

OVERVIEW OF A NEW LIBS DATABASE FOR THE LIBS-RAMAN SENSOR. A. Barlow¹, E. Sawyers², E. A. Lalla^{2,3}, M. Aznar⁴, J. Manrique⁵, G. Lopez-Reyes⁵, M. Veneranda⁵, J. Freemantle², M. G. Daly², M. Konstantinidis⁶ and E.A. Cloutis⁷. ¹University of Waterloo, Waterloo, ON, Canada; ²York University, Toronto ON, Canada; ³Canadensys Aerospace Corporation, Bolton, ON, Canada; ⁴Universidade de Cabo Verde, Praia, Cabo Verde; ⁵Unidad Asociada UVA-CSIC-CAB. Boecillo, Valladolid, Spain; ⁶Institute of Health Policy, Management and Evaluation, University of Toronto, ON, Canada; ⁷Department of Geography, University of Winnipeg, Winnipeg, MB, Canada.

Introduction: Several space missions from NASA and ESA have visited and will land on Mars in the search for life. The last mission to arrive was the Mars 2020 Perseverance rover on 18 February 2021. The next mission to Mars will be the ExoMars Rosalind Franklin rover, expected to touch down sometime after 2028. Both Mars2020 and ExoMars are equipped with spectroscopic systems: SuperCam and SHERLOC on Perseverance and Raman Laser Spectrometer (RLS) on Rosalind Franklin (RF) [1]. MarSCoDe (Mars Surface Composition Detector) is China's first instrument for Mars material analysis, which accompanied the Zhurong Mars rover landing on Utopia Planitia. These instruments will identify rocks, minerals, and potential organic biosignatures on the Martian surface. Many challenges are associated with the qualitative and quantitative analysis of resulting data from current missions on Mars.

In the case of ablation techniques, LIBS relies on the instrumental configuration such as laser pulse duration and intensity, the number of laser pulses, focusing, laser-detector synchronization, detector exposure time and the number of detections. Moreover, the environmental conditions under which the plasma is generated, such as atmospheric pressure and composition affect the emission spectra. To maximize the results to be received from different instruments, we can apply several methods such as preprocessing the signal based on on-board calibration targets signals (i.e., denoising, background removal, wavelength calibration, instrumental response correction). Also, to improve the elemental composition quantification of various targets, we need to have a large number of laboratory sample analyses with the corresponding dataset and database. Major elements compositions are usually predicted using multivariate data analysis tools. A database of LIBS emission lines under planetary mission conditions from laboratory measurements will facilitate the detection of minor elements [2].

In the framework of ongoing and future missions, it is important to develop new databases and datasets not only to increase the flight possibilities of future systems but also to improve the analysis for the results received from Mars. It has been shown that the richness and accuracy of the information provided by the LIRS instrument acquired during the past simulated mission

allowed us to conclude that working effectively in cooperation increases detection capabilities. These new datasets will improve the potential for combined analysis of the different in-situ instruments and possible sharing of data and tools to facilitate the interpretation of the results.

LIRS Instrument: The LIRS instrument is a multi-spectroscopic design featuring two laser sources to provide different excitations for LIBS using a 1064 nm pulsed laser and Raman/LIF through a 266 nm laser. LIRS generates plasma using an active Q-switched 1064 nm laser (Bright Solutions Wedge HF) with 293 kW peak power and 560 ps pulse width at a 10 kHz repetition rate [3]. The laser diameter, spot size, intensity and plasma generation for the minimum and maximum sensing values are depicted in Table 1.

Distance (in cm)	20	35	50
Diameter (μm)	21	33	47
N# of Shots	1000	1000	1000
Intensity (GW/cm^2)	40.9	26	8.3

Table 1. LIRS Measurements conditions for different sensing distances.

The threshold for plasma generation is around 1 GW/cm^2 , the lowest peak intensity encountered is eight times higher than the threshold mentioned above and increases at shorter distances to up 14 times [3]. The beam delivery system for the radiation from the laser to the sample is an optical system composed of four bouncing mirrors, a UV-coated beam expander and an off-axis parabolic mirror. The total optical path is ~ 2 m plus the variable sensing distance between 20 to 50 cm. LIRS possesses a motorized beam expander and a motorized tip/tilt control for alignment before being focused using an off-axis parabolic mirror to the target [3]. The backscattered radiation produced in the sample is collected by an f/2.3 telescope composed of a 10 cm primary mirror and a secondary motorized mirror able to adjust the focus of the collection optics for different sensing distances. Moreover, both the beam expander and the focusing stage are calibrated to any sensing distance (within 20–50 cm) for the IR and UV beams separately. The collected light is delivered through a multimode 50 μm core optic fibre to an Andor echelle spectrograph coupled to an ICCD system. The

