AN IDL ROUTINE FOR PREPROCESSING AND ANALYSIS OF LIBS DATA. M. W. Schaefer[1], M. D.
Dyar[2], S. M. Clegg[3], and R. C. Wiens[3]. [1]Department of Geology and Geophysics, E235 Howe-Russell, Louisiana
State University, Baton Rouge, LA 70803, mws@lsu.edu, [2]Department of Astronomy, Mount Holyoke College, 50
College St., South Hadley, MA 01075, mdyar@mtholyoke.edu, [3]Los Alamos National Laboratory, P.O.Box 1663,
Los Alamos, NM 87545, sclegg@lanl.gov, rwiens@lanl.gov.

Introduction: The LIBS (Laser-Induced Breakdown Spectroscopy) instrument on ChemCam, selected for the Mars Science Laboratory (MSL) will quantitatively probe samples up to 9 m from the rover mast. The LIBS instrument focuses a Nd:YAG laser operating at 1067 nm onto the surface of the sample [1]. The laser ablates material from the surface and generates a plasma containing electronically excited ions, atoms, and small molecules. Light emission from this plasma is collected by three spectrometers for analysis: 240 – 335 nm (UV), 385 – 465 nm (VIS), and 510 – 800 nm (VNIR).

An active project is underway to collect and analyze spectra of geochemically diverse materials to study matrix effects and enable the best selection of flight calibration targets [2]. As part of this project, an IDL (Interactive Data Language [3]) routine for processing and analysis of the spectra is being developed. At present, software development is focused on processing and displaying laboratory LIBS data. Over the coming year, we expect to add advanced chemical analysis capabilities, [4,5].

Processing: The purpose of the IDL routine is threefold: first, to process the raw data for further analysis; second, to conveniently display the raw and processed data in a variety of ways; and third to perform multivariate analysis of the data and perform calculations needed to output chemical data. The first two of these aspects are currently implemented in the software package, as follows.

1. Open, and Average Data. In a typical laboratory experiment, the spectrometer exposure time is set to 1 second and records the plasma emission for 10 laser shots. The spectrometer software is set to average 5 of these exposures for each sample spot probed. Consequently, each probed spot represents 50 laser shots. Five different spots on each pressed pellet are typically sampled to identify any heterogeneity in these powdered samples. Data are collected in each of three wavelength ranges from the three spectrometers. The software reads and processes all of the spectra in the experiment; only one of these is chosen via a dialog box, and the others are automatically opened and read. Each spectrum is normalized to the total integrated intensity to compensate for experimental fluctuations such as shot-to-shot changes in laser power, laser-to-sample coupling, etc. Finally, the normalized

spectra are multiplied by a common scaling factor such that the normalized spectral intensities are similar to the originally recorded intensities. This scaling factor is the double sum of the intensites over both the 2048 channels and 5 spectra, divided by 5.

2. Background Subtraction. Background subtraction is handled via a subroutine based on the `fit_background` subroutine developed by Mark Rivers for the `MCA` analysis library [6]. An enhanced version of the algorithm published by Kajfosz and Kwiatek [7] is implemented.

3. Correction for Spectrometer Nonlinearity. All three spectrometers used in the instrument are slightly non-linear, such that the wavelength spacing between channels is not consistent. To correct for this problem, the software looks for a small subset of relatively-unoverlapped intense peaks (resulting from specific elements: either Ti from a calibration target, or geologically-ubiquitous elements such as O and Si) that occur in the wavelength range for each spectrometer. It fits these peaks to determine their location in channel numbers, and then regresses those values against the known wavelengths for those elements. The resultant curve is used to assign wavelengths (in nm) to each channel. Finally, the entire spectral range is interpolated to create a data set with the same number of channels as the initial file, but with the channels replaced by wavelength values with consistent spacings. The resultant data set can then be passed to the peak fit routine or reformatted and can be passed directly other to statistic routines that do not require information of specific peaks, such as the multivariate technique described in [4].

4. Peak Fits. Because some of the quantitative analysis techniques require peak energies, areas, intensities, and/or widths, the software calculates these. Peak characteristics are calculated via a subroutine based on the `fit_peaks` subroutine from the `MCA` analysis library [8]. This routine fits spectra to a set of Gaussian peaks and compares them to a look-up table containing peak positions for all geologically-reasonable elements. (We are also exploring additional lineshape functions such as Lorentzian and Voigt). Results are then output to an ascii file.

