





Machine learning methods applied to combined Raman and LIBS spectra: Implications for mineral discrimination in planetary missions

Sofía Julve-Gonzalez¹  | Jose A. Manrique^{1,2}  | Marco Veneranda¹  |
 Iván Reyes-Rodríguez¹ | Elena Pascual-Sanchez¹ | Aurelio Sanz-Arranz¹ |
 Menelaos Konstantinidis³ | Emmanuel A. Lalla^{3,4} | María E. Charro¹ |
 Eduardo Rodriguez-Gutierrez¹ | José M. Lopez-Rodríguez¹ |
 José F. Sanz-Requena¹ | Jaime Delgado-Iglesias¹ | Manuel A. Gonzalez¹ |
 Fernando Rull¹ | Guillermo Lopez-Reyes¹ 

¹ERICA Research Group, Universidad de Valladolid (UVa), Valladolid, Spain

²Université de Toulouse 3 Paul Sabatier, CNRS, CNES, Toulouse, France

³Centre for Research in Earth and Space Science, York University, Toronto, Ontario, Canada

⁴Canandensys Aerospace Corporation, Bolton, Ontario, Canada

Correspondence

Sofía Julve-Gonzalez, ERICA research group, Universidad de Valladolid (UVa), Valladolid, Spain.

Email: sofia.julve@estudiantes.uva.es

Funding information

Agencia Estatal de Investigación, Grant/Award Number: PID2022-142490OB-C32; Ministry of Economy and Competitiveness, Grant/Award Number: RDE2018-102600-T; European Union

Abstract

The combined analysis of geological targets by complementary spectroscopic techniques could enhance the characterization of the mineral phases found on Mars. This is indeed the case with the SuperCam instrument onboard the Perseverance rover. In this framework, the present study seeks to evaluate and compare multiple machine learning techniques for the characterization of carbonate minerals based on Raman-LIBS (Laser-Induced Breakdown Spectroscopy) spectroscopic data. To do so, a Ca-Mg prediction curve was created by mixing hydromagnesite and calcite at different concentration ratios. After their characterization by Raman and LIBS spectroscopy, different multivariable machine learning (Gaussian process regression, support vector machines, ensembles of trees, and artificial neural networks) were used to predict the concentration ratio of each sample from their respective datasets. The results obtained by separately analyzing Raman and LIBS data were then compared to those obtained by combining them. By comparing their performance, this work demonstrates that mineral discrimination based on Gaussian and ensemble methods optimized the combine of Raman-LIBS dataset outperformed those ensured by Raman and LIBS data alone. This demonstrated that the fusion of data combination and machine learning is a promising approach to optimize the analysis of spectroscopic data returned from Mars.

KEYWORDS

data combination, LIBS, machine learning, PCA, Raman

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2023 The Authors. *Journal of Raman Spectroscopy* published by John Wiley & Sons Ltd.

1 | INTRODUCTION

In recent years, several spectroscopic methods have been used to characterize the geology of Mars, including Raman spectroscopy and Laser-Induced Breakdown Spectroscopy (LIBS).^{1–5} On the one hand, Raman spectroscopy provides molecular information about the crystallographic structure of the target, thus allowing the discrimination of the mineral phases that compose it. On the other hand, LIBS provides information about its elemental composition thanks to its characteristic emission spectrum. Taking advantage of their complementarity, the SuperCam instrument onboard the NASA/Mars 2020 Perseverance rover is the first instrument operating in space that is capable of performing simultaneous Raman and LIBS analysis on the same target.⁶ SuperCam emerges as an evolution of ChemCam from the Curiosity rover.¹ The instrument combines five different techniques in the same instrument: Time-Resolved Raman and Luminescence (TRR/L), LIBS, Visible-Infrared Spectroscopy (VISIR), Remote Micro-Imaging (RMI), and sound recording.^{7–10}

Since its landing on Mars in February 2021, the analysis performed by SuperCam contributes to the fulfillment of the main scientific objectives of the Mars2020 mission by searching for biosignatures on the rocky surface of Mars.¹¹ To this end, it aims at characterizing the geological units found at the landing site and searching for molecular biosignatures that are potentially preserved within the mineralogical matrix.

In this regard, it must be underlined that the selected landing site (Jezero Crater) hosts one of the largest geological units of carbonates identified on Mars.^{12–14} Knowing that Jezero was formerly an open basin paleolake,¹⁵ it could be inferred that the carbonates detected at Jezero are alteration products generated from the long-term interaction of primary rocks with liquid water. Assuming the Jezero's paleolake could have harbored life in the past, carbonate deposits would represent one of the priority scientific targets were to search for organic biosignatures. As detailed in the final report of the Mars Biosignature Working Group,¹⁶ carbonates possess (together with sulfates, silica, phyllosilicates, and hematite), the key mineralogical features that are relevant for the preservation of biosignature at the geological time scale.¹⁷ Indeed, through a complex combination of organic–inorganic interactions,¹⁸ the mentioned mineral phases are able to shelter organic molecules from the alteration processes triggered by the arch outer environment, as has been already detected in Jezero crater, with associations between organic molecules and sulfates.¹⁹ This hypothesis is supported by a broad variety of studies in which the organic remains of ancient forms of life

have been successfully detected within relevant terrestrial analog materials,²⁰ as is the case of stromatolitic carbonates.^{21–23}

