Spectral Mixture Modeling using Principle Component Analysis

April 18, 2018

Joseph S. Makarewicz

Heather D. Makarewicz

Outline

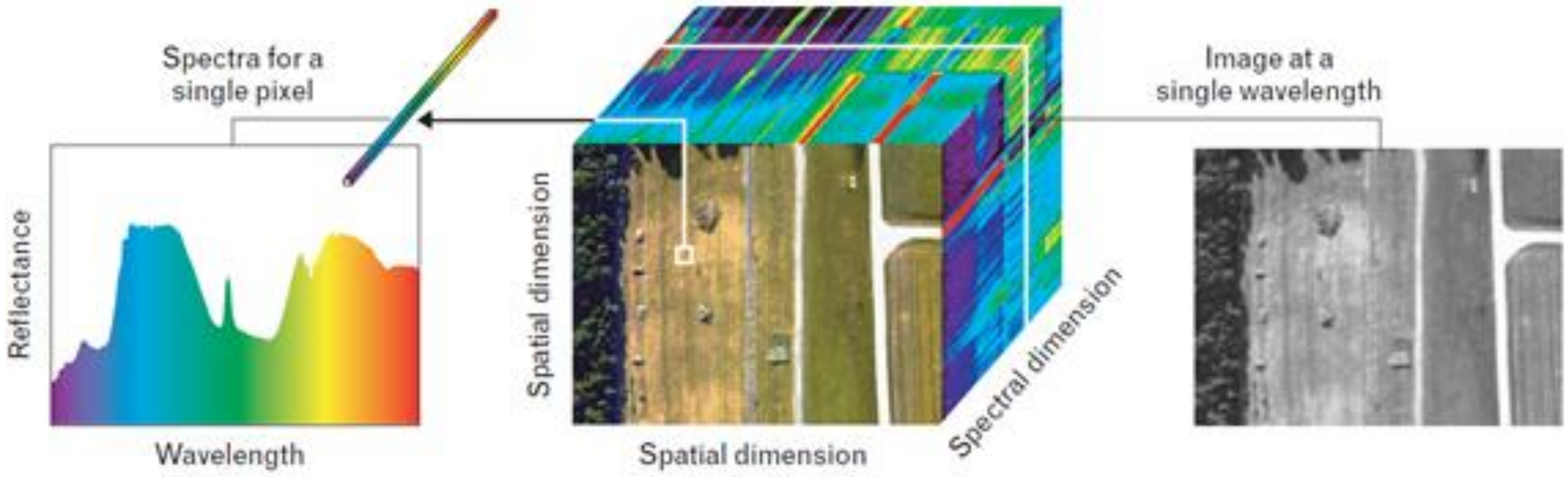
- Background
- Spectra Pre-processing
- Principle Component Analysis
- Principle Component Correlation
- Classification
- Binary Mixture Models
- Discussion
- Future Work

Background

- Mars Reconnaissance Orbiter (MRO)
  - Launched in 2005
  - Capturing images of Mars using several different instruments
  - Stilling sending back images
- CRISM
  - One of several imaging instruments onboard MRO
  - A pushbroom imaging VNIR spectrometer that produces hyperspectral images

www.nasa.gov

Background

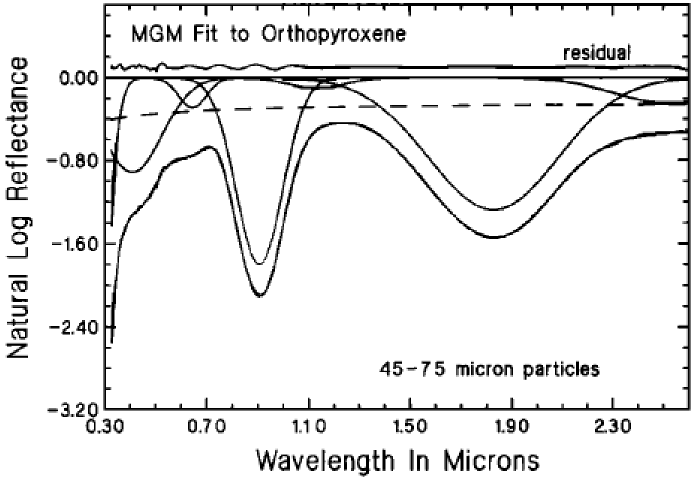


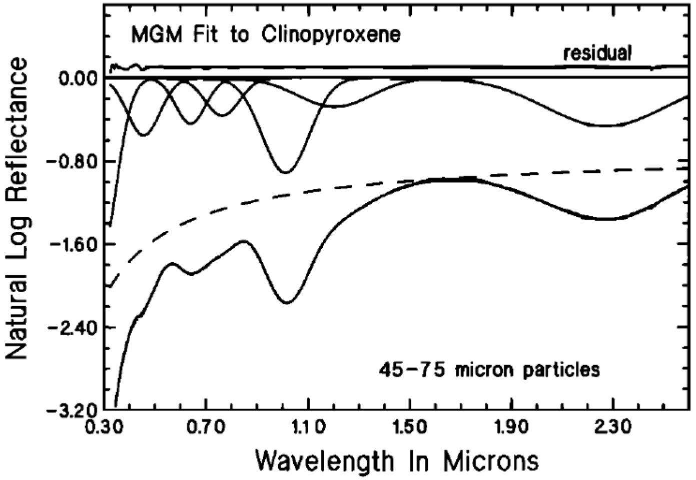
www.tankonyvtar.hu

- Hyperspectral Cube
  - A stack of images of the same scene at many different wavelengths
  - Has 2 spatial dimensions and 1 spectral dimension
  - Acquired with a spectral imager (grid of sensors)
  - Large hyperspectral libraries available (USGS, NASA, etc.)
- Spectra
  - A single pixel in a hyperspectral image
  - A plot of reflective Intensity versus wavelength (or frequency)
  - Can also be acquired with a spectrometer (single sensor)
  - Large spectral libraries available (USGS, RELAB)

Background

- MGM Laboratory Studies
  - Use modified-Gaussian modeling (MGM) to model spectral features
  - Use MGM parameters to determine mineral composition
    - Band Center
    - Amplitude
    - Full Width at Half Max (FWHM)
- Has been performed on:
  - Pyroxenes (Enstatite, Clinopyroxene)
  - Olivine Binary Mixtures (Forsterite, Fayalite)
  - Phyllosilicate Binary Mixtures (Kaolinite, Montmorillonite)

