION

The spectral signature from minerals/rocks on the ground is useful for geological mapping. Recently, hyperspectral sensors have drawn attention because of their capability to capture detailed spectra of surface materials and thus to identify materials and measure their abundances on the ground. However, the spectral signature from each pixel contains a mixture of different spectra due to the multiple components that form the ground surface. This complexity causes incorrect identification and/or misclassification of surface materials. For this reason, it is important to solve this mixing problem in the analysis of hyperspectral data.

Spectral unmixing has been widely studied in hyperspectral data analysis as a possible solution to the mixing problem. Unmixing consists of two steps: 1) extracting characteristic constituent spectra named “endmembers” from image data, 2) calculating their abundances per pixel (Keshava and Mustard, 2002). In general, endmembers are identified as pixels that have characteristic spectra with diagnostic absorption features. However, some minerals and rocks (e.g., basalt) have flat spectra without any characteristic absorption, and such a material would never be selected as an endmember, even if it is actually an important constituent. As a result, the unintended omission of endmembers often causes incorrect identification and surface mapping results. In this case, we can apply a full-pixel classification method such as clustering analysis instead of the unmixing approach. However, only a few studies have attempted to combine mixed-pixel classification (spectral unmixing) and full-pixel classification (Dopido et al., 2011).

In this study, we propose an approach to combine these two methods, spectral unmixing and k-means clustering analysis, for

accurate geological mapping. As mentioned above, it is often difficult to extract a mineral endmember without characteristic absorption and/or to select a true endmember set from candidate endmembers in the first step of spectral unmixing. Therefore, we developed the Improved Causal Random Pixel Purity Index (ICRPPI) by improving the existing method, the Causal Random Pixel Purity Index (CRPPI) (Wu and Chang, 2009), which was modified from the Pixel Purity Index (PPI). In this paper, we adopt the ICRPPI method for unmixing of mixed pixels.

2. METHODS

We focused on the hydrothermal alteration of minerals/rocks, which is an important target in mineral exploration. Generally speaking, spectral unmixing is useful for discrimination and mapping of minerals and rocks with characteristic absorption features, whereas full-pixel classification is more suitable for geological mapping of rock formations without characteristic absorption features. Thus, we first classified the hyperspectral pixel data of the target area into two categories: hydrothermal alteration areas composed of minerals/rocks with characteristic absorptions and unaltered rocks without characteristic absorptions. Separate approaches of spectral unmixing and k-means clustering were applied to each classified area.

2.1 Classification into Hydrothermal Alteration Areas and Unaltered Rocks

Hydrothermally altered minerals/rocks have spectral absorptions around 2.2 μm because of hydroxyl (Hunt, 1974), whereas unaltered rocks show a flat reflectance spectrum in this wavelength region. We normalized the reflectance spectra from 2.0 to 2.5 μm to enhance the absorption features in each

* Corresponding author

spectrum of a pixel and classified the pixels into hydrothermal alteration areas and unaltered rocks.

To normalize the reflectance spectra, we used the continuum removal method (Clark and Rough, 1984). This method generates a continuum-removed reflectance spectrum by dividing a reflectance value R by a continuum line value (convex hull) (Noomen et al., 2006). That is,

$$R' = \frac{R}{R_c} \quad (1)$$

where R' is the continuum-removed reflectance spectrum (normalized spectrum), R is the original reflectance spectrum, and R_c is the continuum. After normalization, R' values from 2.0 to 2.5 μm were averaged for each pixel to obtain an average reflectance value of each pixel. When R has no absorption, R_c is equal to R in all spectral bands, and R' represents 1. Conversely, when R has characteristic absorption features, R_c is not equal to R at these spectral bands, and R' is less than 1. Hence, the pixels whose averaged R' value is higher than a threshold value were regarded as unaltered rocks, and the other pixels were categorized as hydrothermally altered areas. The threshold value was defined from the R' histogram of the target scene.

2.2 K-means Clustering for Unaltered Rocks

Unaltered rocks have no diagnostic absorption features but their brightness changes. For example, basalt has a flat spectrum with low reflectance (brightness), whereas andesite and rhyolite have flat spectra with higher reflectance (brightness). For this reason, brightness is more useful than spectral shape for classifying the pixels of unaltered rocks.