Since it landed on Mars, the SuperCam instrument successfully detected carbonates at different locations of the Jezero Crater.¹⁷ However, the limitations in the acquisition of Raman spectra, or the atmospheric contribution on the emission lines of C detected by LIBS and the mixture with other mineral phases on the other, makes the correct discrimination of the mineral phases detected by SuperCam challenging. To overcome this issue, the SuperCam team is working on the combination of complementary spectroscopic data collected from the same spot of interest as a way to ensure a reliable discrimination of carbonate phases on Mars.^{24,25} Learning from terrestrial applications, the present work aims at evaluating the advantages provided by multiple machine learning methods (Gaussian process regression [GPR],²⁶ support vector machines [SVMs],²⁷ ensemble of trees,²⁸ and artificial neural networks [ANNs]^{29,30}) in the proper discrimination of carbonate phases detected on Mars. After separately analyzing Raman and LIBS datasets, the chemometric models were also applied to combined Raman-LIBS data, to evaluate the advantages provided by a collaborative science approach between both techniques.

Additionally, it was found interesting to use this experiment not only to evaluate the different performances of the mentioned machine learning algorithms but also to evaluate the different performances that could come from the use of different variables to feed those models. Different works from our research group have used different approaches for the dimensional reduction of data that could be grouped in two: principal component analysis (PCA³¹) or extraction of spectral parameters. In the first case, we use the training data set to calculate the set of Principal Components that describe certain level of the variance of the whole system. This methodology includes a higher description of the input spectra but is also sensitive to not useful information, as is an unsupervised method. On the other hand, selection of spectral parameters as the intensity of certain emission lines in LIBS, or the fitting and extraction of parameters for the fitted bands in Raman spectra,³² can be methods to reduce dimensionality but more focused in relevant data fixed by operator (is a supervised method), although at a cost of losing information that operator deems not useful.

Attending exclusively to the combined use of data sets, it is important to evaluate differences between both techniques. On one hand, LIBS spectra covering a typical range of 600 nm at a resolution of 0.1 nm give us a total of 6,000 variables, while this would be significantly lower

for the case of a typical Raman instrument. Previous examples of data combination based on concatenation of spectra prior to dimensional reduction assigned inherently a higher weight to LIBS data, as there were more variables available from LIBS, and some of them presented a higher variance associated. Using similar methods for data combination could lead to a bias from origin towards the performance and capabilities of LIBS, while losing to some point the complementarity of both techniques. In this work we used an independent dimensional reduction per technique, and the combination is done afterwards by creating a combined data set that is later reduced in its number of variables to adequate its size to the dimensionality of the individual data sets. This selection of the final variables is aided by feature selection algorithms as Maximum Representativity minimum Redundancy (MRMR)³³ that provides scores to every input variable in terms of how well or how confusing is related to an objective target.

2 | EXPERIMENTAL METHOD: MATERIALS AND EQUIPMENT

2.1 | Carbonate samples

For this work, a set of binary carbonate mixtures was prepared by using commercial carbonates from sigma Aldrich (reagent grade, >99%) of calcite (CaCO_3) and hydromagnesite ($\text{Mg}_5 [\text{OH} (\text{CO}_3)_2] \cdot 4\text{H}_2\text{O}$) powdered standards as Ca and Mg endmembers. As for the Mg endmember, hydromagnesite was preferred over conventional magnesite due to its higher relevance for current astrobiology research on Mars. Indeed, through the analysis of CRISM data, recently identified several locations at the Mars 2020's landing site exhibit spectroscopic features that are compatible with the presence of hydromagnesite.¹² As multiple studies found the precipitation of hydromagnesite in aquatic environments to be often biologically induced,^{34–36} this mineral phase represents a primary geological target for astrobiology research at the Jezero Crater.

The calibration set was prepared by mixing both powdered minerals in the proportions shown in Table 1. The samples were then milled using an agate mortar to achieve homogeneity and reduce their grain size. Each mixture was finally compacted into a 1.5 cm diameter pellet (with a few mm thick) by using a 10-ton manual press.

2.2 | Analytical instruments

Raman analyses were performed by using a custom-built spectrometer developed by the ERICA research group

TABLE 1 Weight ratio of calcium carbonate and hydromagnesite in each binary mixture tablet.

Calcite (wt%)	Hydromagnesite (wt%)
100	0
99	1
95	5
90	10
75	25
62.5	37.5
50	50
37.5	62.5
25	75
10	90
5	95
1	99
0	100

(University of Valladolid). The instrument consists of the following commercial components: a 532 nm BWN-OEM excitation laser with 50 mW output power, a dual-track grating transmission spectrometer with an Andor Newton detector cooled at -50°C , a Raman head (BWTEK BAC100-532E) and a spectral range of 100 to 4000 cm^{-1} . Using this setup, 20 Raman spectra were acquired per sample at different points with an acquisition time adjusted to 75% of the total dynamic range of the sensor and 10 accumulations per point, thereby achieving a total of 260 Raman spectra. Data were acquired by means of the Andor Solis control software.

Concerning elemental analyses, LIBS data were collected using a custom-built setup equipped with a Nd:YAG (second harmonic) pulsed laser with a maximum pulse energy of 120 mJ and maximum operating frequency of 30 Hz, although 2 Hz was used for the data acquisition. The laser was focused on the sample using a convergent lens of 55 mm diameter and 75 mm focal. The light was collected at a distance of a few centimeters using a catadioptric condenser from Andor, and it was delivered to an echelle spectrometer attached to a $1,024 \times 1,024$ pixels intensified CCD. The instrument covers a total range from 200 to 950 nm, although efficiency in the UV and IR sides of this range is severely impaired. In the time resolving side, the detector includes a Digital Delay Generator and is externally triggered by the Q-switch of the pulsed laser. LIBS spectra were acquired using 80% of the maximum laser power (96 mJ), setting a delay time between the laser pulse and the acquisition of $2\ \mu\text{s}$ and gate width (this is the time window of light acquisition through the intensifier) of $10\ \mu\text{s}$.

