Contents lists available at ScienceDirect

## Spectrochimica Acta Part B

journal homepage: www.elsevier.com/locate/sab

# Effects of univariate and multivariate regression on the accuracy of hydrogen quantification with laser-induced breakdown spectroscopy

### Cai R. Ytsma<sup>a</sup>, M. Darby Dyar<sup>b,\*</sup>

<sup>a</sup> Department of Chemistry, Smith College, Northampton, MA 01063, United States

<sup>b</sup> Department of Astronomy, Mount Holyoke College, South Hadley, MA 01075, United States

#### ARTICLE INFO

Article history: Received 6 July 2017 Received in revised form 24 October 2017 Accepted 16 November 2017 Available online 20 November 2017

Keywords: Libs ChemCam Hydrogen Partial least-squares Lasso

#### ABSTRACT

Hydrogen (H) is a critical element to measure on the surface of Mars because its presence in mineral structures is indicative of past hydrous conditions. The Curiosity rover uses the laser-induced breakdown spectrometer (LIBS) on the ChemCam instrument to analyze rocks for their H emission signal at 656.6 nm, from which H can be quantified. Previous LIBS calibrations for H used small data sets measured on standards and/or manufactured mixtures of hydrous minerals and rocks and applied univariate regression to spectra normalized in a variety of ways. However, matrix effects common to LIBS make these calibrations of limited usefulness when applied to the broad range of compositions on the Martian surface. In this study, 198 naturally-occurring hydrous geological samples covering a broad range of bulk compositions with directly-measured H content are used to create more robust prediction models for measuring H in LIBS data acquired under Mars conditions. Both univariate and multivariate prediction models, including partial least square (PLS) and the least absolute shrinkage and selection operator (Lasso), are compared using several different methods for normalization of H peak intensities. Data from the ChemLIBS Mars-analog spectrometer at Mount Holyoke College are compared against spectra from the same samples acquired using a ChemCam-like instrument at Los Alamos National Laboratory and the ChemCam instrument on Mars. Results show that all current normalization and data preprocessing variations for quantifying H result in models with statistically indistinguishable prediction errors (accuracies) ca.  $\pm$  1.5 weight percent (wt%) H<sub>2</sub>O, limiting the applications of LIBS in these implementations for geological studies. This error is too large to allow distinctions among the most common hydrous phases (basalts, amphiboles, micas) to be made, though some clays (e.g., chlorites with  $\approx$  12 wt% H<sub>2</sub>O, smectites with 15–20 wt% H<sub>2</sub>O) and hydrated phases (e.g., gypsum with  $\approx$  20 wt% H<sub>2</sub>O) may be differentiated from lower-H phases within the known errors. Analyses of the H emission peak in Curiosity calibration targets and rock and soil targets on the Martian surface suggest that shot-to-shot variations of the ChemCam laser on Mars lead to variations in intensity that are comparable to those represented by the breadth of H standards tested in this study.

© 2017 Elsevier B.V. All rights reserved.

#### 1. Introduction

The aqueous history of Mars has been the subject of much study and debate [e.g., [1,2,3,4]]. Neutron, gamma-ray and near-infrared spectroscopies from orbit have conclusively shown that up to 10 weight percent (wt%) of water-equivalent hydrogen (WEH) is present at low to midlatitudes [5,6,7,8]. However, the distribution of that hydrogen (H), whether as subsurface water ice or structural H contained in surface glass or hydrous and hydrated minerals, remains the focus of multiple investigations. Among these, several instruments aboard the Mars

\* Corresponding author. E-mail address: mdyar@mtholyoke.edu (M.D. Dyar). Science Laboratory (MSL) rover *Curiosity* are exploring H contents of Gale Crater rocks. Litvak et al. [9] used the Dynamic Albedo of Neutron instrument [cf. [10,11,12,13]] to detect 1.5–2.5 wt% WEH in the Kimberly Formation. These values are comparable to concentrations measured by the Sample Analysis at Mars instrument [14], which used evolved gas analysis, gas chromatography–mass spectrometry, and tunable laser spectroscopy to find 1.2–2.5 wt% H<sub>2</sub>O at the John Klein, Cumberland, and Rocknest localities [15]. At the same locations, the CheMin instrument observed 18–22% clay minerals [16] and up to 20% smectite minerals in the Sheepbed mudstone [17] using x-ray diffraction; these indicate that  $\approx$  3–4% of the bulk rock is H<sub>2</sub>O. Finally, the ChemCam instrument, which uses laser-induced breakdown spectroscopy (LIBS), has the potential to measure H with a beam size of <0.5 mm to enable remote quantification of rock and soil compositions (see LIBS review by [18]).







