Uncertainty-Based Unmixing of Space Weathered Lunar Spectra. M. Hess[1], T. Wilhelm[1], M. Arnaut[1], and C. Wöhler[1], [1]Image Analysis Group, TU Dortmund University, Germany, 44227 Dortmund. marcel.hess@tu-dortmund.de

Introduction: Quantitatively accurate abundance estimation by means of remotely sensed spectra is a non-trivial task. One approach is to exploit correlations between spectral parameters and laboratory composition data, to either determine the abundances of elements, like Fe or Ti [e.g., 1,2], or to directly estimate mineral abundances [e.g., 3]. A more sophisticated approach is spectral unmixing. Based on known spectra of endmembers the coefficients of a mixture that constitutes the best fit to the measured spectrum can be determined [e.g., 4,5].

In general, it has to be differentiated between linear mixtures, describing spatially separated endmembers, and intimate mixtures, where the light interacts with different minerals grains due to multiple scattering in the medium. These intimate mixtures are then non-linear, because it is no longer a superposition of the endmember reflectance spectra. On the Moon, the surface is covered by a porous layer of mineral grains. Therefore, a non-linear approach has to be employed. It has been shown that while the reflectance spectra have to be unmixed non-linearly, by converting the reflectance values to single scattering albedo (SSA) [6] the problem becomes a linear combination of SSA spectra [4].

On the lunar surface, the main minerals are plagioclase and pyroxenes, and in the mare areas additionally olivine and ilmenite [e.g., 7]. Since near-infrared (NIR) hyperspectral data of high spectral and spatial resolution have become available with the Moon Mineralogy Mapper (M^3) data set [8], the diagnostic absorption bands at 1-μm and 2-μm can be used to estimate the abundances of these minerals. One of the biggest challenges on the Moon, however, is that not only the composition but also the maturity has a strong influence on the measured spectra [9, 10]. Due to the influence of the space environment on the surface of a planetary body, which is not protected by an atmosphere, the spectra become darker and the spectral slope increases in the NIR [9,10]. Spectral libraries that include returned samples from the lunar surface, like the Lunar Soil Characterization Consortium (LSCC) catalog, provide mature endmembers and the mineral and elemental abundances are well characterized [11,12]. However, the simplex of mineral abundances is limited and some spectra, especially in the highlands, cannot be reconstructed.

One common and simple approach is to calculate the best-fit mixture for all possible endmember combinations in the least-squares sense and then select the combination with the lowest error to be the solution. For the LSCC catalog with relatively similar endmembers several combinations produce similar errors, therefore, the choice of the best solution is not clear. Small changes in the spectrum or in the error function lead to different mineral abundances. One approach to account for this issue is to model the uncertainties of the predicted endmember abundances. Then, an informed decision of the most likely solution can be made, and the deficiencies of the model can be understood. Bayesian inference [e.g., 13] provides a coherent framework to estimate the uncertainties of the model parameters and to conveniently include prior knowledge about the problem, without using hard constraints.

Methods and Data Set: This work is based on the M^3 global data set [8]. A global mosaic was created as in [14] at a resolution of 2 pixels per degree with thermal [15] and photometric [16] corrections applied, to remove the influence of topography and thermal emission. Then, a Gaussian Mixture Model (GMM) was used to create 64 clusters [17]. This way a global data set can be created by exploiting that many spectra are very similar. Clustering also reduces the influence of noise of the individual spectra. Each centroid is then converted to a SSA spectrum and the unmixing procedure is employed.

