WHAT LURKS IN THE MARTIAN ROCKS AND SOIL? INVESTIGATIONS OF SULFATES, PHOSPHATES, AND PERCHLORATES

Multivariate analysis of Raman spectra for the identification of sulfates: Implications for ExoMars

GUILLERMO LOPEZ-REYES1-*, PABLO SOBRON234, CATHERINE LEFEBVRE2 AND FERNANDO RULL1

1Unidad Asociada UVA-CSIC-Centro de Astrobiología, C/ Francisco Valles 8, 47151, Bocelli, Spain
2Space Science and Technology, Canadian Space Agency, 6767 Rte. de l’Aéroport, St. Hubert, Quebec J3Y 8Y9, Canada
3MalUva Labs, 822 Allen Avenue A, St. Louis, Missouri 63104, U.S.A.
4SETI Institute, 189 Bernardo Avenue 100, Mountain View, California 94043, U.S.A.

ABSTRACT

We have built three multivariate analysis mathematical models based on principal component analysis (PCA), partial least squares (PLS), and artificial neural networks (ANNs) to detect sulfate minerals in geological samples from laser Raman spectral data. We have critically assessed the potential of the models to automatically detect and quantify the abundance of selected Ca-, Fe-, Na-, and Mg-sulfates in binary mixtures. Samples were analyzed using a laboratory version of the Raman laser spectrometer (RLS) instrument onboard the European Space Agency 2018 ExoMars mission. Our results show that PCA and PLS, can be used to quantify to some extent the abundance of mineral phases. PCA separated hydrated from dehydrated mixtures and classified mixtures depending on the phase abundances. PLS provided relatively good calibration curves for these mixtures. Upon spectral pre-processing, ANNs provided the most precise qualitative and quantitative results. The detection of mineral phases was 100% accurate for pure samples, as was for binary mixtures where the abundance of mineral phases was >10%. The outputs of the ANN were proportional to the phase abundance of the mixture, thus demonstrating the ability of ANNs to quantify the abundance of different phases without the need for calibration. Taken together, our findings demonstrate that multivariate analysis provides critical qualitative and quantitative information about the studied sulfate minerals.

Keywords: Sulfates, ExoMars, Raman spectroscopy, multivariate analysis, qualitative, quantitative

INTRODUCTION

Laser Raman spectroscopy has been proposed as a powerful tool for the identification of minerals in the context of planetary exploration, including Mars (Sharma et al. 2003; Sobron and Alpers 2013; Sobron et al. 2008; Wang et al. 2003; Wiens et al. 2007), Europa (Angel et al. 2012; Sobron et al. 2013), Venus (Lambert et al. 2010), the Moon (Ling et al. 2009), and asteroids (Kong and Wang 2010). In addition, the feasibility of using laser Raman spectroscopy for the detection of biosignatures in terrestrial analogues to Mars has been demonstrated (e.g., Bower et al. 2013; Dickensheets et al. 2000; Edwards et al. 2003, 2011, 2012; Ellery and Wynn-Williams 2003; Steele et al. 2010; Wynn-Williams and Edwards 2000). A Raman laser spectrometer (RLS) is part of the science instrument payload of the European Space Agency 2018 ExoMars mission; the RLS instrument will target mineralogical and astrobiological investigations on the surface and subsurface of Mars (Rull et al. 2011a, 2011b).

The current concept of operation of the RLS instrument is a raster analysis of crushed drill–core materials (Rull et al. 2011a). In this configuration, the geological and morphological context of the spots analyzed by RLS will be lost, as the crushing stage will preclude correlation between RLS spectra and the imagery acquired by the rover Close-Up Imager CLUPI (Jossot et al. 2012). While the synergy between these instruments in the current ExoMars payload configuration has been demonstrated (Lopez-Reyes et al. 2013a), identification of the mineral phases present in the geological targets and quantification of their abundance with RLS will mostly rely on spectral data and not morphology or texture. Therefore, to enable unambiguous identification and quantification of phase abundance, robust spectral processing methods are needed.

Mineral identification using Raman spectroscopy is often performed by comparing acquired spectra to reference spectra available from the literature and different databases, e.g., the RRUFF project database (Downs 2006). Several algorithms have been developed that enable an automated identification of Raman spectra using traditional univariate analysis, i.e., the description of individual variables in a given spectrum (Hermosilla et al. 2013; Kriesten et al. 2008; Perez-Pueyo et al. 2004; Sobron et al. 2008). These algorithms, however, fail to accurately estimate mineral abundance in complex geological samples, although applications for the quantitative analysis of relatively simple mixtures have been proposed (Lopez-Reyes et al. 2013a; Schumacher et al. 2011; Vagenas and Kontoyannis 2003).

Multivariate analysis techniques (MVAT) are statistical techniques that deal with simultaneous measurements on many variables, and aim at understanding the relationships between

---

* E-mail: guillermo.lopez.reyes@cab.inta-csic.es
† Special collection papers can be found on GSW at http://ammin.
geoscienceworld.org/site/misc/specialissue.html.