SPECTRAL MIXTURE MODELING USING PRINCIPLE COMPONENT ANALYSIS APPLIED TO NONTRONITE-FERRIHYDRITE AND KAOLINITE-MONTMORILLONITE MIXTURES

Joseph S Makarewicz[1], Heather D Makarewicz[1], and Janice L Bishop[2,3]

[1]Olivet Nazarene University, Bourbonnais, IL; [2]SETI Institute, Mountain View, CA; [3]NASA ARC, Moffett Field, CA

Introduction

The hyperspectral visible/near-infrared (VNIR) imagers OMEGA [1] and CRISM [2] have detected phyllosilicates in several regions on Mars. Determining the abundance of these phyllosilicates in mixture spectra is more challenging and modeling has been attempted with Shkuratov [3] and radiative transfer [4] techniques. Here we present modeling of mixture spectra using Principle Component Analysis (PCA). Currently, we are examining lab spectra of known mixtures. Our goal is to develop a method for determining abundances of phyllosilicates, sulfates and other alteration minerals in martian spectra.

Several techniques have been evaluated for modeling and unmixing lab spectra of mixtures, e.g. PCA [5], Modified Gaussian Modeling (MGM) [6], MGM with automated initialization parameters [7-8], radiative transfer modeling [9], and checkerboard linear modeling [10]. Challenges commonly faced in mixture studies are non-unqiue solutions and human evaluation as part of the process. In a previous study, we developed a PCA unmixing technique with a Bayesian classifier and percent composition determination using linear regressions that produces reliable results without user intervention [11]. This technique was applied to a set of pyroxene mixture spectra and was capable of detecting abundance and grain size of the samples [11].

Methods

Samples and Spectra. Two suites of mixture spectra were selected for this study: a nontronite-ferrihydrite dataset and a kaolinite-montmorillonite dataset [12-13]. The mixture samples were prepared by weight percent [12-13] and spectra were measured at RELAB as in another mixture study [14].

Pre-processing. Before performing our analyses, the spectra were pre-processed. The nontronite-ferrihydrite spectra were resampled at 0.002 µm increments and were cropped to only include wave-lengths from 0.8 to 2.6 µm. The kaolinite-montmorillonite spectra were resampled at 0.002 µm increments and were cropped to only include wave-lengths from 1.3 to 2.6 µm. Each spectrum was normalized such that the area under the curve was set to 500. Normalization to a constant area under the curve causes spectra to be of similar scale and reduces the impact of albedo variations. Spectra of the mixture datasets after these pre-processing steps are shown in Figs. 1-2.

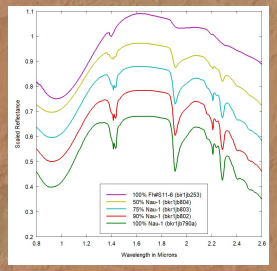

Fig. 1: Relab nontronite (Nau-1) and ferrihydrite (Fh#S11-6) mixture dataset after processing with offsets.

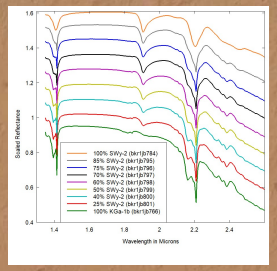

Fig. 2: Relab kaolinite (KGa-1b) and montmorillonite (SWy-2) mixture dataset after processing with offsets.

PCA procedure. PCA was performed separately on each of the pre-processed spectra datasets. The spectral dataset was arranged into a matrix with each row being a pre-processed spectrum. The mean normalized spectrum for the dataset was subtracted from each of the spectra in the matrix. Figs. 3-4 show the mean normalized spectrum for each dataset. The mean normalized spectrum is similar to the 50% mixture of each dataset in Figs. 1-2. Finally, PCA was performed using singular value decomposition (SVD) producing a large set of principle component vectors as well as corresponding principle component values for each spectrum in the dataset. The first principle component vector (PC1) for each dataset is shown in Figs. 3-4. PC1 contains the spectral components that vary the most in the mixture suite. Each spectrum in the dataset can be approximated as the mean spectrum plus the first principle component vector multiplied by the first principle component value.

PCA: Nontronite-Ferrihydrite

For the nontronite-ferrihydrite dataset, the first principle component value monotonically increased as percentage nontronite decreased. The linear relationship for the nontronite-ferrihydrite scatterplot, shown in Fig. 3, can be approximated by the following equation.

$$\%\ Nau\text{-}1 = 62 - 48 \times PC1$$

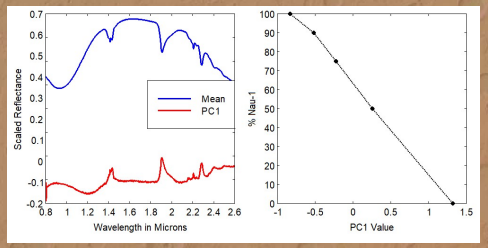

Fig. 3: On the left, nontronite-ferrihydrite mean normalized spectrum and first principle component vector generated by PCA. On the right, a scatter plot linearly correlating percent nontronite composition and first principle component value.