After discarding the first two shots, 20 accumulations per spot were taken, with 20 spots per sample in the case of pellets with extreme concentrations between both compounds (100–0, 99–1, 95–5, 5–95, 1–99, and 0–100) and 40 spots for mixtures with intermediate concentrations of both compounds (90–10, 75–25, 62.5–37.5, 50–50, 37.5–62.5, 25–75, and 10–90).

3 | DATA ANALYSIS

3.1 | Selection of spectral parameters

The first step of data analysis consisted of the selection of key spectral indicators for both Raman and LIBS datasets. To do so, the characteristic LIBS and Raman spectra collected from the sample mixtures described in Section 2.1 were first pre-processed (baseline correction and intensity normalization) using the SpectPro software.^{37,38} Afterwards, the same software was used to evaluate the correlation between the dependence of characteristic spectral features (peak position, intensity, etc.) with the concentration ratio between calcite and hydromagnesite in the sample mixture. In order to optimize the results in the subsequent chemometric model training, parameter selection was carried out through two different approaches. The first step was a data treatment, as described in the introduction, based on PCA.³⁹ Reducing the dimensionality of the system to the first six PCs, we described the system with fewer variables while explaining 99.9% of the variance for Raman data and 98.7% for the LIBS data. The selected PCs for each data set were the selected variables for the different individual models trained.

The second method for variable reduction involved preprocessing of the spectral characteristics. Concerning LIBS data, the emission lines that did not present self-absorption, did not show overlaps, and show good correlations ratio between their normalized intensities and concentrations were chosen as key spectral parameters. In total, 10 characteristic emission lines were selected, 5 of the Ca and 5 of the Mg, as can be seen in Figure 1.

As represented in Figure 2, the characteristic Raman spectra of calcite present the main peak at $1,086\text{ cm}^{-1}$ and secondary peaks at 150, 282, and 712 cm^{-1} . Similarly, the spectra of hydromagnesite present peaks at 230, 327, 725, and $1,118\text{ cm}^{-1}$. Given the similarity between the Raman spectra of the two mineral phases, the reduction of Raman parameters was carried out using an appropriate spectral feature selection algorithm. In total, 10 spectral characteristics (intensity, position, width at mid-height, Lorentzian/Gaussian, and area) of the main peak of calcite and hydromagnesite were selected. Once the LIBS and Raman spectral parameters were selected, these characteristics were sequentially classified using the Minimum Redundancy Maximum Relevance (MRMR) algorithm to identify the six parameters that best described the spectra for each technique.

3.2 | Machine learning methods for data analysis

The identification and classification of spectroscopic data to achieve chemometrics for planetary missions have been one very active topic of research for years.^{40–43} Different approaches using multivariate data analysis and machine learning algorithms have demonstrated good

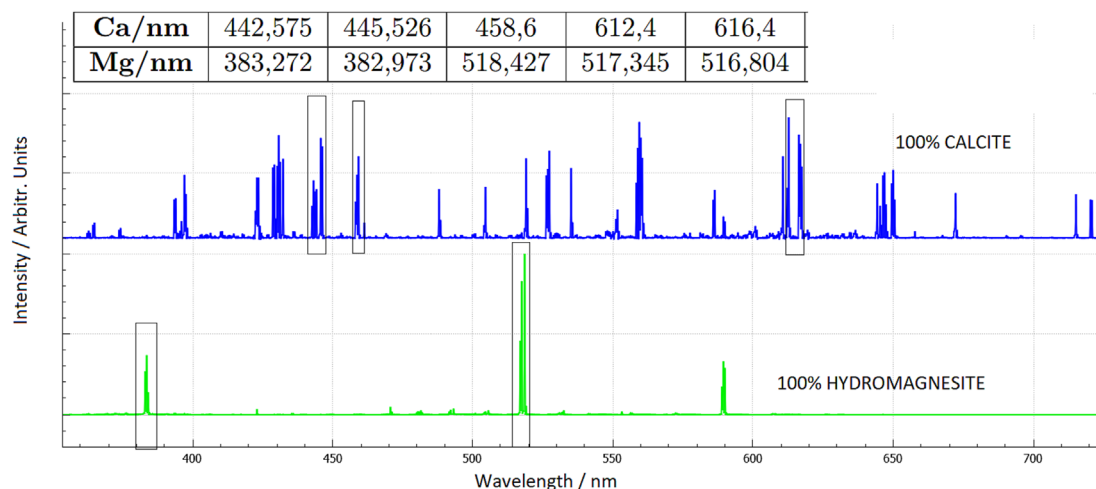


FIGURE 1 LIBS spectra characteristic of Ca and Mg where the characteristic spectral emissions of these two elements are marked and have been used as inputs to the analytical models developed.