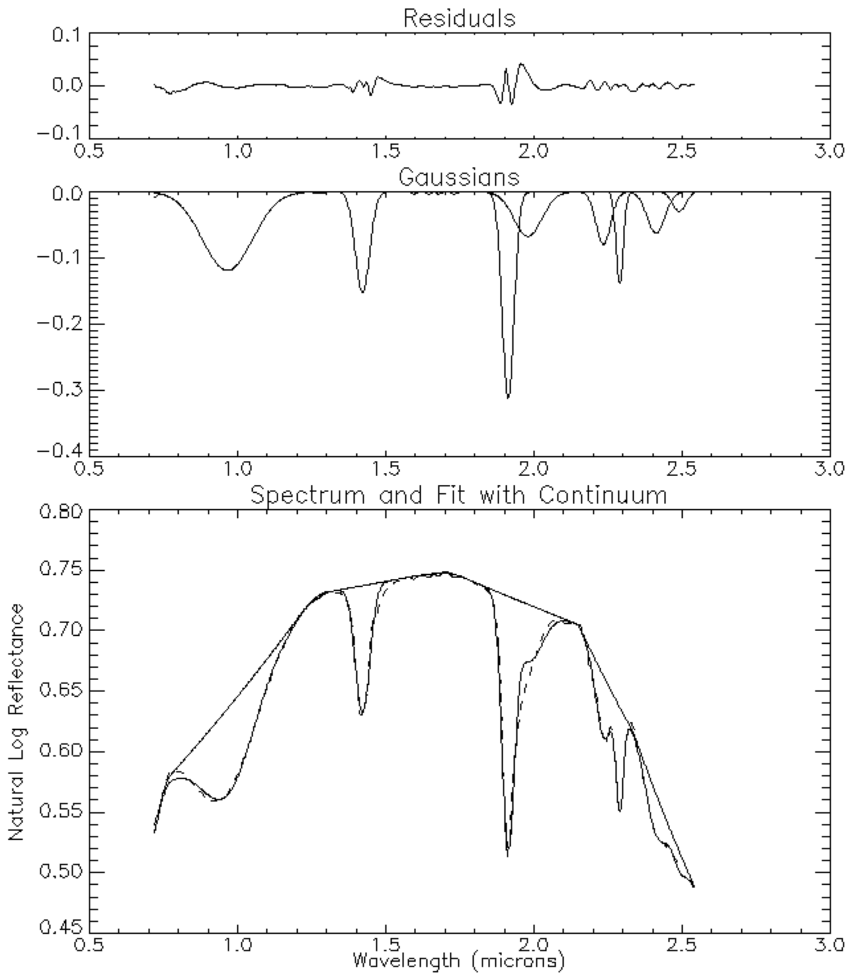




J.M. Sunshine, and C.M. Pieters

Background

- Modified Gaussian Modeling
  - Uses Optimization
    - Levenberg-Marquardt
    - Tarantola and Valette
  - Automated parameter initialization
  - Estimate number of Gaussians and their parameters
  - Applied to Lunar and Martian hyperspectral images

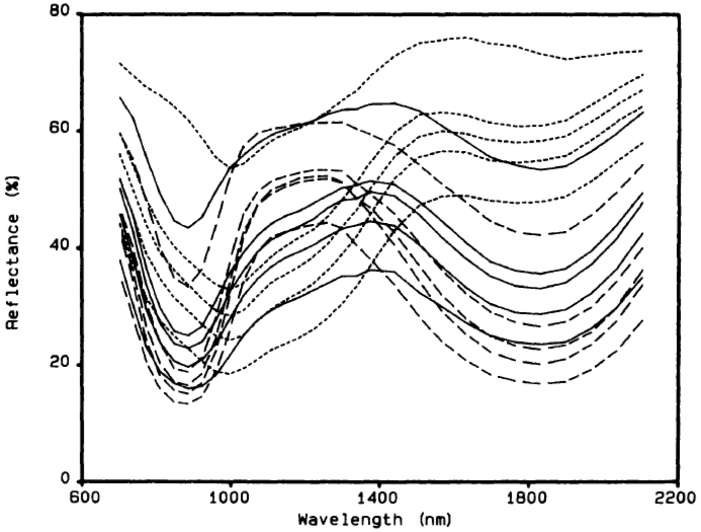


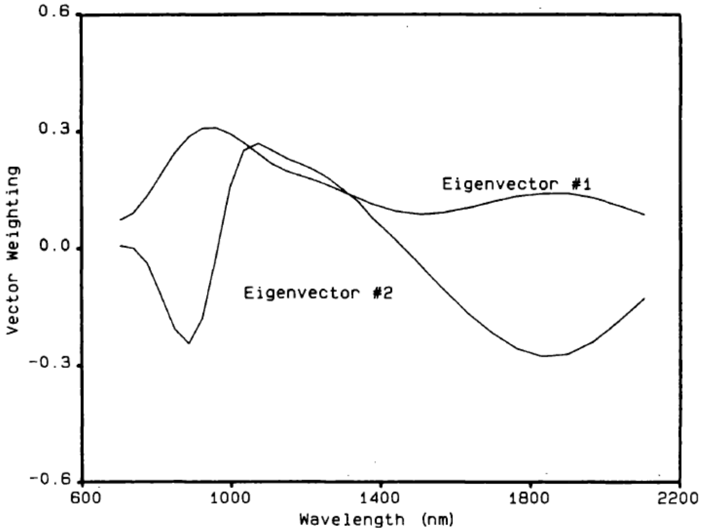
H.D. Makarewicz, M. Parente, and J.L. Bishop

Engineering Accreditation Commission

Background

- PCA Laboratory Studies
  - Use Principle Component Analysis (PCA) to derive a set of PC vectors and PC values
  - Retain a subset of PC values as principle components
  - Use principle components to determine mineral composition
  - No optimization
  - Has been performed on:
    - Tertiary Mixtures (Enstatite, Olivine, and Calcite)

