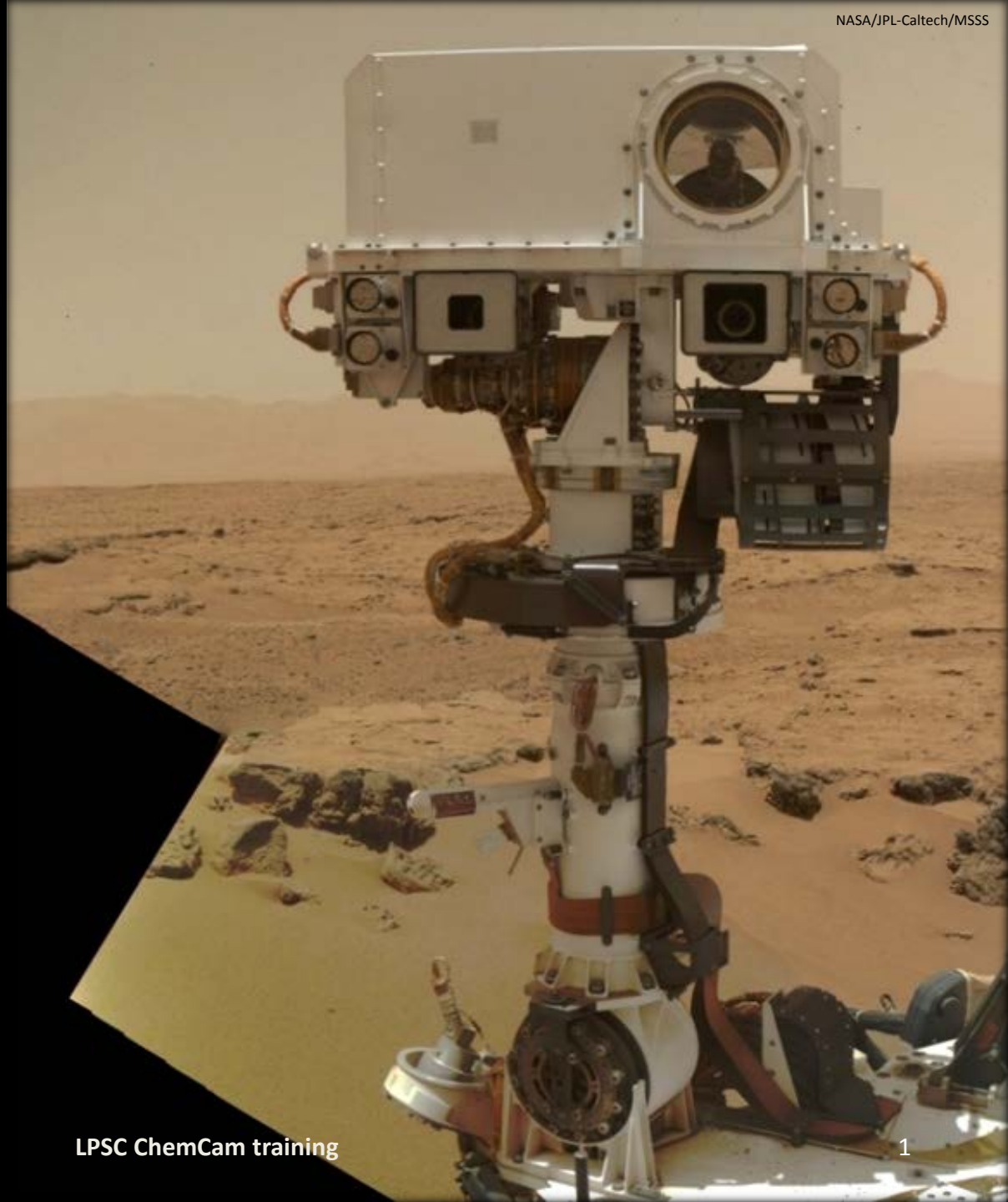
Lunar and Planetary Science Conference, March 18, 2015