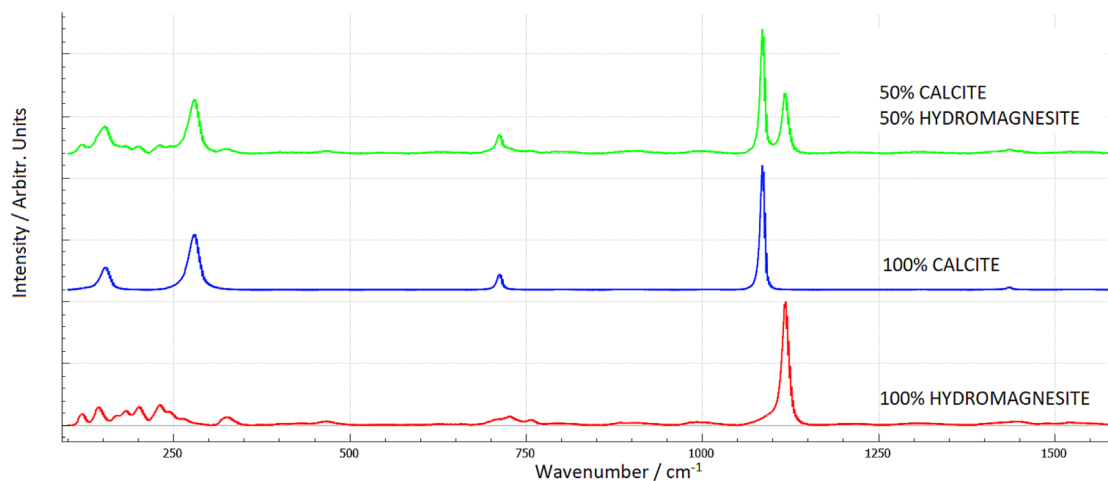


FIGURE 2 Raman spectra of binary mixtures with 100% calcium carbonate, 100% hydromagnesite, and tablet spectra with 50% CaCO₃ and 50% Mg₅ [OH (CO₃)₂] · 4H₂O.

performance extracting quantitative information from spectroscopic sources.^{31,32,44–48} In this work, once Raman and LIBS variables were selected, we evaluated different models generated using a set of machine learning algorithms for regression, then evaluating and comparing the goodness of such models. Specifically, four different techniques based on machine learning methods have been tested: GPR, SVMs, ensemble of trees, and ANNs with an iteration limit set at 1,000 variables per layer. The tool used for the training and evaluation of these methods was the Regression Learner from Matlab™.

For the definition of the test set, we held out a 20% of the data, using the 80% left for training and validation. The validation method selected was the k-folding, in which we separate the training set in k folds, performing the training in k-1 folds and then using the one left for validation, changing afterwards the selected fold to be left out if training. During training, all methods used have been Bayesian optimization type, using 30 iterations and the expected improvement per second plus acquisition function.

Focusing on data analysis, the different techniques were applied for individual Raman and LIBS data, obtaining eight different models for each technique with their corresponding statistical parameters, which are R² of the training dataset and the root mean square error (RMSE) of validation and prediction sets (RMSEV and RMSEP respectively). These parameters were used as tools to describe, in the case of R² the goodness of the fit metric and in the case of, RMSEP the best performing model in the test set.

The combination of data for the joint evaluation based on both techniques was carried out using a concatenation algorithm, considering all the input variables selected for each individual technique obtained when

performing the dimensional reduction. The combination was made both including both the PCA data and the selected spectral parameters. This combination ensured that the resulting dataset contained complete and balanced information on both techniques. Due to the expanded number of variables in the resulting input vector compared to each individual dataset, we employed the MRMR algorithm to identify the most relevant characteristics. This algorithm assigned scores to each input variable based on its connection to the analysis objective. To avoid potential biases caused by differences in analytical fingerprints and different analysis points, we combined all LIBS entries with each Raman entry for each concentration sample, as was done in previous data combining studies.³¹ The reduction process yielded six new parameters for the combinations of PCA and spectral parameters in each case, representing the combined information from both techniques. Specifically, MRMR-based reduction demonstrated that three of these parameters were associated with LIBS information while the other three originated from information obtained through Raman analysis. This comprehensive approach allowed us to effectively integrate the strengths of both LIBS and Raman spectroscopy and obtain valuable insights for the joint analysis of the samples under study.

4 | RESULTS AND DISCUSSION

4.1 | Raman model

The chemometric analyses of the Raman dataset, using both input sets, the one based on PCA and the one based on spectral parameters, is presented in an actual versus predicted plot (Figure 3), focusing on the calcite quantity