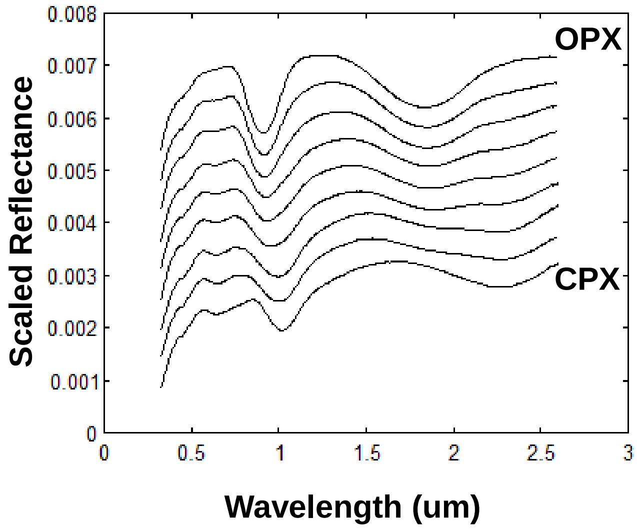


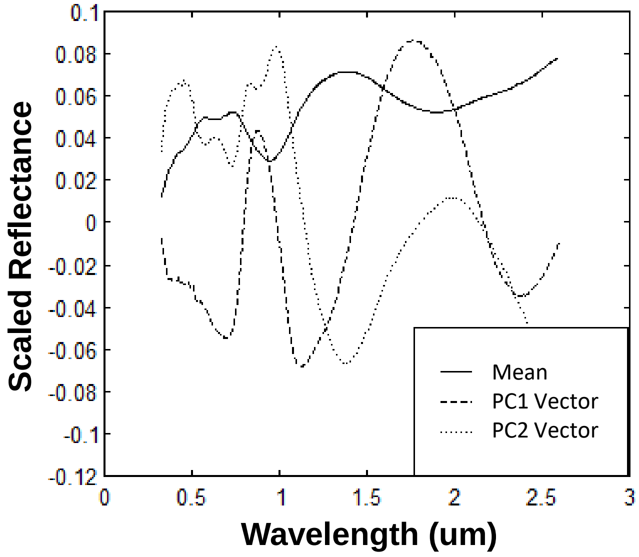


M. O. Smith , P. E. Johnson and J. B. Adams

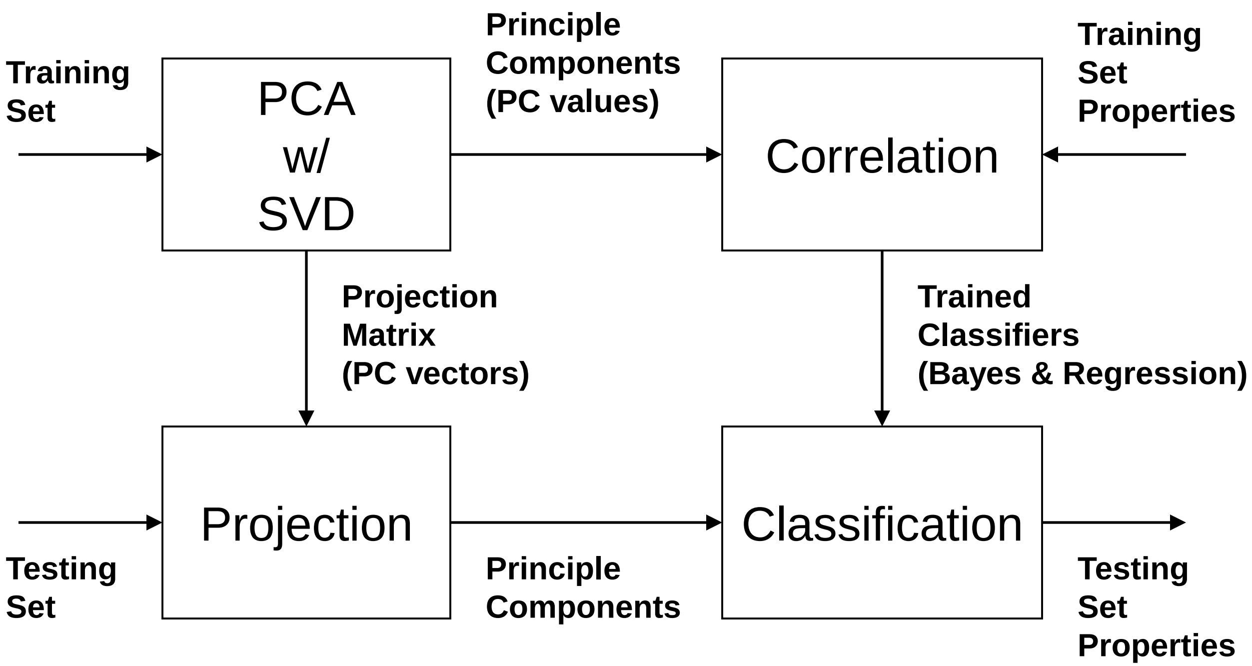
Objective

- Perform principle component analysis to determine mineral composition
  - Use M. O. Smith’s methodology
  - Apply to a variety of mixture data sets





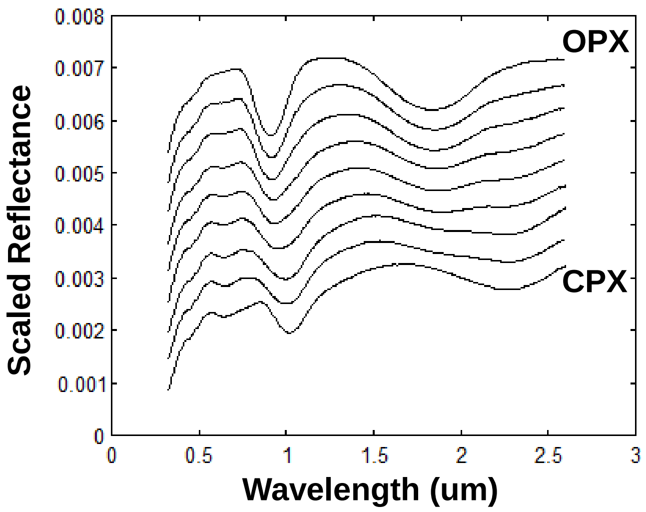
Methods



Training Set
PCA w/ SVD
Principle Components (PC values)
Correlation
Training Set Properties
Projection Matrix (PC vectors)
Trained Classifiers (Bayes & Regression)
Testing Set
Projection
Principle Components
Classification
Testing Set Properties

