Table 1. Summary of Strategies for Extracting Quantitative Chemical Data from LIBS Spectra (using Al as an example)

<table>
<thead>
<tr>
<th>Method</th>
<th>Mathematical Formulation</th>
</tr>
</thead>
</table>
| A. Traditional Spectroscopy | \[\text{Al atomic fraction} = b_0 + \beta X_{308.2} \] \[X = \text{the area of the peak at 308.2 Å}\]
| B1. Complete Multiple Regression | \[\text{Al atomic fraction} = b_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n \] \[X_1, X_2, \ldots, X_n = \text{the intensities of the spectra measured at each individual channel}\]
| B2. Selective Multiple Regression | \[\text{Al atomic fraction} = b_0 + \beta_{234.8} X_{234.8} + \beta_{244.5} X_{244.5} + \beta_{837.5} X_{837.5} + \beta_{923.7} X_{923.7} \] \[X_{234.8}, X_{244.5}, X_{837.5}, X_{923.7} = \text{the areas of all peaks at any wavelength fit above background (not just Al)}\]
| C1. Peak Areas or Intensities | \[\text{Al atomic fraction} = b_0 + \beta_{234.8} X_{234.8} + \beta_{244.5} X_{244.5} + \beta_{837.5} X_{837.5} + \beta_{923.7} X_{923.7} \] \[X_{234.8}, X_{244.5}, X_{837.5}, X_{923.7} = \text{the areas of all peaks at any wavelength fit above background (not just Al)}\]
| C2. Selective Peak Areas or Intensities | \[\text{Al atomic fraction} = b_0 + \beta_{234.8} X_{234.8} + \beta_{244.5} X_{244.5} + \beta_{837.5} X_{837.5} + \beta_{923.7} X_{923.7} \] \[X_{234.8}, X_{244.5}, X_{837.5}, X_{923.7} = \text{the areas of all peaks at any wavelength fit above background (not just Al)}\]
ables, and Gaussian lineshapes were tested and compared (this is the traditional analytical chemistry approach). These results were regressed against known elemental composition (expressed as atomic fraction; see [1]) to calculate conventional concentration curves based upon one or more emission line for a series of standard samples. Best results were obtained when the calibration standards were similar to the unknown, thus requiring a priori knowledge of the sample. Even when the unknown samples are similar to the calibration set, calibration data may be more scattered than desired for a highly accurate and precise quantitative analysis.

For method B, quantitative elemental compositions were extracted using partial least squares analysis. With PLS, a model is developed that describes a statistical relationship between a set of independent variables, $X$, and a set of dependent variables, $Y$. In the current application, the LIBS spectra are the $X$ variables and the elements are the $Y$ variables. PLS analyses employ a linear combination of values to correlate the spectral changes with the elemental compositions (Table 1). Our analysis was done using both the Unscrambler program [5] and SPSS (SPSS, Inc.). Results from this type of analysis are given in detail in [1]. This procedure could also be attempted using backward elimination to avoid the contribution from noise in channels not associated with any given element (Method B2); tests of such analyses are forthcoming.

For methods C1 and C2, we used the set of spectra from samples of various igneous rock types described in [1]. All statistically-significant peaks in each spectrum were fit using PeakFit software (Systat Software, Inc.). The resultant peak areas for all samples were parameterized into a format suitable for input into SPSS, and multiple regressions were run to search for correlations between peak areas and elemental concentrations. For each individual element, the regression analysis revealed which peaks are unrelated to the elemental abundance and which peaks are correlated (Method C1). We removed the unrelated peaks and reran the regression using only those peaks with significant regression coefficients (Method C2; Figure 1). Typically each element can be well-described by ~10-20 partial regression coefficients. Use of any backward-elimination approach requires that care be taken to test our models on other data sets for validation.

Both Methods B and C produce equations for each element that describe all the possible interactions that can occur within combinations of variables (i.e., chemical matrix effects arising from other elements) as well as the variables themselves (i.e., the lines resulting from emission of the element of interest). These methods identify not only channels/peaks directly resulting from emission of the element of interest, but also find channels/peaks resulting from other elements that may be affecting the intensities of the diagnostic channels/peaks. Methods B and C make no assumptions about which channels/peaks correspond to emission by any specific element. They simply fit all the channels/peaks that are present, and then attempt to correlate them with each compositional variable.

We are in the process of developing an IDL routine that will fit the all peaks in a given spectrum in batch mode [4]. The routine will also allow the energy scale of the spectra to be corrected for monochromator non-linearity as described in [4], permitting more accurate determination of peak areas and energies.

Finally, it is important to mention that as with any regression technique, the results from this analysis will be most accurate if the range of concentrations for each element and the number of calibration spectra are large. This underscores the importance of having a broad range of samples in the calibration suite. Further studies will be needed to determine if subsets of calibration suites (e.g., different sets of calibration equations for igneous vs. sedimentary rocks types) will be needed for interpretation of remotely-acquired data.

**Discussion:** It is likely that successful quantification of LIBS data will result from some combination of the approaches laid out here. Considerable additional work is needed to create and test calibration equations, to develop large, well-characterized sample suites for laboratory validation, and to inform choices of calibration standards on the rover.

**Acknowledgments:** This work was supported by NASA grant NNG06GH35G.