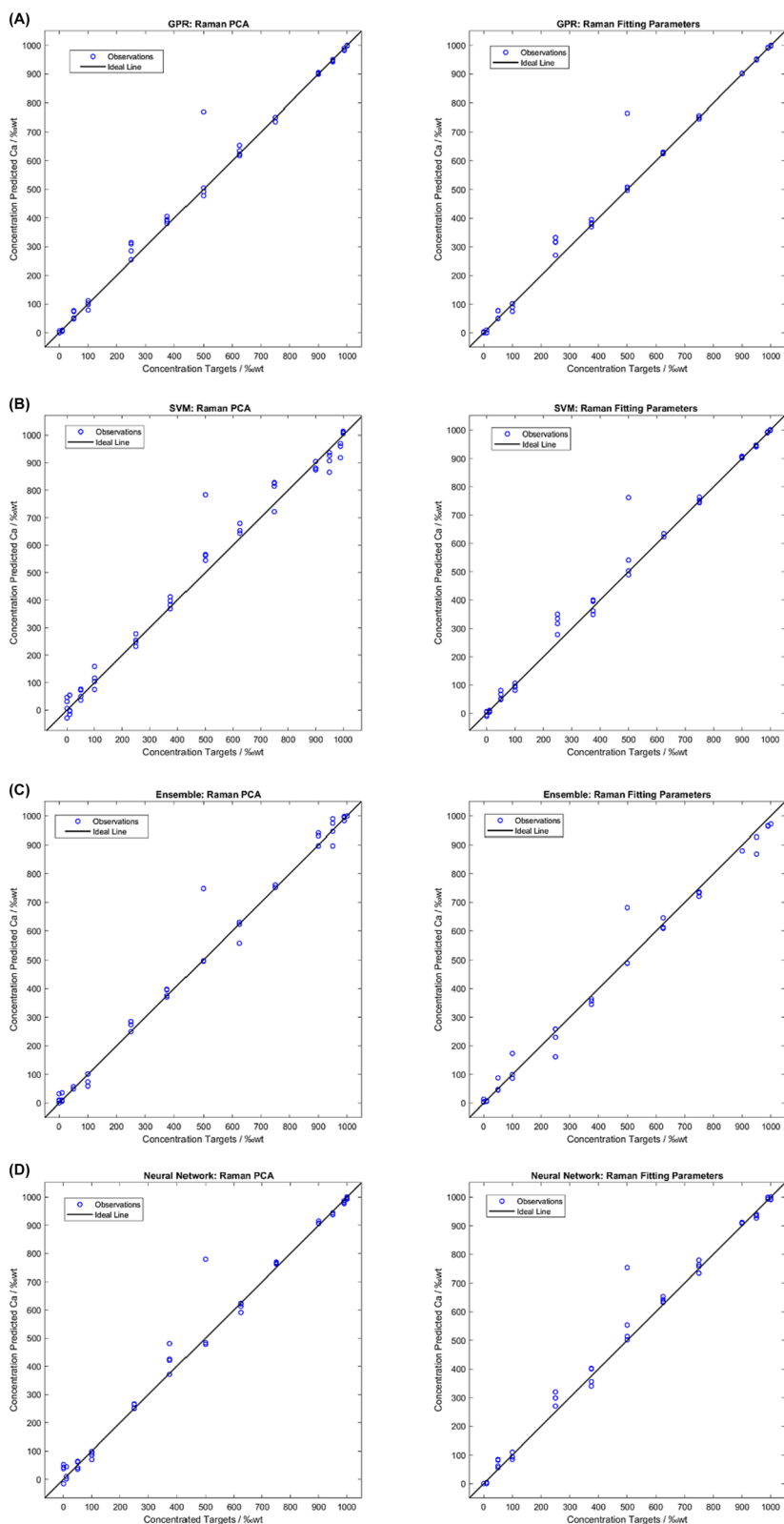


FIGURE 3 Estimation of the target concentration in wt% for selected Raman test samples. Each model is generated according to the multivariate analytical method used, both for the data processed by PCA and the preselection of spectral parameters. (A) GPR, (B) ensemble of trees, (C) SVM, and (D) ANN.

prediction. Starting with the GPR, both for the PCA-based model fitting parameters-model provided an equal coefficient R^2 value (0.99) and a similar RMSEP percentage (4.10 vs 4.16; see Table 2). Looking more in detail, both models performed the worst in the middle region of

the prediction curve. In fact, the highest dispersion occurs for samples with a calcite concentration ratio between 40% and 60%. The model performed significantly better for high calcite concentrations, while for low concentrations, the dispersion is slightly increased.

TABLE 2 Parameters in the regression metric for the different models for both processing (PCA and fitting parameters) based on the machine learning method used for the Raman, LIBS, and combination dataset.

Regression metrics in machine learning						
PCA	Raman		LIBS		Combination	
	R ²	RMSEP (%)	R ²	RMSEP (%)	R ²	RMSEP (%)
Gaussian	0.99	4.10	0.99	3.31	0.99	3.65
SVM	0.98	5.35	0.99	2.95	0.98	5.03
Ensemble of trees	0.96	4.09	0.99	3.50	0.99	3.32
Neuronal networks	0.98	4.56	0.99	4.03	0.99	3.57
Spectra parameters	Raman		LIBS		Combination	
	R ²	RMSEP (%)	R ²	RMSEP (%)	R ²	RMSEP (%)
Gaussian	0.99	4.16	0.99	2.81	1	1.67
SVM	0.99	4.34	0.99	3.16	1	2.42
Ensemble of trees	0.99	4.00	0.99	3.34	0.99	2.96
Neuronal networks	0.99	4.17	0.99	4.17	1	1.93

On the other hand, the models trained using SVM (Figure 3B) presented a R² coefficient of 0.98 and 0.99, respectively. Like the previous model, the error in the estimation of results remain wider for intermediate concentrations. Additionally, unlike the previous model, when preprocessing with PCA was performed, the dispersion of the data increased at the two ends of the prediction curve. In contrast to this result, when SVM was used on the spectral parameters, these variations at concentration extremes were not observed. As a result, this approach allows a perfect distinction between samples with a concentration of calcite-hydromagnesite below 5 wt% for high calcite concentrations.

Regarding the results for ensemble, the value of R² was of 0.99. Figure 3C shows how this model presents results with lower dispersion. Particularly noteworthy was the measurement for 50 wt% of calcite, which showed a significant reduction in the dispersion compared to previous models. However, the dispersion at the extremes was more significant than in both GPR models obtained using GPR.