ChemCam Data Processing - Quantitative Calibration



R. Anderson, S. Bender, S. Clegg, A. Cousin, M.D. Dyar, C. Fabre, O. Forni, J. Frydenvang, O. Gasnault, J. Lasue, S. Maurice, V. Sautter, R. Tokar, R. Wiens and the ChemCam team

Quantitative Analysis

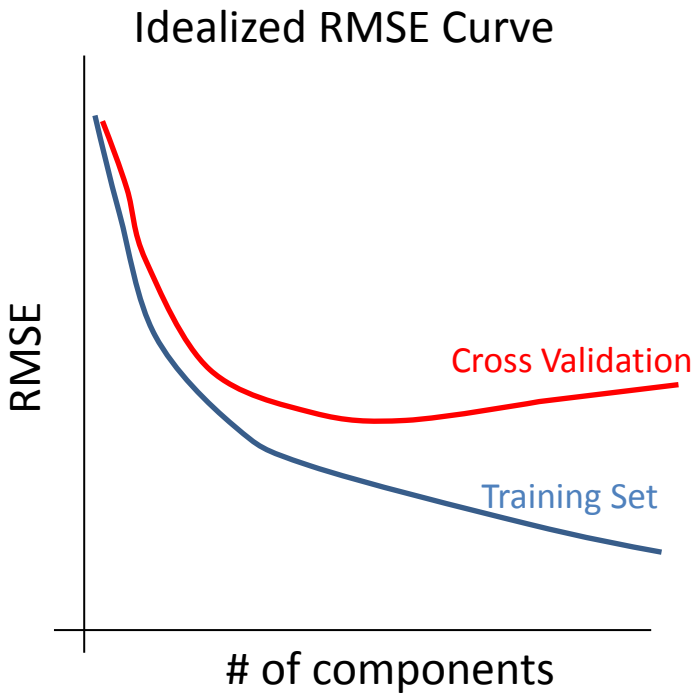
- Goal: to find a regression model that accurately converts from ChemCam LIBS spectra to elemental composition
- Use spectra from targets of known composition to generate the model
- How to evaluate model performance? Use Root Mean Squared Error (RMSE) and cross-validation
- RMSE = $\sqrt{\frac{\sum_n (y_{true} - y_{predicted})^2}{n}}$
  - RMSE is the standard way to estimate the accuracy of a regression in chemometrics
  - Has the same units as the quantity being predicted
  - Squared error strongly penalizes the worst predictions
- Cross validation
  - Many regression models are prone to over-training: they can predict the composition of training samples to arbitrary accuracy, but then will not perform well on unknown data
    - A "simpler" model typically is more generalizable
  - Cross validation mitigates overtraining by iteratively removing samples from the training set and predicting them as if they are unknowns
  - If you have enough data, keeping an independent "test set" is the best way to predict against overtraining

