CHEMCAM DATA ACCESS, PROCESSING, AND INTERPRETATION R.B. Anderson[1], K. E. Herkenhoff[1], R.C. Wiens[2], S.M. Clegg[2], O. Forni[3], J. Lasue[3], A. Cousin[3], O. Gasnault[3], D. Delapp[2], N. Lanza[2], D. Blaney[4], [1]USGS Astrogeology Science Center, Flagstaff, AZ (rbanderson@usgs.gov), [2]Los Alamos National Laboratory, Los Alamos, NM, [3]Institut de Recherche en Astrophysique et Planétologie, Toulouse, France, [4]Jet Propulsion Laboratory

Introduction: The ChemCam instrument on the Curiosity rover uses Laser-Induced Breakdown Spectroscopy (LIBS) to analyze targets up to ~7 m from the rover, collecting atomic emission spectra (240-850 nm) from the laser induced plasma that contain diagnostic emission lines of all major and some minor and trace elements in the sample. The laser is co-boresighted with a Remote Micro-Imager (RMI) for high-resolution (19.6 μrad/pix) imaging. ChemCam has collected >210,000 LIBS spectra and >4200 images on Mars.

The martian environment is ideal for LIBS because the low atmospheric pressure (5-7 Torr) results in a large, bright plasma spark [1]. LIBS is most sensitive (5-100 ppm) to elements that are readily ionized (e.g. alkali and alkali earth metals), and least sensitive (0.1-3%) to nonmetals and halogens.

LIBS Data Processing: Raw ChemCam LIBS spectra must be pre-processed before geochemical analysis [2, 3]. For each active LIBS spectrum collected, an accompanying spectrum is collected without the laser. This "passive" or "dark" spectrum can be subtracted from the LIBS spectrum to remove the effects of ambient light and absorption lines in the solar spectrum. Noise and continuum removal are both accomplished using an undecimated wavelet transform to identify high- and low-frequency signals in the spectra [2, 3]. The continuum in ChemCam LIBS spectra is related to Bremsstrahlung and ion-electron recombination in the plasma and is distance dependent, so continuum removal also partially corrects for distance effects.

An instrument response function and geometric factors are used to convert the spectrum from counts to photons [2, 3]. Normalization to total observed intensity, either by spectrometer (i.e. the sum of the full spectrum equals 3) or across all three spectrometers (i.e. the sum of the full spectrum equals 1) provides an additional correction for distance effects [2, 3].

Wavelength calibration and resampling of each-spectrum is crucial, given the narrow width of atomic emission lines. Spectra from a Ti calibration target on the rover are used to provide a temperature-dependent, channel-by-channel wavelength calibration [2, 3].

Qualitative Data Analysis: ChemCam LIBS spectra contain 6144 spectral channels and hundreds of spectral lines. To aid in identification of emission lines in LIBS spectra (Fig. 1), the ChemCam Quick Element Search Tool (C-QuEST) is available at [3], and can be used to search both the NIST spectral database and a database specific to LIBS spectra collected under Mars-like atmospheric conditions.

Data reduction methods are useful to analyze large spectral data sets and visualize spectral similarity. These methods include Principal Component Analysis (PCA), which reduces high-dimensional data to a lower number of dimensions by identifying axes ("components") corresponding to directions of maximum variation in the n-dimensional data cloud, and Independent Component Analysis (ICA) which is similar to PCA, but seeks to identify components that are statistically independent. ICA has the advantage that each component tends to correspond to a single element, so that each ICA score can serve as a qualitative proxy for signal strength from the corresponding element.

ICA or PCA scores are often used as the input to classification algorithms. Many algorithms can be used, including unsupervised (e.g. hierarchical clustering, K-means clustering), and supervised methods (e.g. Soft Independent Modeling of Class Analogy (SIMCA), PLS Discriminant Analysis (PLS-DA)). [e.g. 3,4,5,6,7]

Quantitative Data Analysis: ChemCam LIBS spectra can be used to determine quantitative abundances of elements of interest. For minor and trace elements, the ChemCam team uses "univariate" calibration [3,8,9], while concentrations of major elements are calculated using "multivariate" methods [2,3].

Univariate Calibration: This method uses the strength of a single emission line to predict the composition of the corresponding element. This method typically uses peak fitting to isolate individual emission lines within fully processed "cleaned calibrated spectra" (CCS). Peak areas can then be plotted against the known composition of the eight geologic ChemCam calibration targets onboard the rover and a calibration curve can be determined. Ratios of peak areas can also be used and help mitigate differences in line intensity on different target types. Advantages of univariate calibration are its simplicity and its independence from terrestrial measurements (i.e., it is based entirely upon spectra collected by the flight instrument under martian conditions). However, univariate calibration cannot correct for "matrix effects": factors that can cause an element's emission line strength to vary independent of elemental concentration [2,3,6].

Multivariate Calibration: Multivariate methods make use of the entire spectral range or a significant portion of it, rather than an individual emission line, to develop a regression model relating the spectrum to a chemical composition. By making use of all available information in this manner, multivariate methods can partially correct for matrix effects [6]. The disad-

vantage of multivariate methods is that they are computationally intensive and it can be difficult to determine how the model arrived at a given result.

