

**PREDICTING OLIVINE COMPOSITION USING RAMAN SPECTROSCOPY THROUGH BAND SHIFT AND MULTIVARIATE ANALYSIS.** L. B. Breitenfeld<sup>1</sup>, M. D. Dyar<sup>1</sup>, C. Carey<sup>2</sup>, T. J. Tague, Jr.<sup>3</sup>, and P. Wang<sup>3</sup>, <sup>1</sup>Dept. of Astronomy, Mount Holyoke College, South Hadley, MA 01075, breit221@mtholyoke.edu, <sup>2</sup>Univ. of Massachusetts Amherst, Amherst MA 01003, <sup>3</sup>Bruker Optics, Inc., Billerica, MA 01821.

**Introduction:** Olivine is well-studied within the planetary science community because of its presence in meteorites and on terrestrial surfaces as a typical component of basalt. Olivine group minerals are typically solid solutions between forsterite ( $\text{Mg}_2\text{SiO}_4$ ) and fayalite ( $\text{Fe}_2\text{SiO}_4$ ) with limited substitution of other cations such as Mn and Ni. Raman signatures of forsterite and fayalite have long been recognized to be distinctive based on studies of small numbers of samples [1-4]. Here, we use 113 well-characterized synthetic and naturally-occurring olivines to develop robust algorithms with quantitatively determined accuracies for measuring Mg and Fe in olivine through Raman spectroscopy. We compare the conventional univariate approach that relates peak position to composition with a multivariate analysis approach that leverages a much broader wavelength range and produces more accurate predictions.

**Olivine Modes:** Generally forsterite Raman bands above  $500\text{ cm}^{-1}$  can be classified as internal movement within the  $(\text{SiO}_4)^{4-}$  tetrahedral. Below this threshold, the peaks are caused by rotation and translation of the tetrahedral as well as magnesium motion [5]. Forsterite has 84 vibrational modes; only 36 are Raman active ( $11A_g + 11B_{1g} + 7B_{2g} + 7B_{3g}$ ) [5]. Key to this study are the two principle Raman peaks that form a doublet composed of five vibrational modes ( $2A_g + 2B_{1g} + B_{2g}$ ) [1,3,6,7]. Peaks in this doublet occur between  $\sim 815\text{-}825\text{ cm}^{-1}$  (DB1) and  $\sim 838\text{-}857\text{ cm}^{-1}$  (DB2) [3]. The  $B_{1g}$  and  $B_{2g}$  modes also have low intensity contributions to the doublet. The energy shift of the  $A_g$  stretch from the  $\text{SiO}_4$  tetrahedra is caused by changes in site geometry due to cation substitutions in the adjacent site. Cation substitutions between forsterite and fayalite thus result in band shifts [3]. Many previous workers (as noted above) have used the DB1/DB2 doublet to derive olivine composition, but this practice does not allow other information in the spectra to be encompassed, such as shifts arising from minor modes that affect the shape of the primary doublet and give rise to other, less prominent features elsewhere in the wavenumber range studied.

**Methods:** The 113 olivine samples examined were either a single crystal or in powdered form. To produce a powdered sample, first each sample was visually inspected and handpicked for purity. Then each grain was treated using oxalic acid (2 tsp. in 2 gal. of water) for 1 hour to remove surface weathering, followed by three cycles of washing and rinsing with clean water. As needed, samples were then crushed in a tungsten shatterbox or ground by hand in a diamond mortar. We chose to study both single crystals and powders because

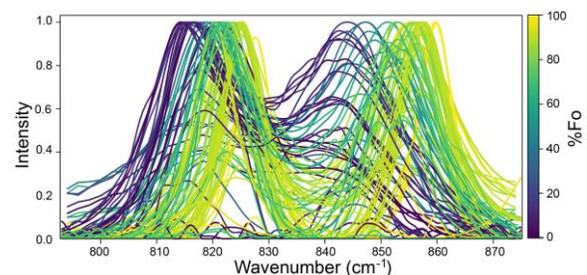
crystal orientation affects the Raman spectrum of a given sample [7] and we sought to evaluate the magnitude of the difference. When the quantity of the sample was too small to produce powder, then only a single crystal was analyzed.