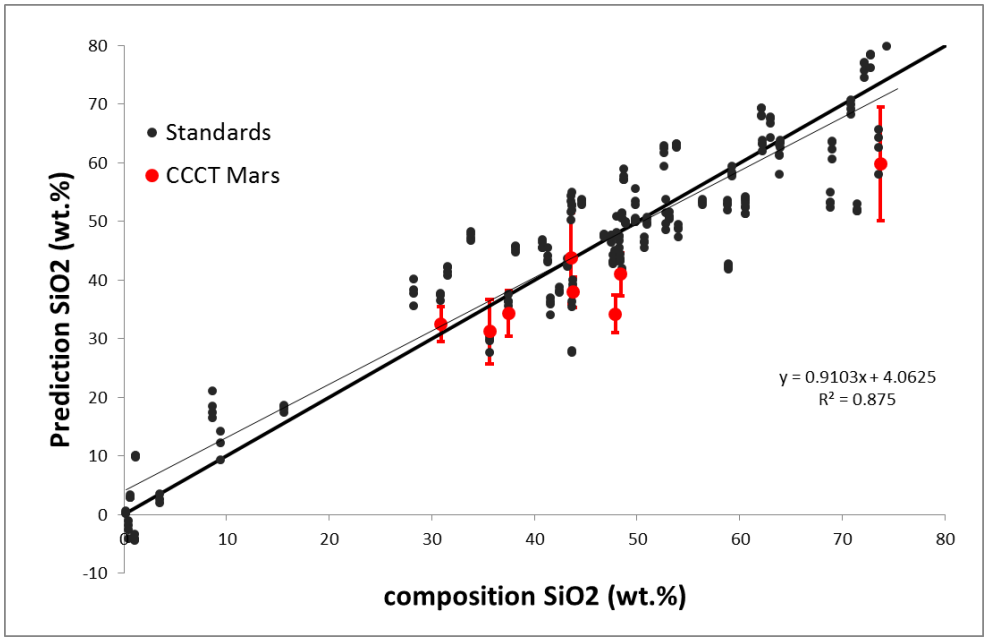
Univariate vs Multivariate Methods

- "Univariate" methods:
  - Regress a single variable (e.g. emission line strength) against composition to come up with a calibration curve
  - Advantages:
    - Simple, easy to understand
    - You can choose which emission line to use
  - Disadvantages:
    - Do not correct for "matrix effects" (factors that cause line strength to vary independent of elemental concentration)
    - Discards much of the information in the spectrum
- Multivariate methods:
  - Use the entire spectrum (or a significant portion of it) to develop a regression model
  - Advantages:
    - Can correct for some matrix effects
    - Uses most available information
  - Disadvantages:
    - Harder to understand what the model is doing
    - More computationally intensive

Current CCAM Quantitative Calibration

- 66 Geochemical Standards Calibration Database, Collected with the ChemCam Flight Model under Mars atmospheric conditions
- Partial Least Squares 1 (PLS1):
  - Multivariate method that regresses multiple X observations (spectra) against a single y value (composition)
  - Closely related to PCA (# of components determines complexity of the model)
- Generate independent optimized models for all major element oxide: SiO_2, TiO_2, Al_2O_3, FeOT, MgO, CaO, Na_2O, K_2O
  - Adjustable parameters:
    - Training spectra
    - # of components
    - Spectral Normalization
    - "Optimum" model defined as minimum leave-one-out cross validation RMSE, with some adjustments made to correct for specific cases





	SiO2	TiO2	Al2O3	FeOT	MgO	CaO	Na2O	K2O	Total
TRAINING SET MIN.	0.2	0	0	0	0	0.1	0	0	
TRAINING SET 1ST QUARTILE	40.8	0.27	5	2.7	0.8	2.5	0.3	0.3	
TRAINING SET MEDIAN	48.6	0.68	13.1	6	2.2	7.1	2.4	0.8	
TRAINING SET 3RD QUARTILE	59.3	1.47	16.1	12.1	6.4	12.8	3.4	1.8	
TRAINING SET MAX.	75.4	5.9	38.8	36.2	49.2	54.9	5.9	6.4	
NORMALIZATION	3	1	1	1	1	3	1	3	
NUMBER OF COMPONENTS	8	10	4	7	8	8	10	4	
RMSEP	7.1	0.55	3.7	4	3	3.3	0.7	0.9	10.1

Quantitative elemental calibration

- As is typical for any analytical instrument, ChemCam's precision (variation in results from repeat analyses of the same target) is considerably better than the absolute accuracy (RMSE)

Table 4. Precisions obtained on synthetic glass Norite and Shergottite rover calibration targets. Exact compositions are given for reference.

	n	SiO_2	TiO_2	$Al2O_3$	FeOT	MgO	CaO	Na_2O	K_2O
Norite[a]		47.9	0.70	14.7	15.9	9.62	12.8	1.53	0.06
Std dev sol 352	9	0.34	0.05	0.12	0.24	0.12	0.32	0.11	-----
Std dev sol 357	9	0.68	0.04	0.21	0.27	0.13	0.50	0.12	-----
Shergottite[a]		48.4	0.43	10.8	17.6	6.39	14.3	1.57	0.11
Std dev sol 271	7	0.60	0.03	0.18	0.26	0.14	0.37	0.10	0.04
Std dev sol 352	9	0.62	0.04	0.14	0.23	0.15	0.30	0.09	0.04
Std dev sol 357	9	0.37	0.02	0.07	0.12	0.07	0.35	0.11	0.04
Mean std dev	5	0.43	0.05	0.13	0.27	0.09	0.30	0.11	0.04
Std dev, all Shergottite obs.	25	1.53	0.14	0.57	1.83	0.49	0.42	0.49	0.14

