Machine Learning as a Tool to Aid in the Interpretation of Spectroscopic Data: Applications to Lunar and Planetary Exploration

Prabhakar Misra[1,2], Dina M. Bower[2,3], and Robert Coleman, Jr.[1]

[1]Howard University, Washington, DC

[2]NASA Goddard Space Flight Center, Greenbelt, MD

[3]University of Maryland, College Park, MD

Abstract

The precise spectroscopic identification of mineral polytypes and specific organic molecules is key in understanding planetary processes and the potential for life beyond Earth in the solar system. For *in situ* exploration, Raman spectroscopy has been chosen for the NASA Mars Perseverance Rover and upcoming ESA ExoMars missions because it is an information-rich, non-contact, non-destructive method for identifying and characterizing compounds. The technique is based on the energetic excitation of lattice bonds, which induces the scattering of light at a shifted wavelength to produce a spectrum unique to each compound composed of spectral peaks that express each vibrational mode. In materials of mixed compositions, like compositionally heterogeneous rocks and sediments in natural environments, structural commonalities of co-occurring minerals and organic compounds can result in similar peak positions that overlap. In such cases the misinterpretation of spectra can result in the misidentification of key information used to reconstruct environmental regimes or the detection of potential biosignatures. In addition, other factors, such as excitation wavelength, surface roughness, crystal orientation, and variations in temperature can affect the expression of individual Raman peaks. Machine learning can provide a means to disentangle the mixed signatures that occur in spectra from heterogenous targets by building algorithms capable of discerning subtle differences. Here we discuss an approach that incorporates a Matlab-based machine learning algorithm to study individual mineral samples as a starting point for more complex algorithms targeted for rocks and sediments. The present study focuses on Raman spectroscopy using visible (VIS) excitation laser (514 nm and 532 nm) and a near IR (NIR) excitation laser (at 780 nm) of an assortment of mineral samples typical for rocks on Mars and the Moon, namely olivines, three types of plagioclase minerals (anorthite, bytownite, labradorite), and pyroxenes (augite and enstatite). We have also begun to study the effect of temperature on the vibrational modes for the same mineral samples over a temperature range 300 – 473 K under NIR excitation. Our preliminary data shows, for example, that olivine samples from two different locations may exhibit the same typical symmetric and asymmetric stretch and bending vibrations for forsterite (Mg_2SiO_4); however, under increasing temperatures the peak intensities of ~ 820 cm^{-1} and ~ 845 cm^{-1} features exhibited by each sample differed. Our results also showed an enhancement of the Raman peak intensity for plagioclase samples as the temperature increased up to 373K, but a decrease at temperatures beyond that. Our project fosters STEM engagement by involving undergraduate and graduate students in machine learning with the step-by-step creation of an algorithm that will lead to a tool to guide Earth and planetary scientists in interpreting Raman spectra.

**Acknowledgments*: P. Misra and R. Coleman, Jr. acknowledge support from NASA Award # 80NCCS20M0019, NSF REU Award # PHY 1950379 & Howard University IDCR Grant# U100189; and D. Bower would like to acknowledge the support of the Internal Research and Development and Fundamental Laboratory Research Programs at Goddard Space Flight Center.