Spectra Pre-processing

- Crop
  - All spectra must have the same range of wavelengths.
  - 0.325 um to 2.6 um
- Resample
  - Fine enough to resolve spectral features
  - 0.05 um
- Normalize
  - Area under spectrum equal 1



Principle Component Analysis

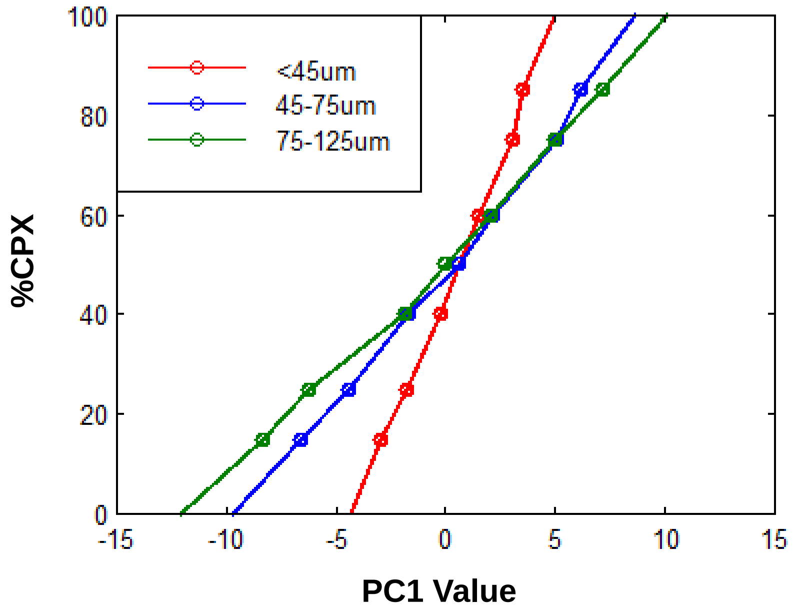
- Subtract the Mean Normalized Spectrum
  - Similar to 50% CPX
  - Features that are present in all sample spectra
- Computed Using Singular Value Decomposition
  - A set of PC vectors
  - Sets of PC values

Table 1 – Training Set Principal Components

Name	Ratio (%CPX)	Grain Size (um)	PC1 Value	PC2 Value
C1PP21	100%	<45	4.9834	14.6796
C1XP16	85%	<45	3.5224	14.1942
C1XP14	75%	<45	3.0545	13.1861
C1XP12	60%	<45	1.4616	17.5569
C1XP10	50%	<45	0.5630	18.3265
C1XP11	40%	<45	-0.1848	18.5786
C1XP13	25%	<45	-1.7919	19.8681
C1XP15	15%	<45	-2.9467	21.0338
C1PE30	0%	<45	-4.3693	22.3367
C1PP22	100%	45-75	8.6066	-5.8343
C1XP26	85%	45-75	6.1334	-7.1253
C1XP24	75%	45-75	5.0372	-9.2222
C1XP22	60%	45-75	2.1198	-6.2064
C1XP20	50%	45-75	0.6355	-6.2429
C1XP21	40%	45-75	-1.7180	-3.9304
C1XP23	25%	45-75	-4.4424	-4.2354
C1XP25	15%	45-75	-6.5749	-4.8977
C1PE31	0%	45-75	-9.7554	-4.4636
C1PP23	100%	75-125	10.091	-11.9483
C1XP07	85%	75-125	7.1403	-10.5483
C1XP05	75%	75-125	4.9318	-10.7730
C1XP03	60%	75-125	2.0399	-14.4355
C1XP01	50%	75-125	-0.0357	-14.2804
C1XP02	40%	75-125	-1.8646	-12.0175
C1XP04	25%	75-125	-6.2256	-11.0001
C1XP06	15%	75-125	-8.3312	-8.1490
C1PE32	0%	75-125	-12.0808	-14.4500

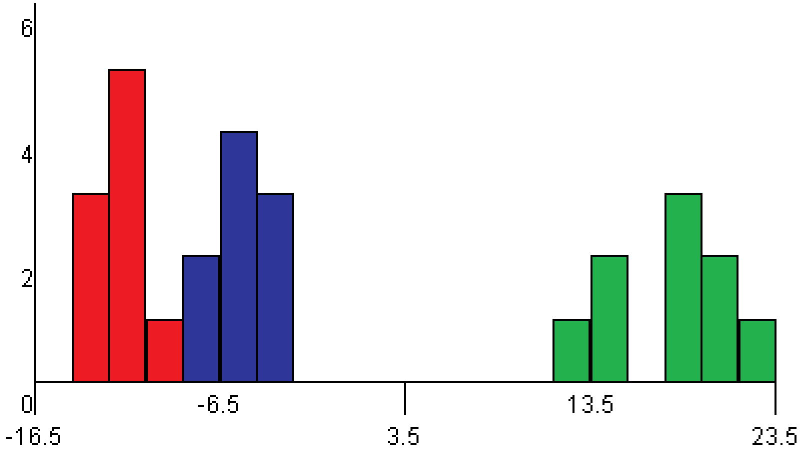
Principle Component Correlation

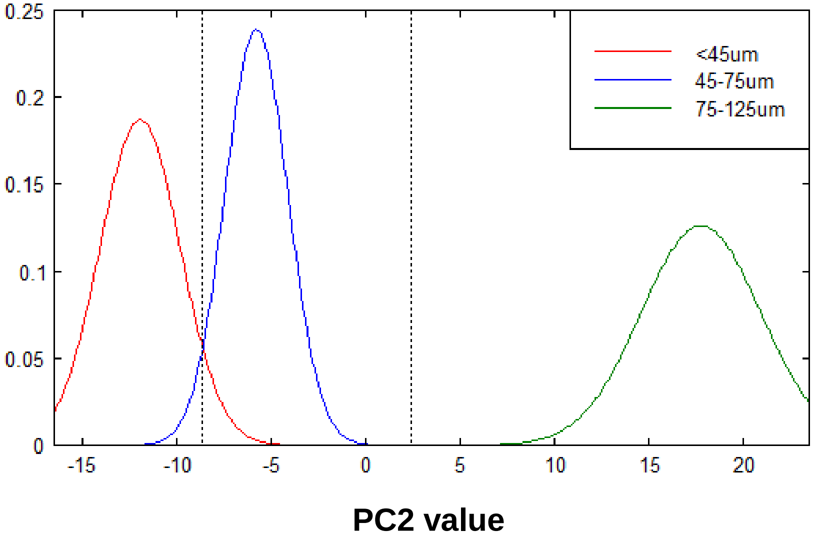
- First Principle Component was correlated with Percent Clinopyroxene
- Linear Regressions
  - <45um
    %CPX = 50 + 10×PC1 value
  - >45um
    %CPX = 50 + 5×PC1 value