As a classification method, we chose k-means clustering (MacQueen, 1967), which is one of the most popular unsupervised classification methods for remote sensing data. This algorithm randomly distributes representative points of k classes in N -dimensional space. Each spectrum in each pixel is assigned to the class whose representative point is the closest to each pixel (Dopido et al., 2011). The average spectrum of each class is considered to be the centroid of each cluster. Then, all pixels are reassigned to the class whose centroid is the closest to the pixel spectrum, and the centroids are recalculated (Schowengerbt, 1997). The k-means clustering continues this iteration until the pixel assignment becomes invariable from the previous assignment. When k-means clustering processes a dataset of flat spectra, the result depends on spectral brightness. In this case, we assumed that the pixels in the same cluster belonged to the same rock formation.

However, it is difficult to identify rock types solely from their reflectance spectra after classification because of the lack of diagnostic absorption features. Therefore, in this study, each unaltered rock class was regarded as one rock type representing one geological unit.

2.3 Spectral Unmixing for Hydrothermal Alteration Areas

The Pixel Purity Index (PPI) method, widely used for endmember extraction, is based on the geometry of convex sets (Ifarraguerri and Chang, 1999). This method considers spectral signatures as vectors in N -dimensional space and extracts the edges of convex sets as candidate endmembers (Plaza et al., 2002). First, the PPI algorithm randomly generates a large number of N -dimensional vectors called “skewers” and adds

one point to the extracted extreme vector per skewer. After several skewer operations, the spectral pixel vectors whose PPI counts are larger than a threshold value are assumed to be candidate endmembers. The Causal Random PPI (CRPPI) method is a modification of PPI (Wu and Chang, 2009). This method repeats PPI with fewer skewers than those in the original PPI, compares the extracted pixels, and retains the common candidate endmembers as renewed candidates.

We developed the Improved Causal Random Pixel Purity Index (ICRPPI) method, which is a further modification of the CRPPI, in order to extract even spectra with weak absorptions (Ishidoshiro et al., 2013). ICRPPI has two operation features: extracting a union of candidates, not an intersection as CRPPI does, and grouping endmembers into commonly existing materials, such as alunite and kaolinite.

The ICRPPI algorithm starts by generating fewer skewers, as in CRPPI. The major difference between these two methods is that ICRPPI extracts the union of candidate vectors selected as extreme vectors at least twice, instead of the intersection of candidates. The reason why this method excludes candidates selected only once is that those might correspond to pixels projected onto the extremes by chance due to randomly generated skewers.

In the next step, endmember candidates are classified into material groups based on a threshold value. In particular, this operation needs to satisfy two conditions in order to collect pixel vectors into the same group: 1) the spectral angle φ between two candidate vectors is smaller than a threshold value, and 2) the spatial distance (Euclidean distance) D between two candidate vectors is smaller than a threshold value. Under these conditions, φ and D are expressed as,

$$\varphi = \cos^{-1} \left(\frac{C_1 \cdot C_2}{|C_1| |C_2|} \right) \quad (2)$$

$$D = \sqrt{(i - i')^2 + (j - j')^2} \quad (3)$$

where C_1 and C_2 are the spectral vectors of the endmember candidates, and (i, j) and (i', j') are the coordinates of two pixels. The spectral angle threshold φ' is calculated according to the following equation and the spatial distance threshold D' is defined empirically. Here, φ' is expressed as,

$$\varphi' = \frac{\varphi_{min}}{2} \quad (4)$$

where φ_{min} is the minimum spectral angle between minerals in a scene.

In the final step, after grouping, a representative endmember from each group is extracted from the selected endmember candidates. These representative vectors are treated as endmember vectors.

After extracting endmember sets, we employ conventional linear spectral unmixing to generate the mineral distribution map, which shows the abundance of each mineral (Plaza et al., 2011). This spectral unmixing result for hydrothermal alteration areas is later combined with the k-means clustering result for unaltered rocks.

3. STUDY AREA AND DATA

The study area is located in Cuprite in western Nevada, U.S.A. Cuprite is one of the most popular sites for testing the performance of new sensors and spectral mapping methods in geologic remote sensing studies (e.g., Abrams et al., 1977; Ashley and Abrams, 1980; Swayze et al., 2014). In this area, the Tertiary volcanic rocks were hydrothermally altered in the Mid- to Late-Miocene epoch. The hydrothermally altered rocks were divided into three groups: silicified rocks, opalized rocks and argillized rocks, as shown in Figure 1 (Ashley and Abrams, 1980).