Electron microprobe analyses of 10 spots on various grains of each sample were acquired at Brown University using standard operation conditions. Mg# number was calculated for each sample by normalizing the contents to contain only Mg and Fe as commonly done with the formula  $\text{Mg\#} = (100 \times \text{Mg}) / (\text{Mg} + \text{Fe})$ . Here Fo represents the total Fe contents ( $\Sigma\text{Fe}^{2+} + \text{Fe}^{3+}$ ). All samples were also analyzed by Mössbauer spectroscopy and the  $\% \text{Fe}^{3+}$  in most samples was zero.

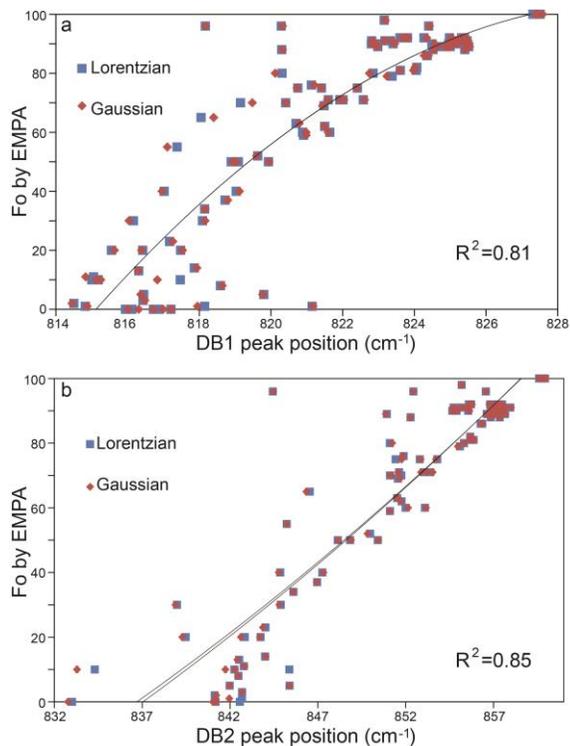
Spectra were acquired on two Bruker instruments: the BRAVO dual laser system (758/852 nm) and the Senterra 532 nm laser. The powdered olivine samples were analyzed on the BRAVO with three sample scans and an integration time of 10 s. Single crystals were analyzed on the Senterra using 10 mW laser power for two sample scans and integrated for 10 s.

All spectra were baseline corrected using AirPLS with smoothness set to 100 and normalized to account for arbitrary intensity differences between the two spectrometers. Finally, the DB1 and DB2 bands of the doublet were peak fitted for each spectrum using a Gaussian and Lorentzian method (**Figure 1**).

**Results:** Fo-Fa ratio can be estimated [3] using the central peak positions of the olivine peaks (DB1 and DB2) as previously used in the literature. Linear and second order polynomial fits relating peak position to composition were performed (**Figure 2**) and  $R^2$  values for fits to all data are provided in **Table 1**. Average coefficients from the 113 fits to each data set (from cross-validation, with one sample held out for prediction in



**Figure 1.** Raman spectra of olivine doublet (DB1 and DB2) of 113 samples acquired on Bruker's Senterra and BRAVO spectrometers. All spectra were baseline corrected using AirPLS and normalized. Spectra are color coded based on Fo content, where forsterite is represented with yellow and fayalite with purple.



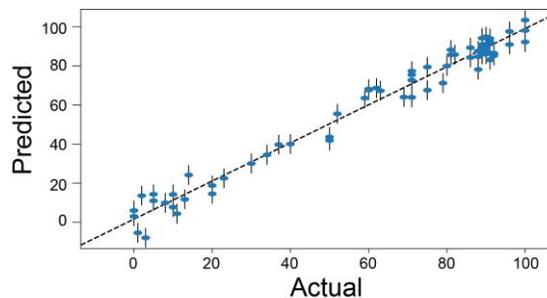
**Figure 2.** Peak centroids positions of the (a) DB1 and (b) DB2 doublet related to Fo content by a second order polynomial fit. Error bars are smaller than the symbols and are given in Table 1. Lorentzian and Gaussian peak shapes were used for all samples.

each trial) are also listed, along with the overall RMSE-CV errors in units of %Fo. The error bars from RMSE-CV on these models are  $\pm 13$ -26 %Fo based on these preliminary results.