Principle Component Correlation

- Second Principle Component was correlated with Grain Size
- Bayesian Classifer
  - <45um
    μ=17.75 , σ=3.16
  - 45-75um
    μ=-5.79, σ=1.67
  - 74-125um
    μ=-11.95, σ=2.13





Classification

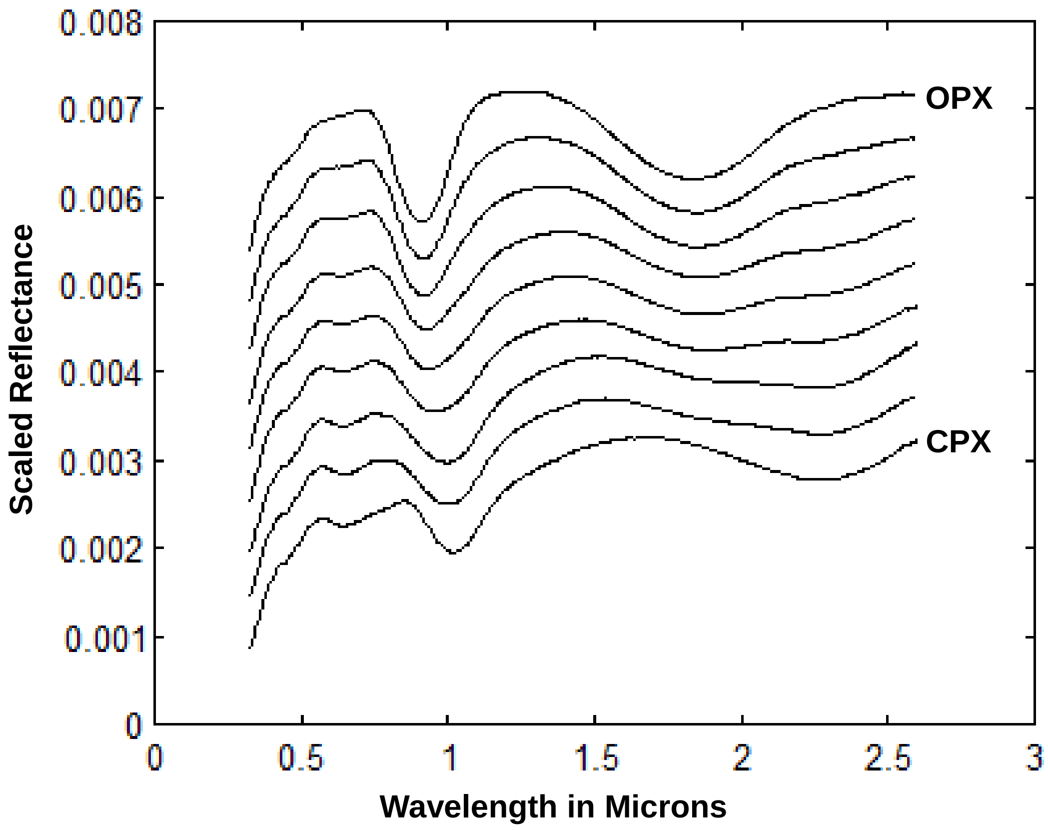
- Process
  - Pre-process spectrum.
  - Compute the first two principle component values using the first two principle component vectors.
  - Determine the grain size using the second principle component value.
  - Determine the percent composition using linear regression and the first principle component value.