We used the hyperspectral data obtained in October 1999 by Hyperspectral Mapper (HyMap), an airborne hyperspectral sensor developed in Australia. The HyMap sensor consists of 126 spectral bands in a span of 0.45 to 2.48 μm . HyMap can obtain a continuous spectrum except in absorption ranges near 1.4 and 1.9 μm due to atmospheric water vapour (Cocks et al., 1998). The spatial resolution is 2 to 10 m.

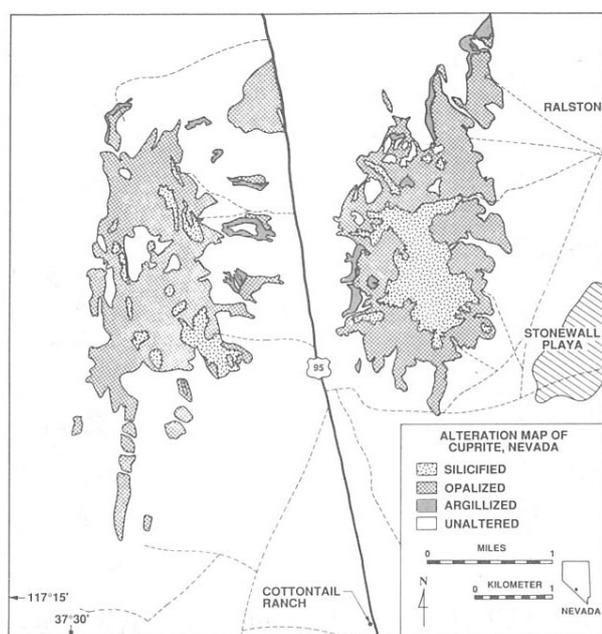


Figure 1. Alteration map of Cuprite, Nevada (Hook and Rast, 1990) redrawn from Ashley and Abrams (1980).

4. RESULTS

The classification result of the hydrothermal alteration areas and unaltered rocks is shown in Figure 2. The distribution of the hydrothermal alteration areas by our method is consistent with the hydrothermally altered zone in the existing map except for the silicified zone in Figure 1 (Hook and Rast, 1990). The silicified zone was misclassified into unaltered areas because the silicified rocks had no or weak diagnostic absorptions from 2.0 to 2.5 μm .

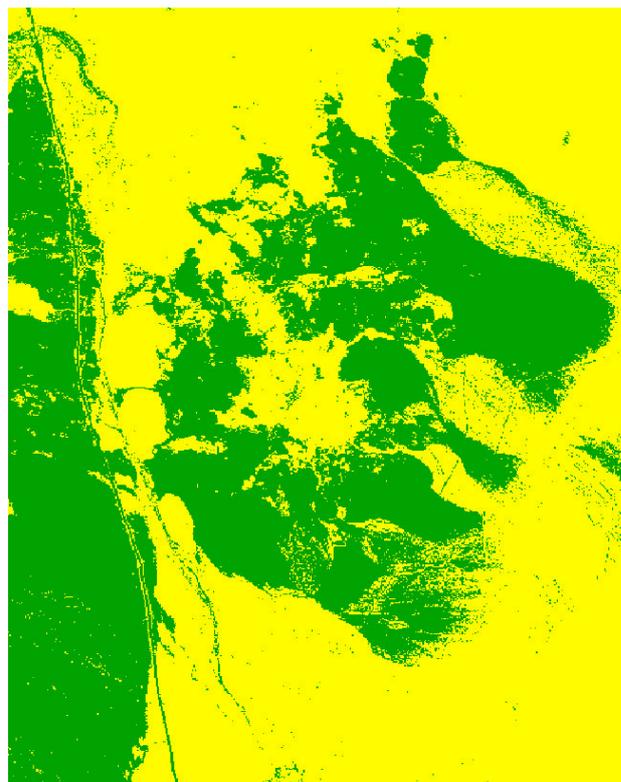


Figure 2. Distribution of hydrothermal alteration areas (green) and unaltered rocks (yellow).

