A MULTIVARIATE MANGANESE CALIBRATION MODEL FOR CHEMCAM. Patrick J. Gasda (gasda@lanl.gov)[1], Ryan Anderson[2], Agnes Cousin[3], Olivier Forni[3], Samuel Clegg[4], Ann Ollila[1], Nina Lanza[1], Sarah Lamm[4], Roger C. Wiens[1], Sylvestre Maurice[3], Olivier Gasnault[3], Adriana Reyes-Newell[1], Dorothea Delapp[1], Roberta Beal[1]. 1) LANL, NM, USA; 2) USGS, Flagstaff, AZ, USA; 3) IRAP, Toulouse, France; 4) Kansas State University.

Introduction: Manganese is an important element for elucidating past habitability, pH, and redox, of water on ancient Mars [1–4]. The ChemCam Laser Induced Breakdown Spectroscopy (LIBS) instrument onboard the NASA *Curiosity* rover can measure the elemental composition of targets on Mars [5–7]. Manganese, like other transition metal elements, has many atomic emission lines in the range of the ChemCam LIBS spectrometers. The Mn emission lines at 403 nm have been used in the past to produce a univariate model for Mn quantification [2,8]. Here, we have produced a multivariate model that seeks to improve on the accuracy of bedrock MnO compositions by using an expanded dataset of 428 standards. The multivariate model uses Partial Least Squares (PLS) and Least Absolute Shrinkage and Selection Operator (LASSO) multivariate techniques with blended sub-models [9].

Methods: *Data Collection and Pre-processing.* A standard set consisting of 428 standards was analyzed using the ChemCam engineering model at LANL from 1.6 m distance (5 average spectra were collected on each standard consisting of 50 shots averaged in each point) under a Mars-like atmosphere. The standard set covers a range of Mn compositions from 27 ppm to ~100% Mn (pure Mn metal) and contains a variety of rock matrices (e.g., rock, mineral, Mn ores, and synthetic materials). We use the Python Hyperspectral Analysis Tool [10] and the associated graphical interface for point spectra analysis [10] to preprocess the data and evaluate multivariate regression models. Each spectrum is normalized by the sum of the total emission for each spectrometer [e.g., 7]. A "peak area" (PA) preprocessing technique is used [e.g., 11], wherein an algorithm finds the local minima and maxima of the average spectra of the dataset. The process then bins the emission between each pair of minima and assigns the result to the wavelength of the corresponding maximum. Thus, the total number of channels is reduced by ~10x, speeding up processing time.

Multivariate Models. The dataset was split into 5 folds of similar distributions of MnO content, 1 was held out as the test set, and the other 4 were used for cross validation and the training set. PLS and LASSO techniques were found to be the best types of models to use via cross validation analysis. Submodels for PLS and LASSO were trained for the full range of data, 0–10 wt.% MnO, and 0–1 wt.% MnO. The blending of the submodels are optimized on the training set. The test set is used to evaluate the accuracy of the optimized models on novel data, and accuracy is quantified using root mean squared error of prediction (RMSEP).

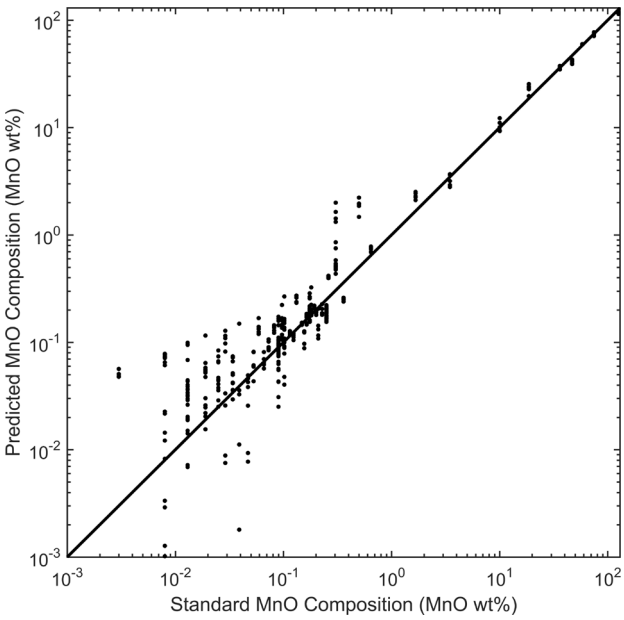

Fig 1: Test Set predictions of the multivariate model compared to a 1:1 line on a log scale plot.

A two-tiered or "double blending" model was used. At the first stage, a PLS full distribution model is used to select between the low 0–10 wt.% MnO LASSO model and the full LASSO model. This initial blended model is then used to select between a lower 0–1 wt.% MnO and itself as a full model.

A local RMSEP method was used to evaluate RMSEP for each predicted MnO value based on the nearest 40 test set samples, and then smoothed. The smoothed local accuracy is more representative of the model performance than a single full model RMSEP value. This method can be used to estimate the quantification limit, as the point where the local RMSEP is indistinguishable from zero, i.e., where the predicted value is greater than the sum of the local RMSEP value and the local RMSEP of the predicted value.

Precision is estimated by taking the standard deviation of each test set prediction of a standard (5 predictions per standard excluding outliers).

