COMPARISON OF UNIVARIATE AND MULTIVARIATE CALIBRATION METHODS FOR GEOLOGICAL TRACE ELEMENTS WITH HANDHELD XRF.
J. L. King[1,2], J. C. Watts[1,2], M. D. Dyar[1], J. Bleacher[3], A. McAdam[3], J. Hurowitz[4], K. Young[5]. [1]Dept. of Astronomy, Mount Holyoke College, South Hadley, MA, 01075, [2]Harvey Mudd College, Claremont, CA, 91711 (pking@g.hmc.edu), [3]Goddard Space Flight Center, Greenbelt, MD 20771, [4]Dept. of Geosciences, Stony Brook University, Stony Brook, NY 11794-2100, [5]Johnson Space Center, ARES Astromaterials Group, Houston, TX, 77058.

Introduction: Laboratory x-ray fluorescence spectroscopy is a reliable and established technique in geochemical analysis. Recently, handheld XRF (hXRF) spectrometers have become available, and their capacity for rapid, *in situ* analysis has great potential for use in both Earth and planetary science. However, available technology focuses on mining and industrial applications, and is not yet well-suited for geochemistry. Moreover, calibration techniques that work well for high resolution lab XRF may not be as well-suited for the handheld units.

Currently, there are few hXRF calibration curves that relate spectral intensity to element concentration for geologically relevant elements and matrices. To realize the full potential of hXRF for planetary science, expanded calibration curves based on geological standards are needed. Here we present preliminary results of univariate and multivariate analysis (MVA) methods to create calibration curves for trace elements with the Olympus Innov-X DELTA Premium hXRF.

Samples and Methods: Samples used in this study were pressed powder pellets from the collections of rock powder standards in the Mount Holyoke College Mineral Spectroscopy Laboratory [1]. About 600 samples in this collection are different rock types doped with up to 10 wt.% of varying trace elements [2-4]; remaining samples are previously-studied rock
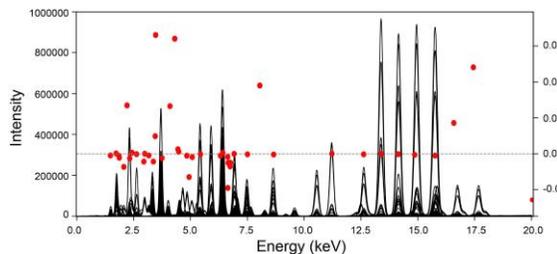
Figure 1. Plotted channel coefficients for Ga calibration.

powders from various localities. In this study, 805 samples were analyzed using the Olympus DELTA Premium hXRF. The instrument fires two beams on each run, with Beam 1 for heavy element analysis and Beam 2 for light element analysis, resulting in a total of 1610 spectra. Elements studied are listed in **Table 1**.

Univariate. PolyReg-CV, a web interface (http://bregman.cs.umass.edu:5693), was used to create univariate calibrations with ≤ 3 polynomial degrees; cross-validated leave one out root mean square errors (LOO RMSE-CV) were then calculated. Univariate calibrations were based on the area of an element's K-α peak using the Summed Area algorithm [6]. L-α and M-α peaks were also analyzed for heavy elements.

Multivariate. The lab website Project Superman [6] (nemo.cs.umass.edu:54321) was used to process the data and create multivariate calibrations. Only spectra with concentrations of the element at close to naturally-occurring levels (typically <5000 ppm) were included in calibrations. Both Lasso [7] and Partial Least Squares (PLS) [8] models were considered. To determine the most relevant channels on the spectrum for predicting element concentration, a Lasso model was trained with $\alpha = 1000$ and the coefficients were plotted along the spectrum, as shown in **Figure 1** for Ga. The higher the absolute value of the coefficient (ratioed to the intensity at that channel), the more important that channel is to the calibration. Through trial and error modeling, peaks occurring near those critical channels were used to determine which combination of spectral ranges yielded calibrations with the lowest LOO RMSE-CV.

Results and Discussion: Univariate models consistently yielded lower errors, although multivariate models occasionally prevailed (**Table 1**). When multivariate models were more accurate than univariate models, however, it was often by a small margin, while univariate models surpassed multivariate models sometimes by more than 300% (**Figure 2**).

Table 1. MVA and Univariate Calibration LOO-RMSE-CV

Element	Z	MVA (ppm)	Univariate (ppm)
Sc	21	81	37
Cr	24	374	145
Mn	25	407	279
Co	27	45	50
Ni	28	230	176
Cu	29	45	45
Zn	30	301	93
Ga	31	21	22
Se	34	112	92
Rb	37	121	38
Sr	38	112	63
Y	39	34	33
Zr	40	190	88
Nb	41	20	20
Mo	42	20	27
Sn (K-α)	50	19	13
Cs (L-α)	55	19	29
Ba (K-α)	56	184	202
La (K-α)	57	226	94
Ce (L-α)	58	89	60
Pb (L-α)	82	301	55

Elements for which the multivariate calibration was more successful are highlighted in blue.
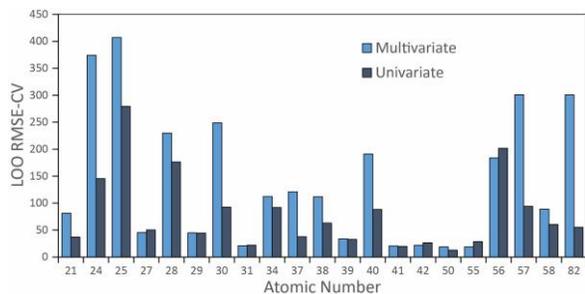
Figure 2. Bar graph of LOO RMSE-CV for multivariate and univariate calibrations.