The number of classes, which was necessary and important input for k-means clustering, was determined as 10, based on the number of geologic formations in the existing geological map (Workman et al., 2002). Figure 3 shows the result of k-means clustering. The distribution of major rock units was consistent with the existing geological map. We also confirmed the extent of the rock formations by a field survey in September 2014. As an example, the clustering result clearly showed the characteristic lithologic pattern formed by the dip slopes in the northeastern part of the eastern hill of Cuprite

By using the ICRPPI method, a total of 15 endmembers were extracted from the HyMap dataset, and eight endmembers were excluded due to noise. Among the 15 endmembers, 13 endmembers indicated spectra similar to pure mineral spectra. These endmembers were judged to contain alunite, kaolinite, montmorillonite, and buddingtonite. As a result of spectral unmixing by using these four representative mineral endmembers, high abundance spots of hydrothermally altered minerals were identified, as shown in Figure 4. The mineral distributions shown on these images corresponded well with the existing clay mineral maps (Swayze et al., 2014).

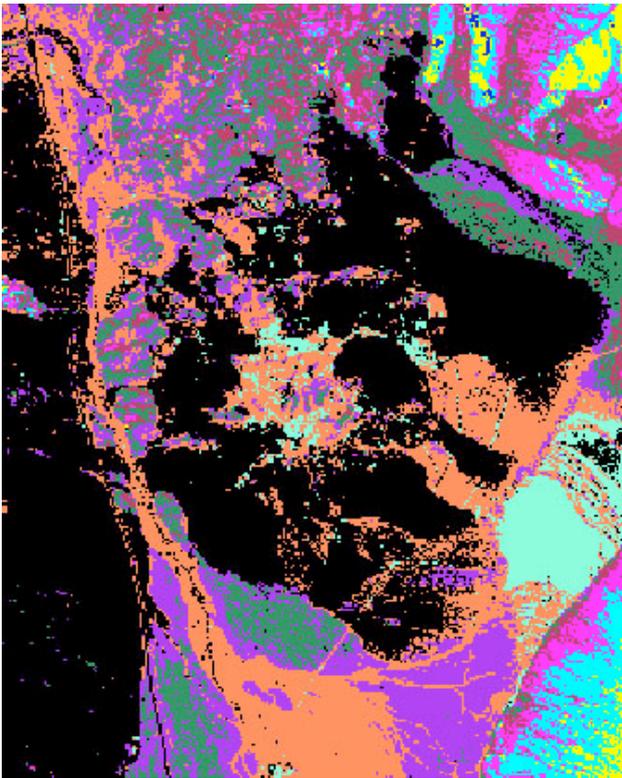


Figure 3. Result of k-means clustering for unaltered rocks.

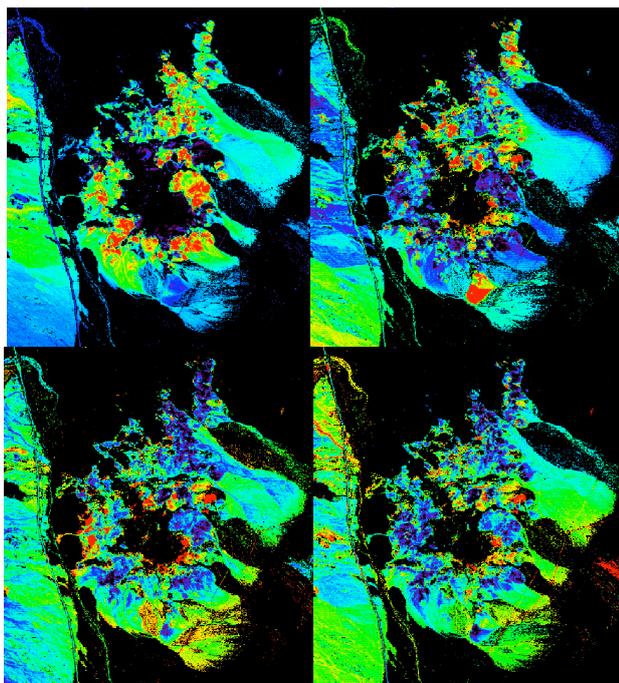


Figure 4. Alteration mineral distributions. Red and blue colours indicate high and low abundances respectively. Upper left: alunite, upper right: kaolinite, lower left: buddingtonite, lower right: montmorillonite.

We conducted field surveys in September 2013 and September 2014, and measured the spectral reflectance of the collected rock samples in our laboratory. The spectra of endmember pixels extracted by spectral unmixing were similar to those of the rock samples, and both spectra showed the existence of the same hydrothermal alteration minerals. This result indicates that spectral unmixing with the ICRPPI method worked properly. However, we must acknowledge that the abundance of each mineral was still not sufficiently accurate, and so we need further modification of the methodologies.