PCA: Kaolinite-Montmorillonite

For the kaolinite-montmorillonite dataset, the first principle component value monotonically increased as percent montmorillonite increased. The linear relationship for the kaolinite-montmorillonite scatterplot, shown in Fig. 4, can be approximated by the following equation.

$$\%\ SWy\text{-}2 = 56 + 52 \times PC1$$

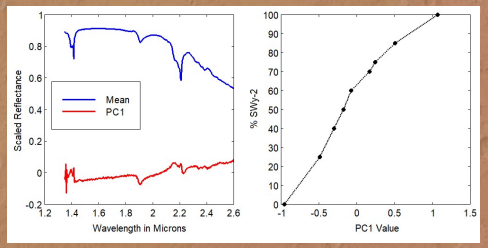

Fig. 4: On the left, kaolinite-montmorillonite mean normalized spectrum and first principle component vector generated by PCA. On the right, a scatter plot linearly correlating percent montmorillonite composition and first principle component value.

Discussion

The technique from this study can be used for other types of studies. The principle component vectors show which spectral features change in the data set. The principle component values can be correlated to the properties of the dataset. In these datasets, percent composition was varied between two endmembers, but datasets can be formed that vary other parameters or include more endmembers.

When previously applied to a set of pyroxene mixture spectra where both percent composition and grain size were varied, a Bayesian classier was used to detect grain size and a linear regression was used to estimate percent composition [11]. In this study involving smectite clay mineral datasets where only percent composition was varied, a linear regression was used to estimate percent composition. Future studies could involve large diverse endmember data sets where mineral group is varied. A Bayesian classifier would be used to predict mineral group affiliation.

This technique can be used with hyperspectral images. When working with small binary and tertiary mixture datasets, detailed percent composition mineral maps can be generated. Trace impurities can be detected by removing the spectral signatures of the prevalent minerals. When working with large diverse endmember datasets, coarse mineral maps can be formed labeling the most prevalent mineral group in different regions of the image.

A benefit of this technique is the insight into spectral features that are correlated with the variance in the dataset. For example, there is a 1.9 micron spectral feature which increases with montmorillonite composition, as seen in Fig. 2. This feature can also be seen in the first principle component vector in Fig. 4.

Since the spectral features that vary amongst the dataset are captured in the principle component vectors, this technique could be used to generate other percent composition spectra not included in the training set. For example, a 10% Nau-1 and 90% Fh#S11-6 spectrum could be generated by adding the Nontronite-Ferrihydrite mean normalized spectrum, in Fig. 3, to the first principle component vector, in Fig. 3, multiplied by a first principle component value of 1.083. The principle component value is computed using the linear regression.

Summary

A technique for modeling mixtures between two end-members using PCA and linear regressions was applied to a nontronite-ferrihydrite mixture dataset and a kaolinite-montmorillonite mixture dataset. For both studies, a linear relationship was found between the first principle component value and the percent composition. This study confirms that this new technique produces linear results for mixtures of pyroxene [11] and phyllosilicates.

Future work will involve applying this technique to other small binary and tertiary mixture collections. When used with remotely sensed data where the composition is similar to the laboratory mixture collections, percent composition estimation and trace impurity detection can be achieved. The technique will also be applied to larger diverse endmember collections. When applied to remotely sensed data, mineral group detection can be achieved.

References

[1] Poulet F. et al. (2005) Nature, 438, 632-627.
[2] Mustard J. F. et al. (2008) Nature, 454, doi: 10.1038/nature07097, 305-309.
[3] Poulet F. et al. (2014) Icarus, 231, 65-76.
[4] Ehlmann B. et al. (2011) 42nd Lunar and Planetary Science Conference, Abstract #1704.
[5] Smith M. O. et al. (1985) JGR, 90, C797-C804.
[6] Sunshine J.M., et al. (1990) JGR, 95, 6955-6966.
[7] Makarewicz H.D., et al (2009) IEEE Whispers doi: 10.1109/WHISPERS.2009.5289046.
[8] Brown A.J. (2006) IEEE Trans.Geosci.Remote Sens-ing, 44, 1601-1608.
[9] Robertson K. M. et al. (2016) Icarus, 277, 171-186.
[10] Stack K. M., R. E. Milliken (2015) Icarus, 250, 332-356.
[11] Makarewicz J.S. and H.D. Makarewicz (2013) IEEE Whispers doi: 10.1109/WHISPERS.2013.8080604.
[12] Parente, M., et al. (2008) Martian Phyllosilicates: Recorders of Aqueous Processes, 7039.
[13] McKeown, N. K., et al. (2008) Martian Phyllosilicates: Recorders of Aqueous Processes, 7033.
[14] Roush T. L. et al. (2015) Icarus, 258, 454-466.

www.nasa.gov