To develop calibrations for elements of interest for the ChemCam LIBS, geologic standards with known compositions have been analyzed in the laboratory using parameters similar to those found on Mars [19, 20]. The ChemCam science team initially focused its efforts on quantification of major rock-forming elements including Si, Al, Ti, Fe, Mg, Na, Ca, and K [19–27]. More recently, the ChemCam team has turned its attention to the study of minor elements including P [28], Mn [29], salt anions [30], F and Li [31], Cr [32], Sr [33], S [34], Li and Ba [35] and H [36–40]. Of these, H is particularly important given its detection by the other instruments and because LIBS provides in situ analyses that can potentially relate variations in H contents to outcrop-scale phenomena such as veining or alteration rinds.

LIBS is ideally a useful method for detecting H because a strong emission line occurs in the region that LIBS measures: the  $\alpha$  Balmer line at 656.6 nm. However, emission line intensities for light elements may be particularly susceptible to variations in bulk chemistry in the plasma due to chemical matrix effects [41], and standards for hydrogen analysis in geological samples are few were quantified by variable methods. H analyses reported for most rock and mineral standards are given in terms of loss on ignition (LOI) results that lump all volatiles (e.g., H<sub>2</sub>O, CO<sub>2</sub> etc.) together, and thus are too imprecise to be used to calibrate H individually. H or H<sub>2</sub>O has been directly and specifically measured in only a few dozen existing commercially-available geostandards. Thus, previous studies have relied upon H analyses using synthesized glass characterized by Fourier-transform infrared spectroscopy (FTIR) and flash combustion, estimated H<sub>2</sub>O contents based on stoichiometry, and samples created by mixing stoichiometric hydrous phases with basalt that has thermogravimetric H measurements [36,39]. As a result, previous work has been constrained to small numbers of standards, many with identical matrices, with varying accuracy on their H<sub>2</sub>O determinations.

Fortuitously, laboratory collections at Mount Holyoke College (MHC) include hundreds of standards with structural/internal H or H<sub>2</sub>O characterized by the uranium extraction technique for previous studies [e.g., [42–44]]; these cover a range of minerals, rocks, and glasses. Errors on these determinations of wt% H<sub>2</sub>O are  $\leq$ 0.1 [43]. Six samples were analyzed using FTIR spectroscopy in the laboratory of Sheila Seaman at the University of Massachusetts, Amherst [45]. U.S. Geological Survey reference materials are also included in these collections, determined with varying methods and errors (see Table S1 for data sources). Finally, the 100 samples studied by Tucker et al. [46] were analyzed for hydrogen using the mass spectrometer in the laboratory of Matt Kohn at Idaho State University. There the NBS30 biotite standard has an H<sub>2</sub>O content reproducibility of  $\pm$ 0.05–0.1 wt%. From all of these, 198 standards were available in sufficient quantities for LIBS analyses.

This study thus creates calibrations for H<sub>2</sub>O from LIBS data collected under Mars conditions using the largest known suite of standards available. Two different LIBS instruments were used to analyze a subset of 16 samples, for which data were acquired both at Mount Holyoke College (ChemLIBS) and at Los Alamos National Laboratory (LANL) with a ChemCam-analog LIBS; results from these instruments are also compared against Mars data from ChemCam. Varying combinations of data preprocessing are tested, including an array of normalization techniques suggested by other workers to improve prediction accuracy for H analyses by LIBS [36,37,39]. These previous models for H<sub>2</sub>O have used sets of standards with fewer than 50 samples covering a restricted compositional range. Here, model performance is estimated using leave-one-out cross validation rootmean-square errors (LOO-RMSE-CV) to determine the accuracy of these calibrations when predicting H<sub>2</sub>O in other "unseen" samples, such as those on Mars. Data sets collected for this study are used to compare univariate calibrations based solely on hydrogen peak areas with multivariate analyses (partial least squares, or PLS, and the least absolute shrinkage and selection operator, or Lasso) based on larger portions of the spectrum.