Finally, the two results of k-means clustering and spectral unmixing were combined to generate one distribution map of rocks and minerals. In this study, because of the importance of mineral exploration, we paid more attention to the hydrothermal alteration areas, which were analyzed by spectral unmixing. Therefore, we generated an RGB colour composite image of the hydrothermally altered minerals (R: alunite, G: kaolinite, B: montmorillonite), and combined the composite image with a monochrome image of the unaltered rock distribution (Figure 5).

The distribution map of rocks and minerals generated by the proposed method allowed us to obtain detailed information of both the hydrothermal alteration minerals and the unaltered rock formations. Moreover, the clustering result was consistent with the geological map in terms of the distribution of geological formations.

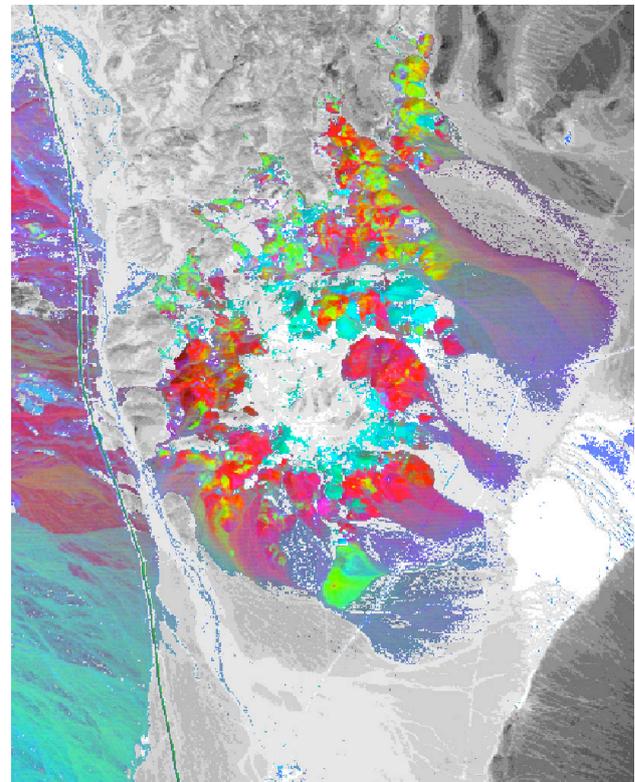


Figure 5. Combined result of distribution map of rocks and minerals. The colour composite image was generated by assigning R to alunite, G to kaolinite, and B to montmorillonite abundances. The monochrome image was the k-means clustering result for unaltered rocks.

5. CONCLUSION

Most of the previous approaches for hyperspectral data analysis have been based on spectral unmixing, which focuses on the characteristic absorptions of reflectance spectra. For this reason, it has been difficult to generate a geological map for unaltered rocks that have no characteristic absorptions. In this study, we propose a novel approach consisting of two separate methods and the following four steps: 1) data division into hydrothermally altered areas and unaltered rocks in order to apply a suitable method for each area, 2) clustering analysis for unaltered rocks, 3) spectral unmixing for hydrothermally altered areas, and 4) finally, combining the two results to generate a single distribution map of rocks and minerals.

We conclude that the combination of spectral unmixing for hydrothermal alteration areas and k-means clustering for unaltered rocks is a realistic approach to learn the accurate distributions of rocks and minerals.

When we analyze high-spatial-resolution data such as HyMap data, the proposed method allows us to obtain detailed information of both the hydrothermal alteration minerals and unaltered rock formations. However, in the data division step, a remaining problem is that the silicified zone is classified as unaltered rocks because their spectra have no or weak diagnostic absorptions from 2.0 to 2.5 μm . This problem may be solved by using the thermal infrared data. This is one of the future issues in the data division step.

ACKNOWLEDGEMENTS

This work is part of a mineral exploration renovating program conducted by Japan Oil, Gas and Metals National Corporation (JOGMEC) and is fully funded by the Ministry of Economy, Trade and Industry, Japan. This work is also partly supported by JSPS KAKENHI Grant Number 15H04225.

REFERENCES

Abrams, M. J., Ashley, R. P., Rowan, L. C., Goetz, A. F. H. and Kahle, A. B., 1977. Mapping of hydrothermal alteration in the Cuprite mining district, Nevada, using aircraft scanner images for the spectral region 0.46 to 2.36 μm . *Geology*, 5(12), pp.713-718.