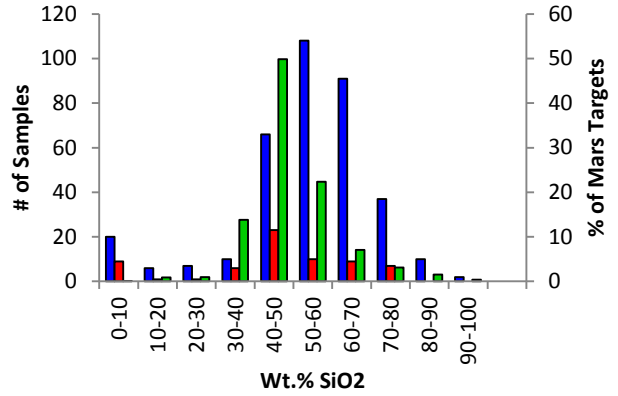
[a]Norite and Shergottite compositions are from Wiens et al. (2013).

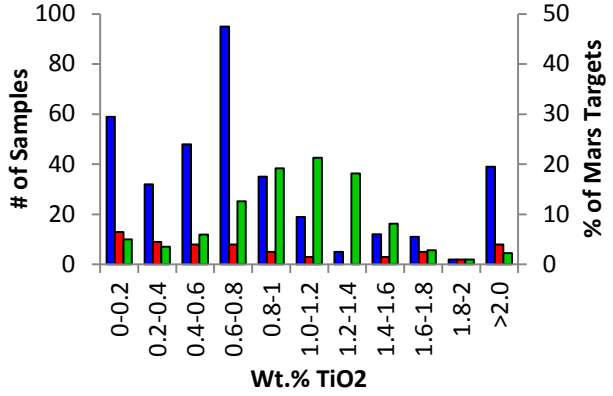
Blaney et al., 2014

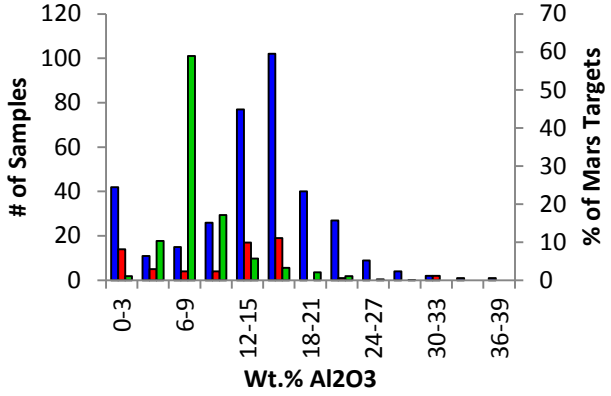
Improving CCAM's Quantitative Calibration

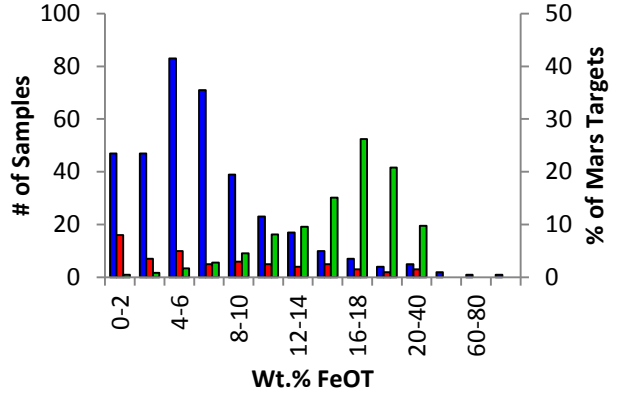
- How are we improving quantitative CCAM results?