Bayesian inference can be used to estimate the parameters of a model and to simultaneously estimate their uncertainty. In the case of unmixing, the model is the linear superposition of the endmember SSA spectra weighted with the abundances. The probability of the parameters is then the posterior distribution. It is proportional to the prior distribution and to the likelihood, and thus also to their product. The likelihood is a measure on how well the current parameters of the model describe the measured data. The prior distribution includes assumptions about the distribution of the parameters. For this work, we are using uninformative priors uniformly distributed between zero and one for the abundances. To sample from the posterior distribution, a Metropolis-Hastings sampler [18] is used. We are not enforcing the sum-to-one constraint, because of possible differences in grain size, compaction, or due to a possible offset in the sensor. However, we include a normally distributed prior for the sum of the weights, centered at 1.0 with a standard deviation of 0.05. Thus, solutions close to a sum to one constraint are favored by the sampler.

The endmembers are taken from the LSCC catalog [11,12]. Additionally, a pure plagioclase sample taken from the RELAB library (http://www.planetary.brown.edu/relab/, ID: PL-EAC-029) was added in order to improve the reconstruction of nearly featureless highland spectra. This laboratory plagioclase sample was artificially space weathered with the model of [19] to fit the average continuum slope of the three brightest LSCC endmembers.

Results: The advantage of Bayesian modeling is that a posterior probability distribution of the parameters given the measurement is obtained, which contains information about the most likely solution as well as the uncertainties of the model parameters (abundances). Figure 1 shows the measured spectrum and the 95% confidence interval of the reconstructed spectra. While the variations in the spectrum are small, the abundances of the endmembers are varying strongly (Figure 2). If all samples are converted from endmember to normalized mineral abundances, the uncertainties of the mineral abundances can also be determined. As an example, the histogram of the plagioclase abundance for the shown spectrum is displayed in Figure 3. Finally, global maps can be created by using the cluster centroids and selecting the mean of the posterior distribution as the most likely solution. The plagioclase map is shown in Figure 4.

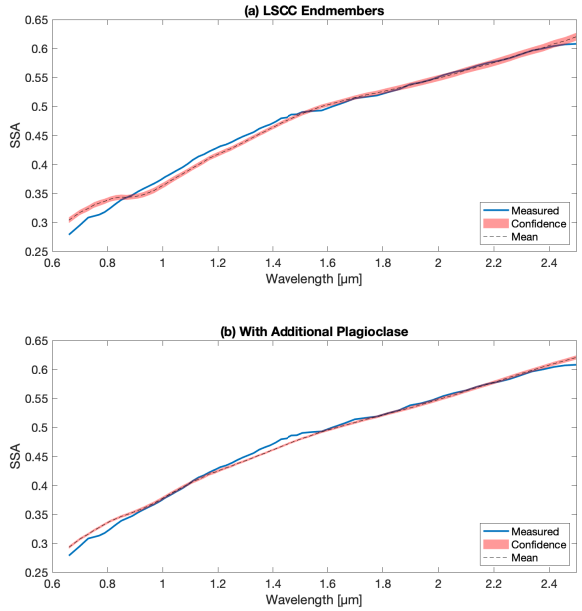

***Figure 1**: Reconstruction of typical highland spectrum. When using only the LSCC catalog (a) the absence of the 1-µm absorption is not represented in the reconstruction. When including an additional plagioclase endmember (b) the reconstruction is more representative for a typical highland spectrum.*

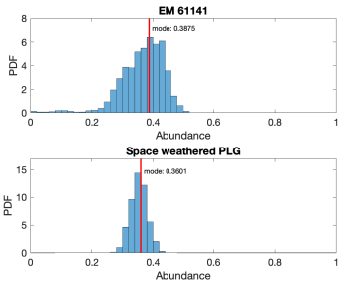

***Figure 2:** Histogram of the sampled posterior distribution of the two endmembers (EMs) 61141 from the LSCC catalog and the additional artificially space weathered plagioclase endmember. The sum of the twenty endmember means is 0.8704, the sum of the modes is 0.7508. Therefore, the other endmember contributions are negligible compared to the two shown in the histograms.*

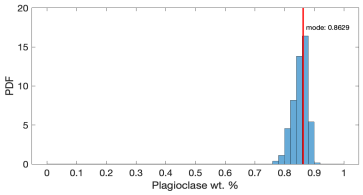

***Figure 3:** Histogram of the posterior distribution of the mineral abundance of plagioclase.*

Conclusion: In this work, a Markov Chain Monte Carlo (MCMC) approach to spectral unmixing has been introduced. Compared to classical optimization-based techniques the uncertainties of the model are also estimated, enabling an informed decision about the best solution. The inclusion of an artificially space weathered laboratory plagioclase spectrum improves the reconstruction results of a typical featureless highland spectrum.