Regarding the ANN models, the regressions had R² of 0.99. The reduced concentration dispersion for high calcite concentrations allowed, similar to GPR, the distinction between 5 wt% of calcite in both models. For very low calcite concentrations, the dispersion for the PCA-based model was quite significant, making it impossible to identify this compound when its concentration was less than 10 wt%. When using the fitting parameters-based model, low contents of calcite became detectable, although the use of these spectral features complicated the detectability of calcite in intermediate regions with concentrations around 25–40 wt%, something achievable in the model trained from PCs.

Across the eight trained models, a consistent overestimation was observed at one of the data point where the intended target concentration was 50 wt%, as the estimated values centered around 80 wt%. This overestimation pattern persisted uniformly for Raman test data. However, it is worth noting that this discrepancy did not impact the final outcome significantly. The statistical analysis remained in good agreement with the expected value for the complete dataset, underscoring the overall resilience of the models despite the recurrent anomaly at that specific point.

4.2 | LIBS model

Compared to Raman results, all LIBS models exhibited very similar performance in regression (R² is 0.99 in all cases; see Table 2). As displayed in Figure 4, all models ensured low dispersion of data at both high and low concentrations of calcite, except for the GPR and SVM models when using PCA, where distinguishing low Ca concentration became difficult. Additionally, high concentrations of calcite were also challenging to discriminate using the spectral parameters data set. Overall, in the intermediate region of the prediction curves, the dispersion of the results has increased compared to the curves obtained with Raman, with the dispersion of the test data often overlapping, diminishing confidence in our models.

The regression metrics presented by each of these models are shown in Table 2. Evaluating the degree of variation in the test data, it can be observed that, for ANN models, they provided the worst results in the test set as they had the highest value of RMSEP compared to the rest of the models. For the GPR and ensemble of trees

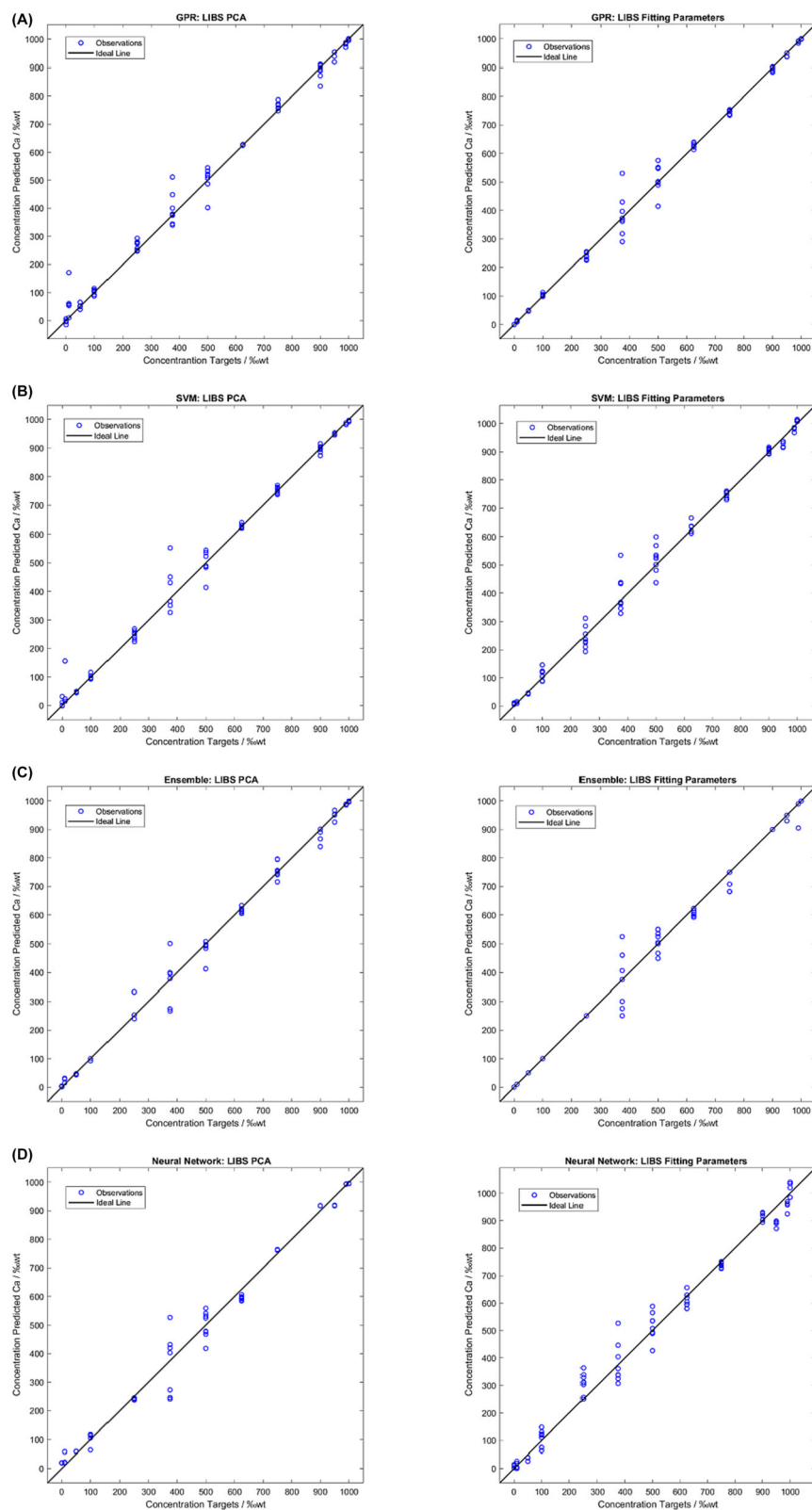


FIGURE 4 Estimation of targets concentration in wt% for test samples selected for LIBS data. Each model is generated according to the multivariate analytical method used, both for the data processed by PCA and the preselection of spectral parameters. (A) GPR, (B) ensemble of trees, (C) SVM, and (D) ANN.

models, the best results were obtained with the use of PCA, while in the models obtained by SVM and ANN, the spectral parameters data set yielded better results. Even though the difference in dispersion in the LIBS

models for the intermediate concentration regions was higher, the regression metrics have shown that the models trained using LIBS data generate, globally, better results than those obtained by Raman.