#### 2. Background

There are hundreds of published papers using LIBS to study hydrogen in non-geological applications (review by [47,48]). For example, Oztoprak et al. [49] found that the H/C line ratio was a useful indicator of the presence of organic compounds such as uric acid. Wang et al. [50] studied variations of the atomic emission line intensities of C, H, O, and N in explosives. Gaddam et al. [51] and many others have studied H in coal. Mansour et al. [52] studied H/N ratios in combustion systems. Quantification in such systems is often quite successful because the matrix is generally the same material, and the element(s) of interest is/are the only variable(s) in the system.

However, due to the complexities of variable rock compositions, only a handful of papers have addressed the issue of measuring H in geological samples (see review by [53]). Meslin et al. [40] studied a "ubiquitous hydrogen signature" in Martian soils, and concluded that it results from the hydration of amorphous phases in the soil that were identified by the CheMin Instrument [16,17]. Schröder et al. [37] investigated the H signature in LIBS data from several different rock and soil samples along *Curiosity*'s traverse, observing systematic variations. Soils were found to have the highest H signatures, likely due to alteration in their fluvial lacustrine environment. Almost all samples studied had an observable H peak.

Because bulk composition is quite variable, these observations are affected by the spectral normalization schemes used to process LIBS data. For example, ChemCam uses the total spectral intensity in each of three wavelength ranges to normalize its data. A sample high in transition metals will have many emission lines in the ChemCam (ultraviolet (UV)-visible (VIS)-near-infrared (NIR)) wavelength range, while a more alkali or silica-rich spectrum would have fewer lines. The H emission line in either of those matrices would have very different intensities following normalization to total counts, even when H remains constant, creating an uncertainty with this type of normalization that must be accounted for. Thus many workers have experimented with alternative approaches to normalization to attempt to lower this uncertainty.

These normalization methods have included division of the H peak area by total spectrum intensity in the visible near-infrared (VNIR) range, continuum intensity [37], and the areas of three atomic emission peaks. Normalization to carbon and/or oxygen peaks were tested, based on the belief that they constant in concentrations in Mars's CO<sub>2</sub> atmosphere (and in the experimental chambers used in analog experiments). Neutral C 247.9 was simultaneously fit with Fe 248.3 and ionic C 657.8 was simultaneously fit with the H peak, while the neutral O 778 triplet was fit as a single peak [36,39]. Rapin et al. [36] found that the relative standard deviation between measurements of the H peak area was smallest when normalizing to 0 778 ( $\pm$ 1.7 wt%) or C 248 ( $\pm$ 1.2 wt%) peaks. Normalization to the C 657.8 ion was found to be less reliable than to the neutral emission lines due to its adjacency to the H peak and subsequent influence on small H signals. Based on these results, Rapin et al. [36] surmised that emission lines intensities are more dependent on pressure than on VNIR and continuum normalization, which instead heighten signal variation. Thomas et al. [39] used identical methods of normalization to Rapin et al. [36], resulting in similar findings: normalization in all cases except continuum decreased scatter of H peak area, with O 778 and C 248 having the highest linear regression R<sup>2</sup> values, finally recommending O 778 due to its success at low weight percent H. This paper tests these methods using a larger, more diverse set of H standards to determine the generalizability of their conclusions.

#### 3. Materials and methods

#### 3.1. Sample selection

From over 10,000 diverse rock and mineral standards at the MHC Mineral Spectroscopy Laboratory, 198 natural samples with known Download English Version:

https://daneshyari.com/en/article/7673921

Download Persian Version:

https://daneshyari.com/article/7673921

Daneshyari.com