Supplementary Materials

Angela F. Gao, Brandon Rasmussen, Peter Kulits Eva L. Scheller, Rebecca Greenberger, Bethany L. Ehlmann Caltech

{afgao, brasmuss, kulits, eschelle, rgreenbe, behlmann}@caltech.edu



(a) PCA + K-Means

(b) Autoencoder + GMM on Unratioed Data

(c) Autoencoder + GMM on Ratioed Data

Figure 1: Map projected Claritas rise (CRISM) clustering Note that the lime-green class in (a) (which appears to be composed of cross-track dependent noise) is absent in both (b) and (c).



Figure 2: Spectra comparison on Claritas rise (CRISM) image Comparison of mean spectra for ratioed and unratioed clustering on Claritas Rise image (offset for clarity) (a) Cluster means on spectral data from Fig. 1b (left) with ratioed duplicate (right) (b) Cluster means of ratioed data from clusters in Fig. 1c. Key unique absorptions are at 1900 nm (water in minerals), 2300 nm, 2500 nm, and 2100 nm (serpentine) , 2200 (Aluminum-rich phyllosilicates). Important features are correlated with the vertical dotted lines.











(d) AE + GMM on CR Data

K-Means

(c) AE + GMM on Spectral Data

Figure 3: Oman core clustering with different methods (a) has 11 classes, while (b), (c), and (d) have 20. Red boxes show ROI presented in Fig. 1 of the main text. (c) clustering captures much more important mineralogical information than (b), and even contains additional zeolite veins that are not identified in (a).