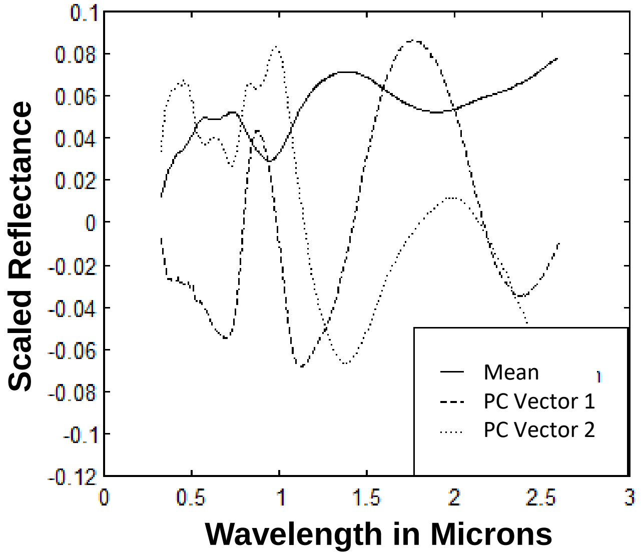
Results

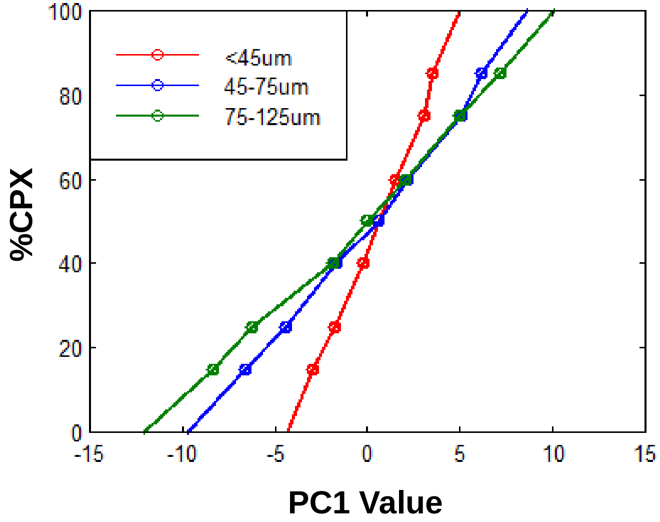
Table 2 – Test Set Grain Size				
Name	Actual Ratio (%CPX)	Actual Grain Size (um)	PC2 Value	Predicted Grain Size (um)
C1PE30	0%	<45	22.3367	<45
C2PE30	0%	<45	22.3361	<45
C4PE30	0%	<45	24.9357	<45
C5PP21	0%	<45	13.4838	<45
C2XP15	15%	<45	21.0338	<45
C5PE30	50%	<45	25.3391	<45
C4XP14	75%	<45	13.3947	<45
C1XP17	75%	<45	15.1154	<45
C2XP16	85%	<45	14.1942	<45
C4XP16	85%	<45	13.7501	<45
C1PE31	0%	45-75	-4.4636	45-75
C2PE31	0%	45-75	-4.4653	45-75
C3PE31	0%	45-75	-4.5904	45-75
C3PP22	100%	45-75	-7.9079	45-75
C1PP22	100%	45-75	-5.8343	45-75
C5PE32	0%	75-125	-10.9787	75-125
C1PP23	100%	75-125	-11.9483	75-125
C2PP22	100%	75-125	-5.8349	45-75
C2PP23	100%	75-125	-11.9487	75-125
C5PP23	100%	75-125	-11.8181	75-125

Table 3 – Test Set Percent Composition				
Name	Actual Ratio (%CPX)	Actual Grain Size (um)	PC1 Value	Predicted Ratio (%CPX)
C1PE30	0%	<45	-4.3693	6.3067
C2PE30	0%	<45	-4.3693	6.3074
C4PE30	0%	<45	-4.1384	8.6156
C5PP21	0%	<45	5.3783	103.783
C2XP15	15%	<45	-2.9467	20.533
C5PE30	50%	<45	-4.1980	8.0199
C4XP14	75%	<45	3.1054	81.0539
C1XP17	75%	<45	2.9716	79.7157
C2XP16	85%	<45	3.5224	85.2244
C4XP16	85%	<45	3.475	84.75
C1PE31	0%	45-75	-9.7554	1.2229
C2PE31	0%	45-75	-9.7553	1.2236
C3PE31	0%	45-75	-10.2877	-1.4383
C3PP22	100%	45-75	9.0428	95.2139
C1PP22	100%	45-75	8.6066	93.033
C5PE32	0%	75-125	-12.2364	-11.1822
C1PP23	100%	75-125	10.0919	100.4594
C2PP22	100%	75-125	8.6066	93.0332
C2PP23	100%	75-125	10.0919	100.4596
C5PP23	100%	75-125	10.3279	101.6394