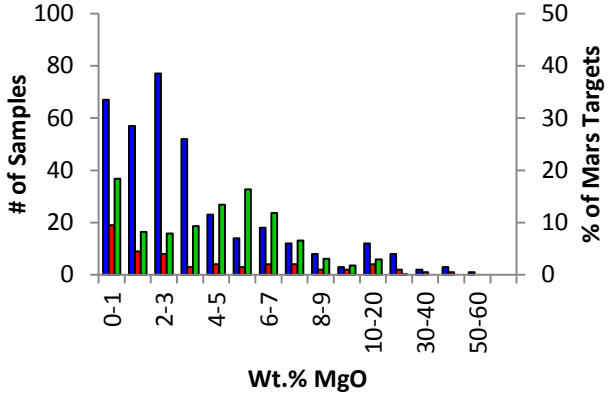
1. Expand the training database:
   - Current results are based on a database of 66 geostandards
   - We have collected data for an expanded database containing 482 standards, including many minerals
   - Larger database allows more robust cross validation (k-fold vs leave-one-out)
2. Improve calibration methods:
   - We are currently experimenting with several different methods:
   - "Sub-model" Partial Least Squares
     - Use a ratio-based "calibration transfer" method to make terrestrial database spectra more similar to spectra collected on Mars
     - Create PLS models that "specialize" in a certain composition range (e.g. a "low", "mid", and "high" SiO2 model)
   - Univariate
     - Use calibration target info from Mars to develop regressions
   - Independent Components Analysis (ICA) regression
     - Similar to univariate, but uses ICA score rather than individual line strength
   - Validate against Mars targets for which we can constrain composition