Ashley, R. P. and Abrams, M. J., 1980. Alteration mapping using multispectral images – Cuprite mining district, Esmeralda County, Nevada. *U. S. Geological Survey Open-File Report*, 80-367, 17 pps.

Clark, R.N. and Roush, T.L., 1984. Reflectance spectroscopy: quantitative analysis techniques for remote sensing applications. *Journal of Geophysical Research*, 89(B7), pp. 6329-6340.

Cocks, T., Jenssen, R., Stewart, A., Wilson, I. and Shields, T., 1998. The HyMap airborne hyperspectral sensor: the system, calibration and performance. In: *Proc. 1st EARSel Workshop on Imaging Spectrometry*, pp.37-42.

Dopido, I., Villa, A. and Plaza, A., 2011. Unsupervised clustering and spectral unmixing for feature extraction prior to supervised classification of hyperspectral images. *Proc. SPIE*, 8157, pp.1-8.

Hook, S. J. and Rast, M., 1990. Mineralogic mapping using airborne visible infrared imaging spectrometer (AVIRIS) shortwave infrared (SWIR) data acquired over Cuprite, Nevada. In: *Proc. 2nd Airborne Visible Infrared Imaging Spectrometer (AVIRIS) Workshop*, JPL Publication 90-54, pp.199-207.

Hunt, G.R., 1979. Near-infrared (1.3-2.4 μm) spectra of alteration minerals –potential for use in remote sensing. *Journal of Geophysical Research*, 44, pp. 1974-1986.

Ifarrauerri, A. and Chang, C.-I., 1999. Multispectral and hyperspectral image analysis with convex cones. *IEEE Transactions on Geoscience and Remote Sensing*, 37(2), pp.756-770.

Ishidoshiro, N., Uezati, T. and Yamaguchi, Y., 2013. A new endmember extraction method in spectral unmixing for hyperspectral data. In: *Proc. International Symposium on Remote Sensing 2013*, 4 pps.

Keshava, N. and Mustard, J.F., 2002. Spectral unmixing. *IEEE Signal Processing Magazine*, 19(1), pp. 44-57.

MacQueen, L., 1967. Some methods for classification and analysis of multivariate observations. In: *Proc. Fifth Berkeley Symp. Mathematical Statistics and Probability*, 1(14), pp. 281-297.

Noomen, M. F., Skidmore, A.K., van der Meer, F.D., and Prins, H.H.T., 2006. Continuum removed band depth analysis for detecting the effects of natural gas, methane and ethane on maize reflectance. *Remote Sensing of Environment*, 105(3), pp. 262-270.

Plaza, A., Martinez, P., Perez, R. and Plaza, J., 2002. Spatial/spectral endmember extraction by multidimensional morphological operations. *IEEE Transactions on Geoscience and Remote Sensing*, 40(9), pp.2025-2041

Plaza, A., Martin, G., Plaza, J., Zorteza, M. and Sanchez, S., 2011. Recent developments in endmember extraction and spectral unmixing, In: *Optical Remote Sensing: Advances in Signal Processing and Exploration Techniques*, Prasad, S., Bruse, L., and Chanussot, J., Eds., pp.235-267.

Schowengerdt, R.A., 1997. *Remote Sensing: Models and Methods for Image Processing (second edition)*. Academic Press, 522 pps.

Swayze, G. A., Clark, R. N., Goetz, A. F. H., Livo, K. E., Breit, G. N., Kruse, F. A., Sutley, S. J., Snee, L. W., Lowers, H. A., Post, J. L., Stoffregen, R. E. and Ashley, R. P., 2014. Mapping advanced argillic alteration at Cuprite, Nevada using imaging spectroscopy. *Economic Geology*, 109(5), pp.1179-1221.

Workman, J. B., Menges, C. M., Page, W. R., Taylor, E. M., Ekren, E. B., Rowley, P. D., Dixon, G. L., Thompson, R. A. and Wright, L. A., 2002. Geologic map of the Death Valley groundwater model area, Nevada and California. *U. S. Geological Survey Miscellaneous Field Studies Map*, MF-2381-A.

Wu, C.-C. and Chang, C.-I., 2009. Causal Pixel Purity Index (PPI). *Proc. SPIE*, 7334, pp. 1-13.