spectrograph and camera allow for a wide spectral range (200–850 nm) and a high resolution (0.04–0.16 nm). Slits are not required because the fibre core size is 50 microns. The detector is an intensified nanosecond gated CCD (Andor iStar) coupled to the spectrometer and cooled to -25C. The signal detection and instrumentation synchronization (between laser and camera) are possible through a delay generator which triggers the spectrometer-camera system with the laser. The light entering the detector is directed onto an 18 mm photocathode, generating visible photoelectrons, which are amplified in a phosphor intensifier tube and reach 1024 × 1024 pixel CCD with 13 µm wide pixels. A NIST (National Institute of Standards and Technology) calibrated deuterium lamp is used for spectral intensity calibration of the spectrometer, allowing it to remove the effect of the edge filter and saw-tooth effect of the echelle spectrometer system. The wavelength calibration is done with a HgAr fibre-coupled lamp from 240 nm to 890 nm following the standard procedure of the manufacturer.

Database Description. The database collection used for the recalibration of the LIRS-LIBS mode was obtained from the certified standard Brammer Standard Company Inc, OREAS Standards, Exolith Labs, NIST, GSJ, USGS, Sigma Aldrich, CSA, NCS Standard and National Resources Canada. Note that the standards from Brammer Standard Company Inc and Oreas were certified and developed in-house by the companies.

Other samples were characterized by E. A Cloutis at the University of Winnipeg. Moreover, the samples included in this study have been characterized by several techniques for composition (e.g., XRF, wet chemistry), structure (e.g., XRD), and spectroscopic techniques such as Raman and UV-vis-NIR reflectance spectroscopy. Details of the techniques applied to sample characterization can be found in [4, 5], and analytical data at the University of Winnipeg's Centre for terrestrial and Planetary Exploration (C-TAPE) website: <http://www.uwinnipeg.ca/c-tape>.

The samples included in this study have been discussed in several previous studies that focus on specific samples [4, 5, 6, 7], as well as in support of analysis of data from planetary missions [8, 9].

Overview of the Calibration methods. The calibration methods used in this research are those from [10]. Moreover, Alix et al. (2021) [10] detail the efficacy of various machine learning models in analyzing the chemical attributes of (extra-)terrestrial geological samples. The data used was considering the concentrations of major elements found in the samples. Our previous study compared the performance of partial least squares (PLS), extreme gradient boost (XGB), various neural networks, and linear models and their

prediction accuracy with reference to the collected dataset from ChemCam. Ensemble prediction models were also created and tested to combine the best-performing models. Ultimately, it was indicated that the PLS and XGB models performed with the highest predictive accuracy and fastest runtime for the tested data set. Also, it was found that the predictive ensemble model was often the superior model among the other prediction methods. Based on the previous findings, it was recommended that the ensemble method be used specifically when analyzing spectroscopy data sets.

The spectroscopic data used for training and testing the various models was collected with the goal of providing a wide range of concentrations of the following elements: SiO₂, TiO₂, Al₂O₃, FeO, MgO, CaO, Na₂O, and K₂O.

The performance of our database included several steps. To first pre-process the data for each element in the dataset, Principal Component (PCA) was conducted to allow for faster model calibration. The dataset was then segmented into training, validation, and testing sets. Using the newly generated training set, each of the models was trained using fivefold cross-validation, and optimal hyperparameters for each model were obtained. Each model was then run using half of the testing set, with the purpose of determining the best-performing models to be used in the ensemble model. Using the remaining half of the testing dataset, the best three performing models and their predictions were then aggregated to be representative of the ensemble model's predictions.

Acknowledgments: We thank the Canadian Space Agency, the Natural Sciences and Engineering Research Council of Canada, the Ontario Centre of Excellence (OCE), the European Research Council and the Ministry of Economy and Competitiveness (MINECO).

References [1] Aznar et al. (2023) *Adv. Space Res* (in press). [2] Lalla et al. (2021) *Appl. Spectrosc.* 75(9):1093-1113. [3] Konstantinidis et al. (2021) *Icarus* 358, 114113. [4] Cloutis et al (2006) *Icarus* 184, 121-157. [5] Cloutis et al. (2015) *Planet. Space Sci.*, 119, 155-172. [6] Cloutis et al. (2002) *J. Geophys. Res.* 107, 5-1,5-19. [7] Cloutis et al. (2021) *Planet. Space Sci* 208, 105336. [8] Bell et al. (2022) *Sci. Adv.* 8, eabo4856 (18). [9] Liu et al. (2022) *Sci.* 377, 6614, 1513-1519. [10] Alix et al. (2022) *J. Chemom.*, 2022; e3400.