In multivariate models, spectral regions most important to predicting element concentrations were sometimes unexpected but show evidence for geochemical affinities among trace elements. Se, for example, was best predicted by the Fe K-β peak rather than by a Se peak. **Table 2** lists the peaks that appeared in the best multivariate calibration for elements tested.

Ga, Nb, Mo, and Sn formed a group of elements that were useful in predicting one another's concentrations with multivariate analysis (highlighted in green in **Table 2**). This phenomenon is likely because these trace elements naturally appear in similar concentrations, perhaps because of their similar atomic radii. In an octahedral geometry, which is common in geological samples, Ga, Nb, Mo, and Sn have corresponding atomic radii of 0.62, 0.64, 0.65, and 0.69 Å. This phenomenon echoes trace element camouflage, where a trace element often forms minerals alongside a more common element with the same ionic potential. However, while Ga, Nb, Mo, and Sn have similar atomic radii, their ionic potentials are too different for camouflage to be their cause of their grouping. More research is needed to better understand why certain elements are useful in predicting others.

Cropping element concentrations to naturally occurring concentrations significantly lowered RMSE values compared to models using the full suite of doped standards (concentrations up to 10 wt.%) [6].

Conclusions and Future Work: Errors on these models still lack the accuracy necessary for proper quantitative analysis with hXRF units. While this abstract offers preliminary insight into univariate and multivariate calibrations for hXRF spectrometers, much work remains to be done.

Multivariate analyses works best in types of spectra where many different features useful to prediction occur or where subtle differences in peak shapes are observed. In hXRF, there are only a few peaks in the observed energy range that can result from any single element, so there is little other information that can be leveraged by MVA. XRF spectra acquired using a more sensitive, higher resolution full-scale laboratory unit using longer count times might provide better resolution and more detailed peak shapes. MVA models based on those data might utilize variations in peak shapes and possibly boost MVA accuracies beyond those of univariate.

Combined approaches might also prove useful for quantification. MVA could be used to find the most useful channels for any given element, and then peaks associated with those channels could be combined and applied to univariate calibrations to improve RMSE. For example, an improved Se calibration results when a univariate calibration uses the Fe K-β peak area.

Table 2. Element Peaks found to be Useful in Multivariate Calibrations of Trace Elements

Element	Z	Element Peaks in Most Successful Multivariate Calibration
Sc	21	K, Ca, Sc
Cr	24	Mg, Cr (K-α, K-β)
Mn	25	K, Ca, Ti, Cr, Mn
Co	27	Co (K-α, K-β), Zn (K-α, K-β), Mo
Ni	28	Ni (K-α, K-β)
Cu	29	Mg, Cu, Zn
Zn	30	Zn (K-α, K-β), Ga
Ga	31	Ga, Nb, Mo
Se	34	Fe (K-β)
Rb	37	Fe, Rb
Sr	38	Ca, Sc, Ti, Rb, Sr
Y	39	La (L-α), Ce (L-α), Rb, Y (K-α, K-β)
Zr	40	Zr (K-α, K-β), Mo
Nb	41	Ga, Nb (K-α, K-β), Mo, Sn
Mo	42	Ga, Nb, Mo, Sn
Sn	50	Ga, Mo, Sn
Cs	55	Ga, Sn, Cs
Ba	56	Sn, Cs, Ba, La, Ce
La	57	Ca, Sc, Ti, Cr, Ce (L-α), La (L-α)
Ce	58	K, Pb (L-β), Rb, Y (K-β), Zr (K-β), Nb (K-α, K-β), Mo (K-α, K-β)
Pb	82	Pb (L-β)

All peaks are K-α peaks unless otherwise specified in parentheses. Ga, Nb, Mo, and Sn, highlighted in green, were frequently useful in creating calibrations for one another.

This study utilized several standards for which only approximate bulk compositions were available. Complete geochemical characterization of those standards is currently underway; once completed, they may change our preliminary conclusions given here.

Overall, this study sheds light on some of the work needed to maximize the potential of hXRFs for research in geological and planetary science.

References: [1] Dyar M. D. et al. (2016) *LPS XLVII*, Abstract #2205. [2] Breves E. A. et al. (2015) *LPS XLVI*, Abstract #2338. [3] Lepore K. et al. (2017) *LPS XLVIII*, Abstract #1293 [4] Mackie J. et al. (2017) *LPS XLVIII*, Abstract #1292. [5] Carey C. et al. (2017) *LPS XLVIII*, Abstract #1097. [6] King J.K. et al. (2018) *LPS XLIX*, Abstract #1172. [7] Tibshirani, R. (1995) Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society, 58, 267-288. [8] Wold, S., Martens, H., and Wold, H. (1983) The multivariate calibration problem in chemistry solved by the PLS method. Lecture Notes in Mathematics, 973, 286-293.