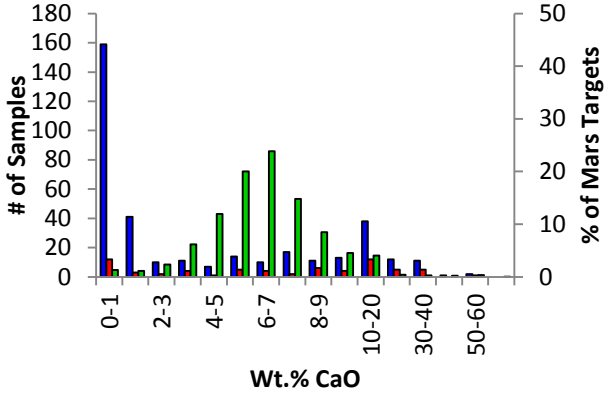


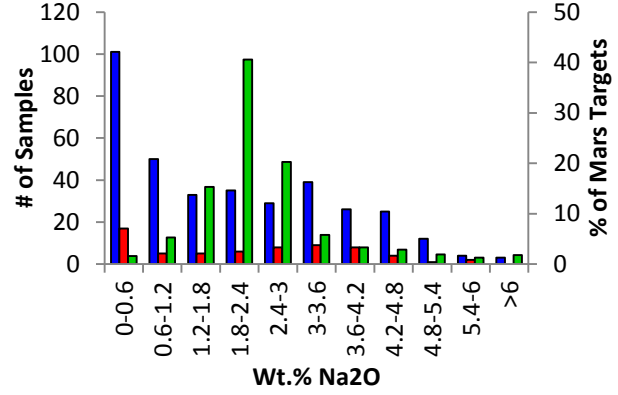


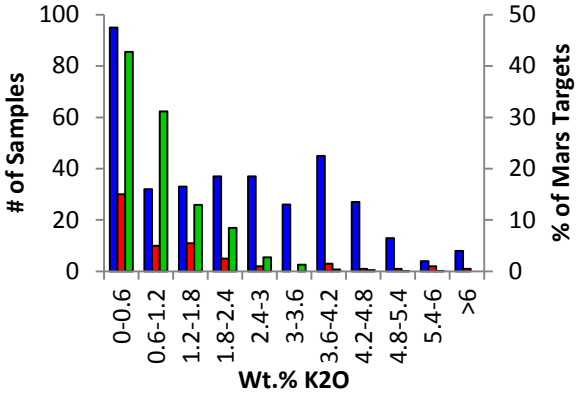








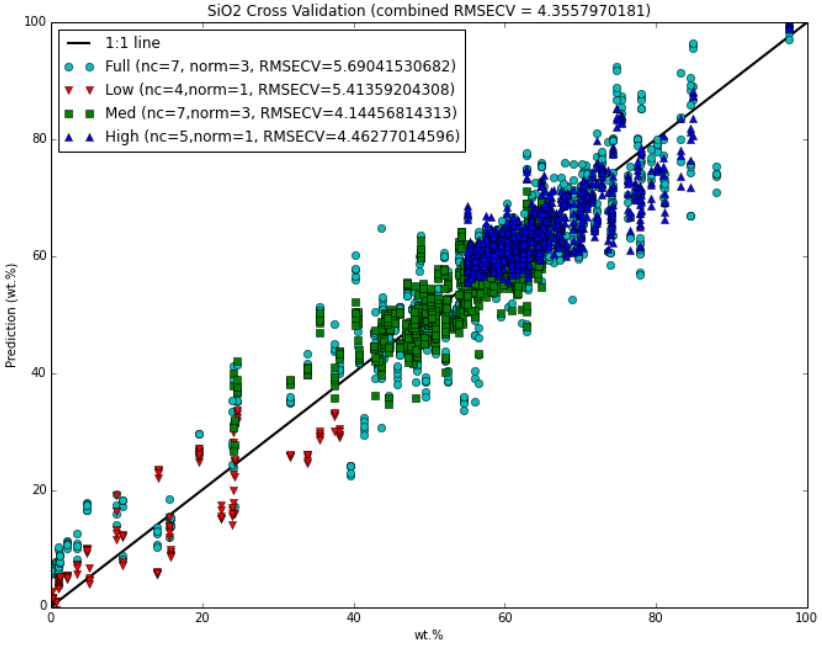


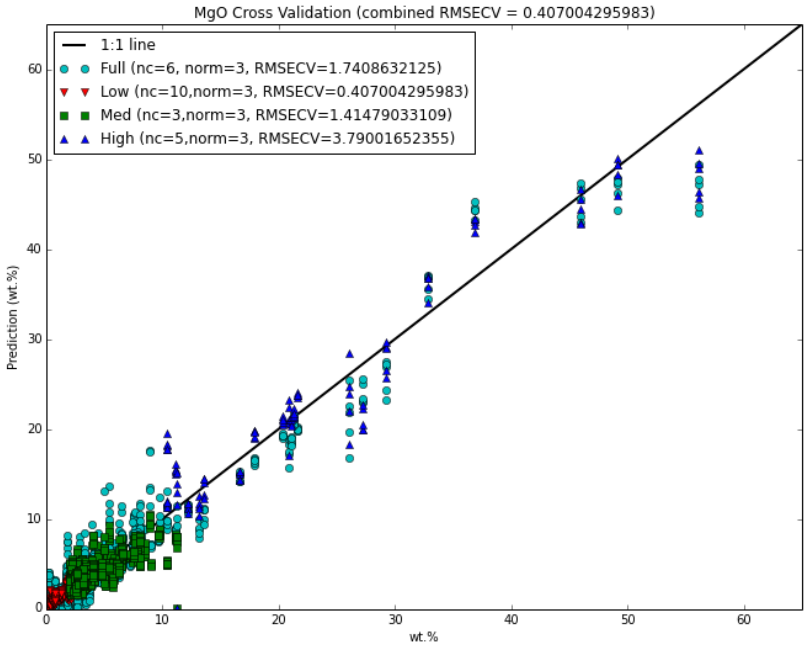


New Database

Old Database

Mars (based on current PLS1 calibration)