**Table 1.** Comparison of Regression vs. Multivariate Methods for %Fo Predictions

Equation	RMSE-CV	R <sup>2</sup>
<b>DB1 ~820 nm</b>		
$y = 8.494x - 6914.4$	16.07	0.78
$y = -0.445x^2 + 738.96x - 306401$	26.28	0.80
<b>DB1 ~850 nm</b>		
$y = 4.725x - 3958.1$	12.83	0.85
$y = 0.0325x^2 - 50.415x + 19446$	24.11	0.85
<b>PLS prediction</b>	5.05	0.98

Multivariate analyses provide an alternative method for Fo-Fa predictions. The partial least squares (PLS) method used here regresses one response variable (%Fo) against multiple explanatory variables (intensity at each channel of the spectra). PLS predictions utilize every channel of the spectral range (here, we used 800-870  $\text{cm}^{-1}$ ), assigning coefficients to every single channel. Because PLS utilizes all available variables (channels) and eliminates multicollinearity (peaks whose intensities are dependent, as is the case for the doublet peaks in the Raman spectra of olivine), PLS generally



**Figure 3.** PLS predictions using both peaks DB1 and DB2, where both were distinguishable ( $n = 77$ ); a wavenumber range of 800-870  $\text{cm}^{-1}$  was used.

outperforms univariate methods for predictions in spectroscopic applications [8,9]. Spectra included in the PLS prediction were acquired on Bruker's Senterra spectrometer and were analyzed using tools on the superman web site [10]. When compared with univariate models (Figure 2), PLS results (Figure 3) show far better prediction accuracy, as demonstrated by the  $\pm 5$  %Fo error bar from RMSE-CV.

**Discussion:** It is quickly apparent that the multivariate method significantly outperforms the regression models, as observed in LIBS [8,9] and XAS [11,12] spectroscopies, among many others. In applications where the olivine being studied is pure and its diagnostic doublet is well-resolved, PLS is the preferred technique if optimum prediction accuracy is desired. However, in planetary applications where olivine may be one of an assemblage of phases mixed together in a Raman spectrum, then the univariate methods based on peak positions will be necessary. In either case, Table 1 gives a quantitative estimate of resultant errors on predictions of %Fo in olivine.

Further research, including additional screening of samples/spectra, experimentation with different energy ranges, and application of other multivariate methods such as the least absolute shrinkage and selection operator method (lasso), is ongoing. Thus coefficients reported in Table 1 should be considered preliminary until a penultimate model is achieved.

**Acknowledgments:** We are grateful for support from the Massachusetts Space Grant Consortium and the NASA RIS<sup>4</sup>E SSERVI node.

**References:** [1] Chopelas A. (1991) *Amer. Mineral.*, 76, 1101-1109. [2] Wang A. et al. (2004) *J. Raman Spectros.*, 35, 504-514. [3] Kuebler K. E. et al. (2006) *Geochim. Cosmochim. Acta*, 70, 6201-6222. [4] Ishibashi H. et al. (2011) *J. Raman Spectros.*, 43, 331-337. [5] McKeown, D.A. et al. (2010) *Amer. Mineral.*, 95, 980-986. [6] Iishi K. (1978) *Amer. Mineral.*, 63, 1198-1208. [7] Price, G.D. et al. (1987) *Mineral Mag.*, 51, 157-170. [8] Tucker J. et al. (2010) *Chem. Geol.*, 277, 137-148. [9] Dyar M. D. et al. (2016) *Spectrochim. Acta B*, 123, 93-104. [10] Carey C. et al. (2017) *LPS XLVIII*, Abstract #1097. [11] Dyar M. D. et al. (2012) *Amer. Mineral.*, 97, 1726-1740. [12] Dyar M. D. et al. (2016) *Amer. Mineral.*, 101, 1171-1189.