References:

[1] Lucey, P. G., et al., (2000). JGR: Planets, 105(E8), 20297-20305. [2] Bhatt, M. et al., (2019). Astronomy & Astrophysics, 627, A155. [3] Sunshine, J. and Pieters, C. (1998), JGR Planets, 103, (E6), 13675-13688. [4] Keshava, N. and Mustard, J. (2002), IEEE SPM, 19, 1, 44-57. [5] Heylen, R. et al., 2014, IEEE JSTAEORS, 7, 6, 1844-1868. [6] Hapke, B., (2002). Icarus, 157, 2, 523-534. [7] Papike, J. J. et al., (1982). Reviews of Geophysics, 20(4), 761-826. [8] Pieters, C. M. et al. (2009). Current Science, 96, 4, 500-505. [9] Hapke, B. 2001, J. Geophys. Res., 106, 10039. [10] Lucey, P. G., & Riner, M. A. 2011, Icarus, 212, 451. [11] Taylor, L. A. et al., (2001). J. Geophys. Res., 106 (E11), 27,985-28,000. [12] Taylor, L. et al. (2010). J. Geophys. Res., 115, E02002. [13] Gelman, A., et al. (2013). Bayesian data analysis. CRC press. [14] Wöhler, C., et al., (2017b), Sci. Adv., 3, e1701286. [15] Wöhler, C., et al. (2017a), Icarus, 285, 118. [16] Wöhler, C. et al., (2014). Icarus, 235, 86. [17] Arnaut, M. et al., (2020). LPSC LI, abstract #3008. [18] Hastings, W. K. (1970). Biometrika. 57, 1, 97–109. [19] Wohlfarth, K. et al. (2019). Astron. J., 158, 2.

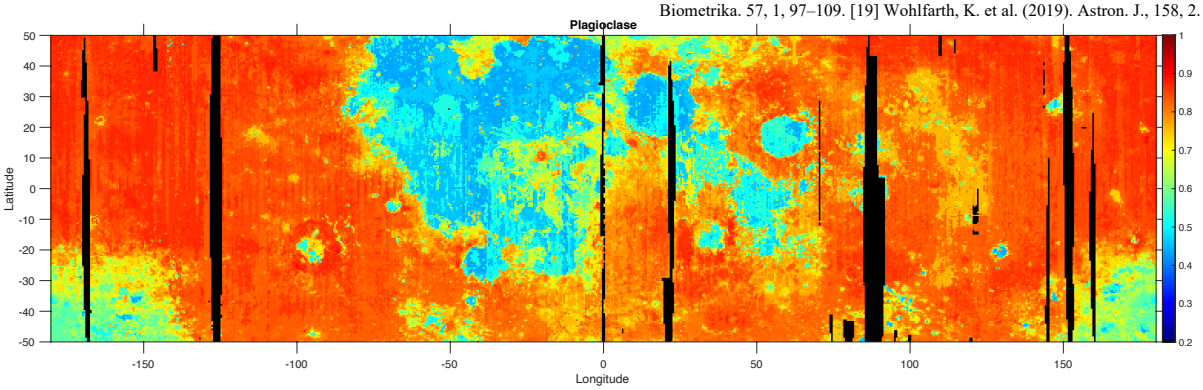

***Figure 4:** Map of the means of the posterior distribution for plagioclase based on the Bayesian inference approach. Black pixels indicate missing data.*