- ***Preliminary*** results from "submodel" PLS are promising:
  - "Old" SiO2 RMSE = 7.1 wt.%
  - "New" SiO2 RMSE (full model) = 5.7 wt.%
  - "New" SiO2 RMSE (using submodels) ~ 4.4 wt.%
  - "Old" MgO RMSE = 3 wt.%
  - "New" MgO RMSE (full model) = 1.8 wt.%
  - "New" MgO RMSE (using submodels) ~ 0.41 wt.%
- Samples with "bad" compositions in the "old" calibration (e.g. SiO2 >80%, MgO <0%) are significantly improved

Conclusions

- We are currently using PLS1 for quantitative results
- Accuracy is rigorously estimated using cross validation and RMSE
  - CCAM precision is better than accuracy
    - Changes in measured composition are more reliable than absolute compositions
- Lots of ongoing work to improve calibration!
  - Expanded database shows improved results
  - "submodel" PLS, ICA, Univariate methods being investigated and validated
- New calibration should be available by August 2015
- A calibration paper describing the improvements in detail is in the works. Stay tuned!
- Feel free to contact me with questions about quantitative calibration: rbanderson@usgs.gov

References

- Anderson, et al., (2015) “ChemCam Results from the Shaler Outcrop in Gale Crater, Mars” Icarus, Vol. 249, pp. 2–21
- Anderson, et al. (2013) “Spectral Classification and Variability in ChemCam Data from Bradbury Landing to Rocknest”, 44th LPSC, #2750.
- D. Anderson et al. (2015) “Quantification of Salt Anions Using Laser-Induced Breakdown Spectroscopy (LIBS)”, 46th LPSC, #2724
- Boucher et al. (2015) “Calibration Transfer of LIBS Spectra to Correct for Mars-Earth Lab Differences”, 46th LPSC, #2773
- Dyar et al. (2015) “Calibration Suite for Mars-Analog Laser-Induced Breakdown Spectroscopy”, 46th LPSC, #1510
- Fabre et al. (2014) “In situ calibration using univariate analyses based on the onboard ChemCam targets: first prediction of Martian rock and soil compositions” Spectrochim. Acta B, Vol. 99, pp. 34–51
- Forni, O., et al. (2015), First detection of fluorine on Mars: Implications for Gale Crater's geochemistry, Geophys. Res. Lett., 42, doi:10.1002/2014GL062742.
- Forni, O. et al. (2015) “Fluorine and Lithium at the Kimberly Outcrop, Gale Crater”, 46th LPSC, #1989
- Forni, et al. (2013) “Independent component analysis classification of laser induced breakdown spectroscopy spectra” Spectrochim. Acta B, Vol. 86, 31–41. doi:10.1016/j.sab.2013.05.003
- Gasnault, O. et al. (2015) “Updated Perspective on ChemCam Data Through Clustering”, 46th LPSC, #2789
- Lanza et al. (2015) “Oxidation of Manganese at Kimberly, Gale Crater: More Free Oxygen in Mars’ Past?”, 46th LPSC, #2893
- Lanza, et al. (2015) “Understanding the signature of rock coatings in laser-induced breakdown spectroscopy data” Icarus, Vol. 249 pp. 62–73
- Lanza, N. L., et al. (2014), High manganese concentrations in rocks at Gale crater, Mars, Geophys. Res. Lett., 41, 5755–5763, doi:10.1002/2014GL060329.
- Lasue et al. (2015) “Detection of Zn With ChemCam on Mars”, 46th LPSC, #1413
- Ollila, et al. (2013) “Trace Element Geochemistry (Li, Ba, Sr, and Rb) using Curiosity ’ s ChemCam: Early Results for Gale Crater from Bradbury Landing Site to Rocknest” Journal of Geophysical Research: Planets. doi:10.1002/2013JE004517
- Ollila, et al. (2014) “Trace Elements (Strontium, Barium, Rubidium, and Lithium) Analyses by ChemCam for the First 360 Sols in Gale Crater, Mars”, 45th LPSC, #2490
- Tokar et al. (2015) “Relationship Between MSL/ChemCam Laser Focus, Plasma Temperature, and Compositional Calibrations”, 46th LPSC, #1369