4.3 | Combined Raman-LIBS model

The results derived from the application of the machine learning methods to the combined Raman-LIBS dataset are represented in Figure 5. The combination of Raman

and LIBS parameters provided a data set with more variables that were not redundant, as those variables come from different techniques and different physical processes. The increase in the variables used to analyze the data causes the uncertainty of sample classification to

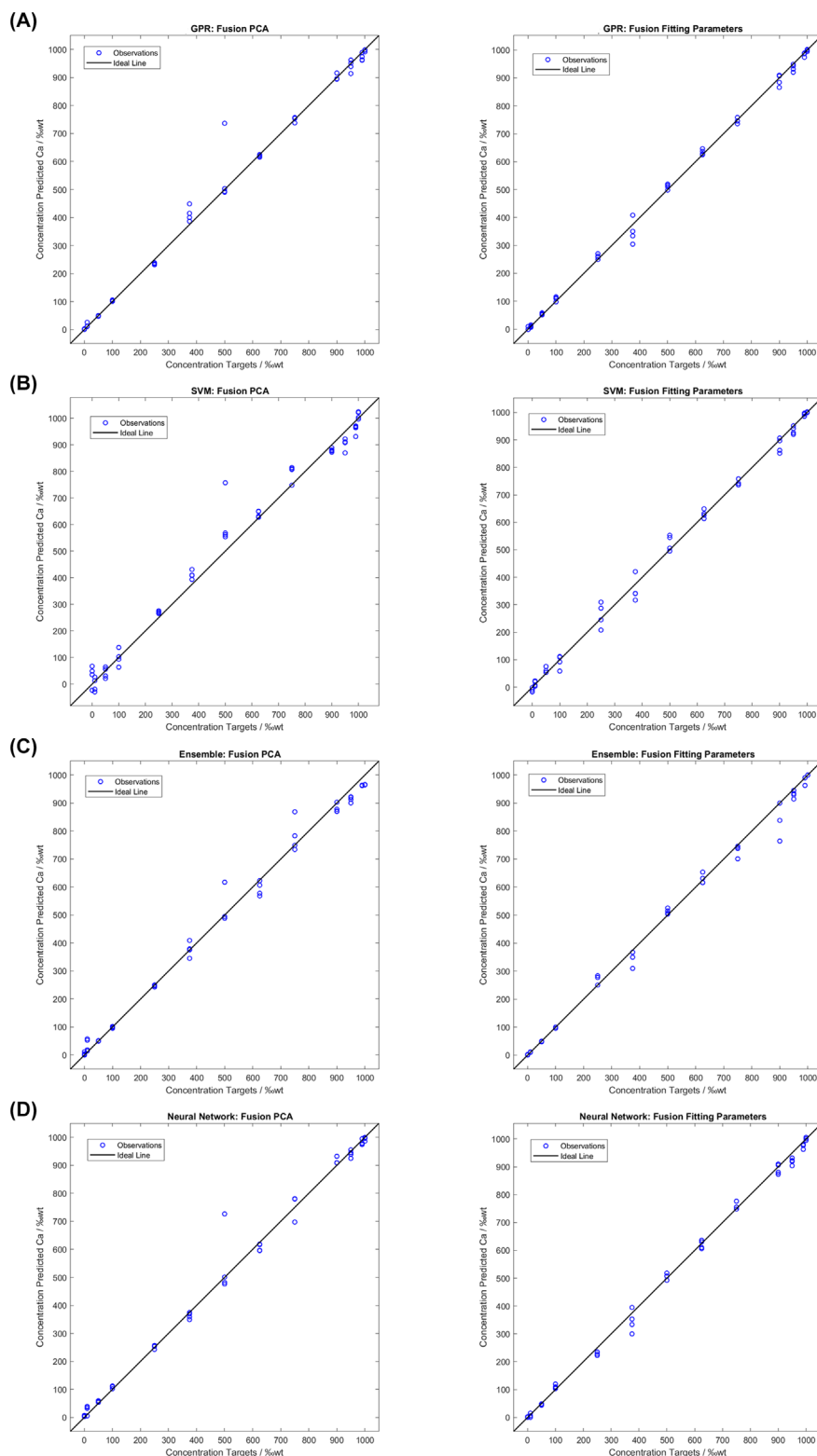


FIGURE 5 Estimation of targets concentration in wt% for test samples selected for data combination. Each model is generated according to the multivariate analytical method used, both for the data processed by PCA and the preselection of spectral parameters. (A) GPR, (B) ensemble of trees, (C) SVM, and (D) ANN.

decrease, this resulting in a general improvement in all the trained models. Beyond the expected statistical effect, this result underlined the complementarity of the information provided by Raman and LIBS techniques upon the prediction capability of the developed models. To prevent the possible improvement from the fact that we had more input variables available, we used, as described, MrMR algorithm to select the top six variables from this combined set. The algorithm selected three from each data set in the top six, indicating the complementarity of the techniques.