Binary Mixture Models – Clinopyroxene-Orthopyroxene

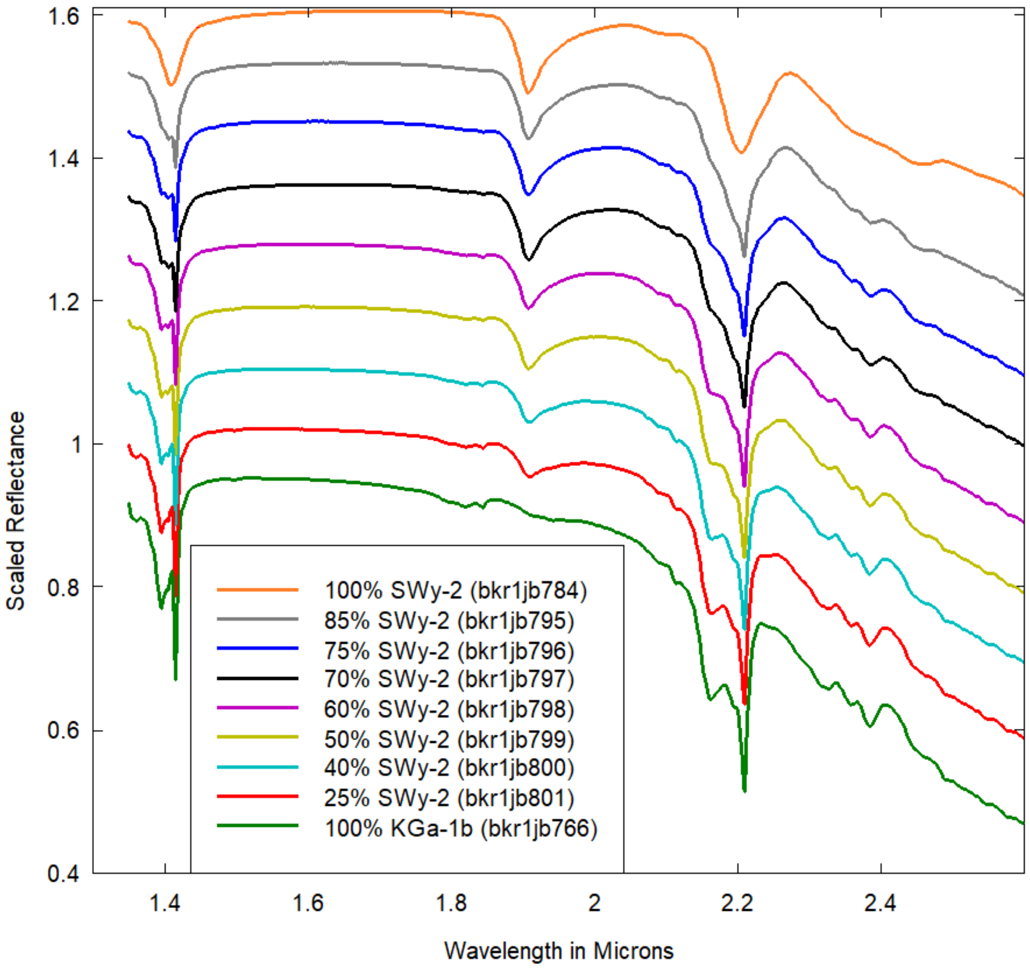


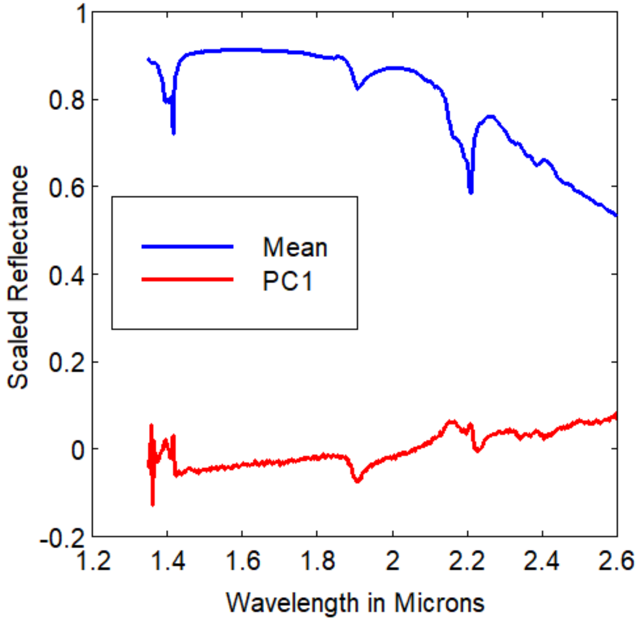


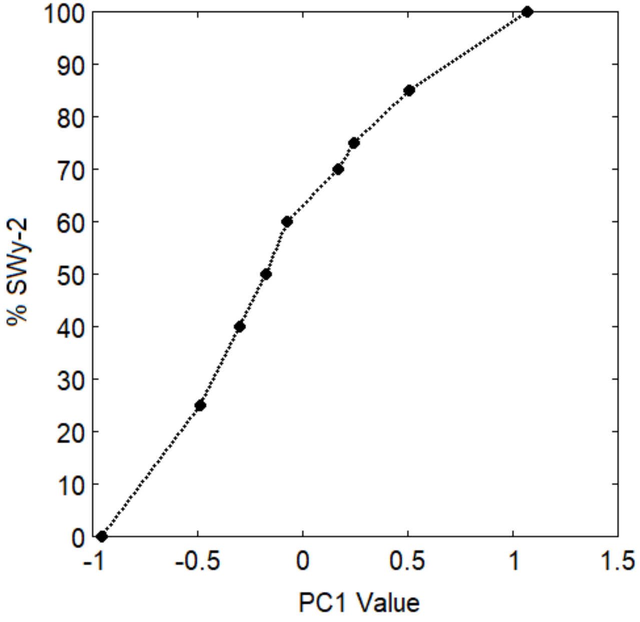


<45um: %CPX = 50 + 10 × PC1 value
>45um: %CPX = 50 + 5 × PC1 value

Binary Mixture Models – Kaolinite-Montmorillonite

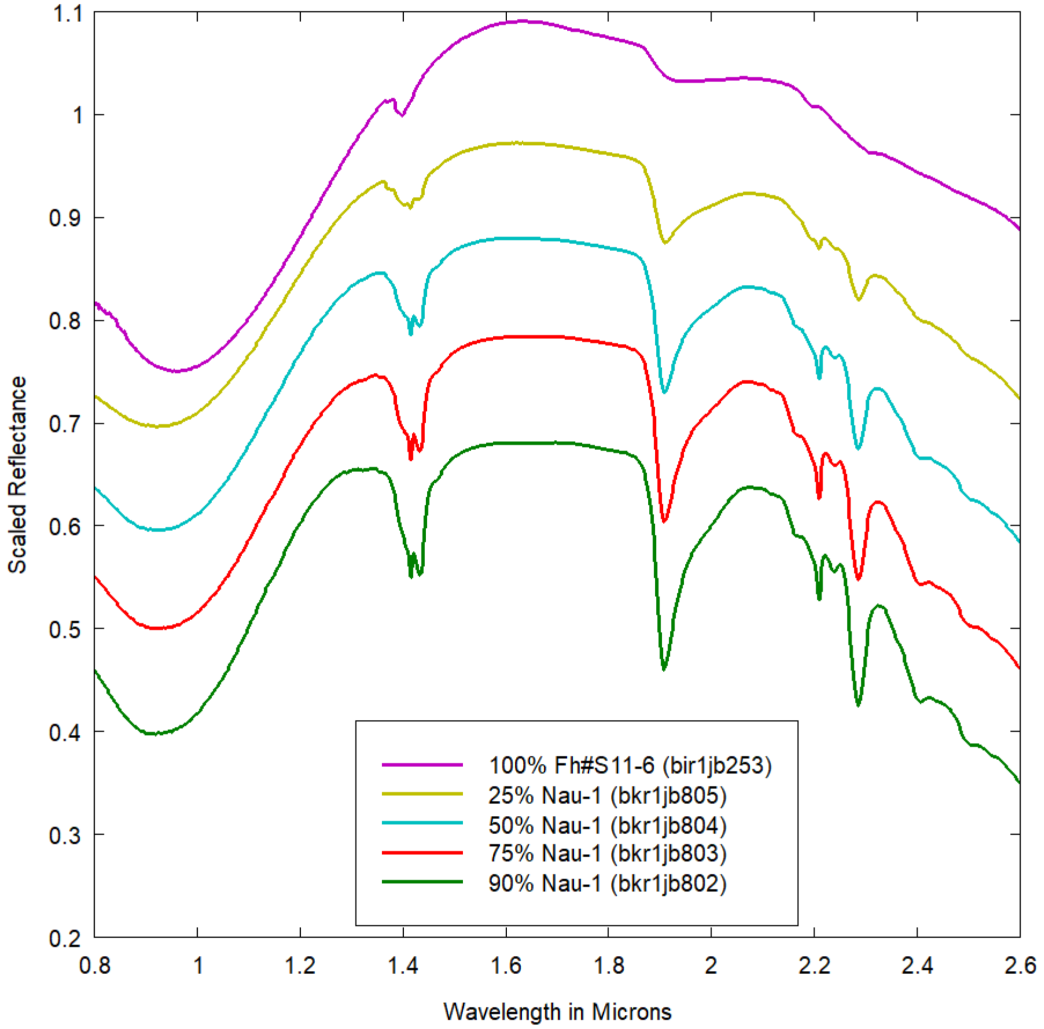


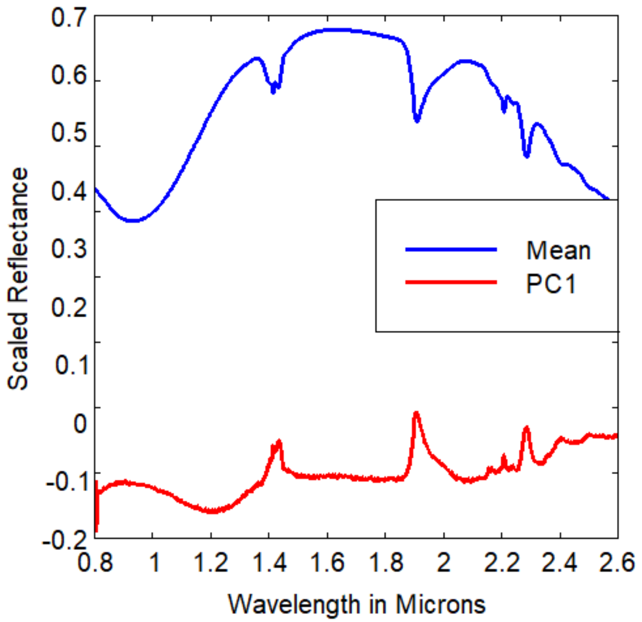


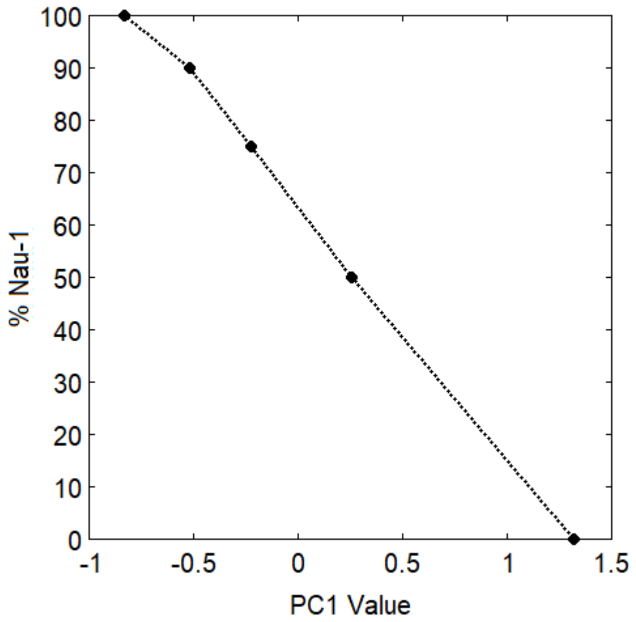


$\%SWY\text{-}2 = 56 + 52 \times PC1$ value

Binary Mixture Models – Nontronite-Ferrihydrite







%Nau-1 = 62 - 48 × PC1 value

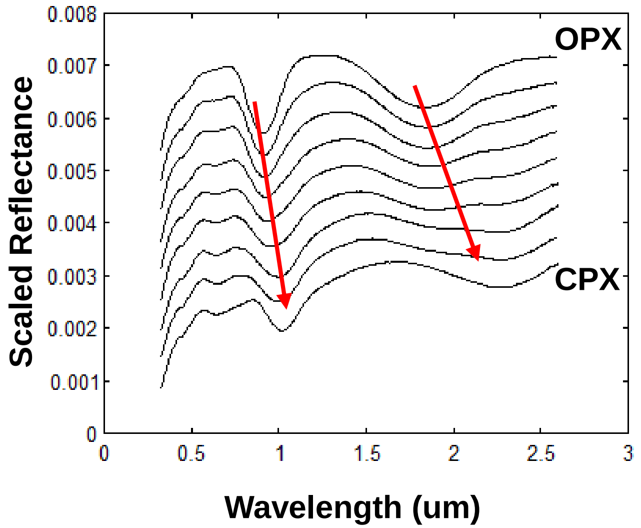
Engineering Accreditation Commission

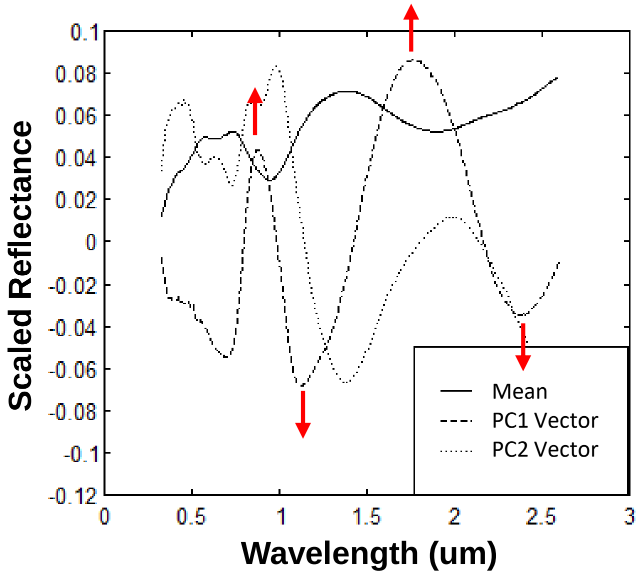
Discussion

- Normalization
  - Lose information about albedo
  - Allows analysis to be invariant to light source intensity
  - Important for generalizing algorithm to remotely sensed data
- Grain Size
  - Strongly correlated with albedo
  - May not always be detectable
  - If spectral features change with grain size, then it will still be detectable after normalization.

Discussion

- PC vectors
  - Used to form the projection matrix
  - Contain reflection and absorption band information
- Example
  - As PC1 value increases, %CPX increases, and PC1 vector is weighted more and absorption bands at 1.02um and 2.29um become stronger, but bands at 0.91um and 1.83um become weaker.





As %CPX increases

Future Work

- Apply algorithm to small sets of binary and tertiary mixtures for mass percent composition models.
- Apply algorithm to larger sets of various minerals for coarse identification of mineral presence.
  - Variance will increase amongst sample spectra
  - More principle components vectors and values will need to be retained
  - PC vectors will represent bands that are unique to different mineral groups
- Use models to identify minerals and estimate mineral abundance in lunar and Martian hyper spectral images.
  - Investigate residuals from model to identify trace minerals.