Results: The results for some of the cross validation are summarized in Table 1, and the Test Set is shown in Figure 1. The full double blended model RMSEP accuracy is 1.39 wt.% MnO and an average precision of 0.2 wt% MnO (both 1σ). These mean values do not fully reflect the model performance, and the local RMSEP accuracy is shown in Figure 2. The local RMSEP accuracy and precision scales with MnO abundance, resulting in a model accuracy of 0.03 wt.% at the quantification limit (0.05 wt.% MnO), and other values are listed in Table 2.

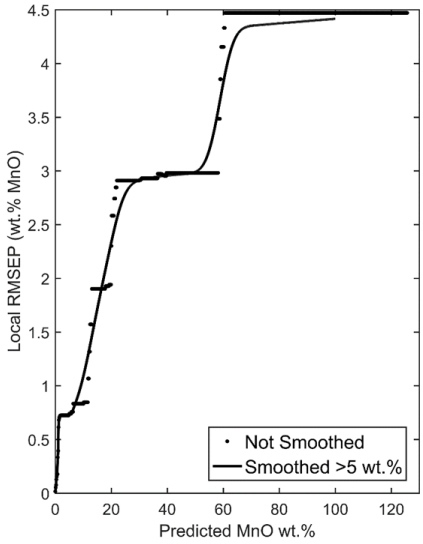

Fig 2: The local RMSEP based on 40 nearest test set predictions (squares) and smoothed curve for values >5 wt.% MnO (black curve).

Discussion: We have applied this model to Mars data as a sanity check. This analysis found that the obtained abundances scale with Mn peak intensity of the known Mn features in the LIBS spectra. The new abundance values for MnO, compared to the model developed in [2], are ~40 rel.% lower. Figure 3 shows the distribution of new MnO predictions using the double blended model for Mars targets up to sol 2729, showing that 95% of all Mars data lies in within the 0–0.22 wt.% MnO range. The highest values predicted from Mars data range up to ~22 wt.% MnO.

Table 1: Summary of Root Mean Square Error of Cross Validation (RMSECV) for some of the models tested.

Model Range (MnO wt.%)	*Method*	*RMSECV*	*Alpha/Fold*
Full	LASSO	2.486	8.43E-5
Full	PLS	2.620	15
0-10	LASSO	0.296	4.53E-5
0-1	LASSO	0.066	5.93E-6

Conclusions: This new robust multivariate model validated with laboratory data has improved accuracy for MnO predictions of ChemCam data for the *Curiosity* rover compared to previous methods [2,8]. The improvements, especially for values <0.4 wt.% MnO where the RMSEP accuracy is ≤0.08 wt.% MnO and precision is ≤0.05 wt.% MnO, will help with distinguishing the variations in MnO within the typical bedrock within Gale crater and help with interpretations of data from Mars.

Acknowledgements: NASA Mars Exploration program, CNES, France; Carlsberg Foundation

References: [1] Lanza et al. (2014) GRL, 41(16), 5755–63. [2] Lanza et al. (2016) GRL, 43(14), 7398–407. [3] Gasda et al. (2019). 50th LPSC #1620. [4] Frydenvang et al. (2020), JGR:P, 125(9). [5] Wiens et al. (2012) SSR, 170(1–4), 167–227. [6] Maurice et al. (2012) SSR, 170(1–4), 95–166. [7] Clegg et al. (2017) Spectrochim B, 129, 64–85. [8] Gasda et al. (2020) LPSC Abstract 1644. [9] Anderson et al. (2017) Spectrochim B, 129, 49–57. [10] https://github.com/USGS-Astrogeology/PyHAT; https://github.com/USGS-Astrogeology/PyHAT_Point_Spectra_GUI [11] Clegg et al. (2017) 48th LPSC 2439.

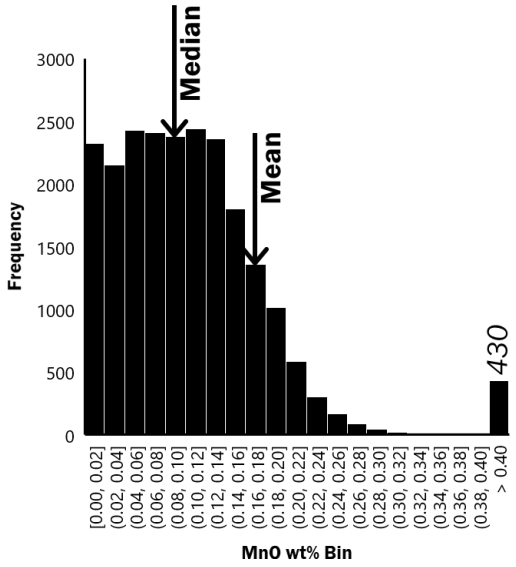

Fig 3: Histogram of MnO abundance for Mars targets up to sol 2729.

Table 2: Estimated accuracy and precision for selected MnO (wt.%) compositions, all 1σ.

Abundance	*RMSEP Accuracy*	*Precision*
0.05	0.03	0.01
0.1	0.05	0.02
0.4	0.1	0.05
1.0	0.4	0.09
100	4.4	2.1