The ChemCam team currently uses the Partial Least Squares (PLS) method to derive major element compositions from target spectra [2,3]. PLS and most other multivariate regression methods use a "training set" of known spectra and corresponding compositions to predict the composition of an unknown target. To avoid overfitting the model to the training data, cross-validation is used to choose the number of components, and the estimated accuracy of the model is expressed as the root-mean-squared error (RMSE). All PLS-based quantitative ChemCam results available on the Planetary Data System (PDS) as Major Oxide Calculation (MOC) files list the estimated accuracy, along with the quartiles of the training set used for each element. Predictions that are near or outside the range of compositions in the training set are less reliable. The precision of ChemCam-derived compositions is better than the accuracy [10] (i.e., changes in measured composition are more reliable than absolute compositions).

Quantitative results available on the PDS are based on calibration that uses 66 geostandards [2,3]. Work is ongoing to develop an updated calibration based on an expanded database of 482 standards [3].

Image Data: For each LIBS observation, ChemCam also collects at least two RMI images: one before and one or more after LIBS, depending upon the number and geometry of analysis locations. These images provide context for the LIBS analyses and can be used to locate laser ablation pits and characterize the geology of the targets [11]. Mosaics of the RMI images (Fig. 2) associated with each LIBS observation, annotated with approximate LIBS analysis locations, are available on the ChemCam website [12] under the "results" tab. The RMI is also occasionally used to collect "standalone" images of targets independent of LIBS. Repeated RMI observations at different focus settings (Z-stacks) can be used to create focal merges and derive 3D information [11].

Data Access: To date, ChemCam data through Sol 804 are released on the PDS. These include active and passive spectral data (raw and processed), MOC files, RMI images (raw and processed, including standalone, Z-stack data, and mosaics), and LIBS spectra collected in the laboratory (used for calibration). ChemCam files on the PDS follow the naming convention in Fig. 3. ChemCam "quicklook" products are also available on the PDS, and the PDS provides the MSL Curiosity Analyst's Notebook, which provides a user-friendly way to access mission data [3,12].

The ChemCam team encourages scientists interested in working with the ChemCam data on the PDS to contact members of the team to assist in analyzing the data. A spreadsheet with contact information for the ChemCam science team is available at [3].

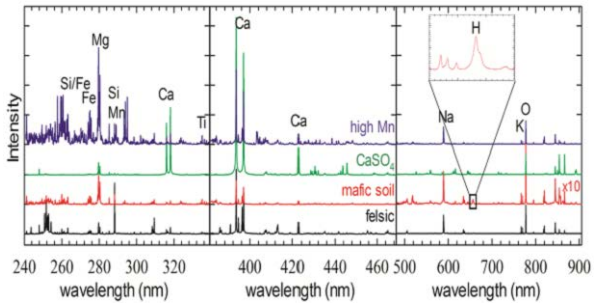

Fig. 1: Example processed ChemCam spectra of compositionally diverse targets on Mars.

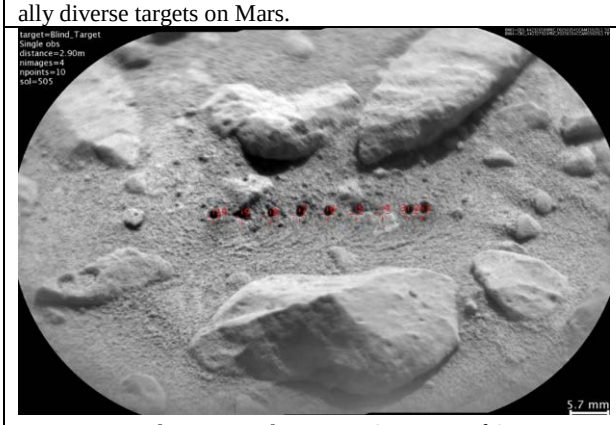

Fig. 2: Example annotated post-LIBS mosaic of 3 RMI images. Mosaics of this type are available at [11].

Fig. 3: ChemCam file naming convention. **1.** Data Type: CL5 = LIBS, CL9=Passive, CR0 = RMI, CL0 = Passive (averaged); **2.** Spacecraft clock; **3.** Processing level: EDR = raw, RDR = Level 1a, CCS = "Cleaned Calibrated Spectra" Level 1b, MOC = Level 2, PRC = processed RMI; **4.** Flight software version; **5.** Sequence ID; **6.** Processing version (always use the highest P# available); **7.** File type

References:[1] Knight, A.K., et al., 2000. Appl. Spectrosc. 54, 331–340. [2] Wiens, R.C., et al., 2013. Spectrochim. Acta B. 82, 1–27. [3] http://pds-geosciences.wustl.edu/workshops/ChemCam_Worksho p_Mar15.htm [4] Gasnault, O. et al. (2015) 46th LPSC, #2789 [5] Anderson, et al. (2013, 44th LPSC, #2750. [6] Clegg, S.M., et al. 2009. Spectrochim. Acta B. 64, 79–88. [7] Ollila A.M. et al., (2012) Applied Optics 51, B130-B142. [8] Fabre et al. (2014) Spectrochim. Acta B, Vol. 99, pp. 34–51. [9] Ollila, et al. (2013) JGR, 119, 255-285. [10] Blaney et al. (2014), JGR, 119, 2109-2131. [11] Le Mouélic, S. (2015) Icarus, 249, 93–107. [12] http://results.msl-chemcam.com [13] http://an.rsl.wustl.edu