5. Quantiative analysis. As described in [4,5], several different methods for extracting quantitative
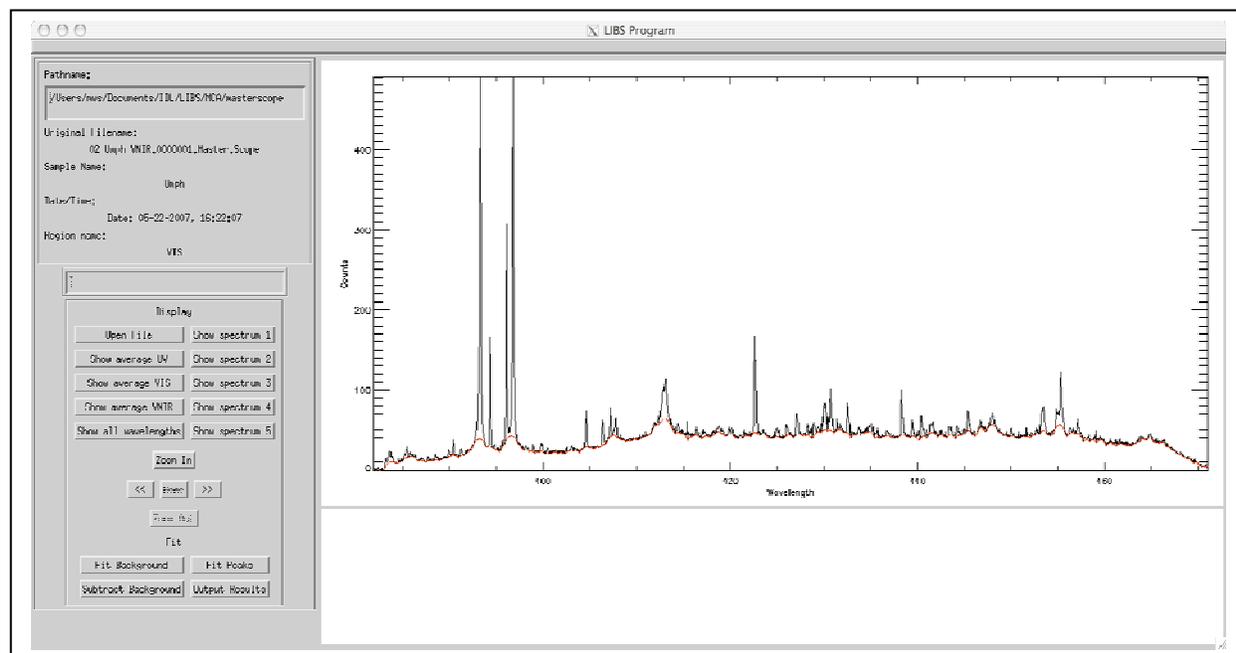
chemical data in terms of both atomic fraction and weight percent oxides are currently being evaluated. As these strategies are developed, we will implement them into the IDL code.

Display Capabilities. The IDL routine is capable of displaying the data in a variety of formats. Once data is read in, information about the data set is displayed on the left, and the spectrum itself plotted on the right. The graphical display may be zoomed in and out, and panned from side to side, to better inspect the data.

At any time each individual spectrum from the wavelength range read in may be displayed, or the normalized and averaged spectra from the three wavelength ranges may be displayed. The fit background button is accessible no matter which version of the spectrum is displayed, and one chooses within the fit background dialog box which wavelength range to fit. The fitted background is displayed overlaid on the spectrum in red. After fitting the background, it may be subtracted before fitting the peaks in a separate operation. When peaks are fit, they are displayed overlaid onto the spectrum in red, and also may be output as a ascii file.

Discussion and future modifications: At present, this software is designed to accommodate our laboratory data, but ultimately it could be easily adapted to processsing of remotely-acquired LIBS data from ChemCam. Additional factors that could be introduced in the flight software would include adjustments to correct for laser power and optical coupling to the target, distance, focus quality, and detector linearity and saturation.

References: [1] Wiens et al. (2005) *LPS XXXVI*, Abstract #1580. [2] Clegg et al. (2007) *Seventh Int. Conf. on Mars*, Abstract #3216. [3] www.ittvis.com [4] Clegg, S. et al., *Spectroch.*, submitted. [5] Dyar et al. (2008) *LPS XXXIX, Abstract # 2146*. [6] http://cars9.uchicago.edu/software/idl/mca_utility_ routines.html#FIT_BACKGROUND [7] Kajfosz, J., and Kwiatek, W. M. (1987) *Nucl. Instr. Methods* B22, 78-81. [8] http://cars9.uchicago.edu/software/idl/ mca_utility_routines.html#FIT_PEAKS.