Analyzing in detail each of the resulting models, the predictions obtained for both the GPR and ANN methods (Figure 5A,D) provide improved results with a wide confidence interval across the entire concentration range. In both models, it was possible to distinguish with a high degree of accuracy between samples ranging from 1–99 to 0–100 wt%. Within this range, and only when the calcite concentration in the mixtures was very high (95–99–100 wt%), we can say that the GPR-based technique using spectra parameters presents more accurate and precise values. However, the GPR model exhibits more significant dispersion for intermediate concentrations over ANN-based models.

The SVM and ensemble of trees models based on the combined Raman-LIBS data analysis provided greater data dispersion for extreme concentrations compared to the previous models, except for the SVM model trained with spectra parameters, where extreme values also generated results with good discrimination power. In particular, the graph represented in Figure 5C (ensemble model with PCA) tends to underestimate the concentration of calcite above 90 wt%. Overall, all models exhibited less dispersion than those obtained from LIBS data with the same number of input parameters to be used. It is worth noting that, when studying through PCA parameters, all models showed a fairly scattered test result for a concentration of 50 wt% calcite, similar to what was observed with the Raman-trained models. This dispersion completely disappeared when using filtered and selected spectral parameters, as observed in Figure 5.

4.4 | Comparison of results

The regression metrics obtained from the application of the machine learning methods to Raman, LIBS, and combined Raman-LIBS datasets for both preprocessing approaches are shown in Table 2. The results are represented as the mean in the concentration prediction for each set of samples used for testing, depending on the actual theoretical value of the concentration of each mixture. Data analyses were performed for each individual

spectroscopic technique as well as for the combination of both data sets. Summary metrics for the goodness of fit of the different models are shown in Table 2, where R^2 is shown as a goodness of the regression fitting on the training set. The root mean square error of prediction (RMSEP) has also been calculated to assess which model performed best in the test set that would give a hint of the expected accuracy using new data.

By combining the detailed results presented in Sections 4.1, 4.2, and 4.3 with the regression metrics provided in Table 2, it can be inferred that the content prediction model based on the combined Raman-LIBS dataset was always superior to that achieved solely by Raman data when using spectral feature selection after filtering characteristic parameters. However, when measurements were performed using only PCA parameters, the ensemble of trees and ANN models were slightly better than those obtained solely from LIBS. The data combination yielded improved results when employing GPR and ANN based on spectral parameters. In these cases, the regression coefficient was close to unity, while the RMSEP value is below 2%.

5 | CONCLUSIONS

In this work, several chemometric methods have been used to develop models for the analysis of Raman, LIBS, and combined Raman-LIBS data for the evaluation of the quantification capabilities in mineral binary mixtures.

Despite the apparent higher dispersion exhibited by the models trained for LIBS compared to those of Raman, the RMSEP calculated for the Raman models is greater than that of LIBS or the combined data, indicating that the concentration values estimated by these methods better fit real values with higher precision. It is also important to highlight that the detection limit and behavior of the LIBS models improved for extreme concentration values. Consequently, the LIBS models effectively distinguished samples with 0 and 1 wt% of calcite. Indeed, Raman effect is fainter than the emission lines of an induced plasma, and while we used different spectral lines from LIBS, this could not be the case for Raman for different reasons. First, the high difference in SNR from the main band to secondary bands makes that in lower concentrations, these bands would be confusing for the training of models. On the other hand, in the scope of this work, more focused on planetary exploration instruments, these instruments show a limited Raman range closer to the Rayleigh, as SHERLOC, or have lower SNR that impairs the detection of secondary bands.

Furthermore, the use of the same number of input variables in every case allowed us to infer that the models

using spectral parameters generated better results than PCA-based models, ensuring a lower RMSEP and R^2 close to unity in all cases. In this particular case, it looks like the selection by operator of the informative parameters to be used outperforms the statistical definition of variables through PCAs.

When analyzing the combined Raman-LIBS PCA dataset, the results of all machine learning methods improved significantly over all the models based on Raman data alone. Data combination also displayed improved results over LIBS dataset when Ensemble and ANN are used. In detail, the RMSEP obtained in both models is around 3%, this being slightly lower than the RMSEP values obtained from the analysis of LIBS (between 3.5 and 4%) or Raman (from 4 and 4.5%) datasets alone.

On the other hand, models trained using preprocessing with spectral parameters demonstrate how the combination of Raman-LIBS data significantly improves upon the data obtained from each individual technique, achieving an RMSEP better than 2% and an R^2 equal to unity for models trained with Gaussian and neural networks. This result concurs with the previous observation on the PCA based versus the spectral parameters-based models, in this case on the combined data set. Combined models could harness the best aspects of all variables, having information from different techniques and physical effects, and the combined models outperformed those based on individual datasets. PCA or spectral parameters variables can be, at the end, reiterative beyond one point, in the case of PCA, as we increase the number of PCs, we reduce the variance associated with the new variables added to the input set, introducing information that can be confusing or just redundant. The use of variables from the two techniques introduces more information that is not redundant or confusing to the models, as is demonstrated. Presented results demonstrate that the Raman-LIBS data combination can enhance the quantification of mineral abundance in binary mixtures, if the appropriate machine learning method is employed.

Models performing best in the middle range of concentrations are the ones performing worst for the extreme values. As such, an ensemble of trees models, assigning weights to the prediction of each model depending on the calculated concentration, could offer a good compromise for the whole range of concentrations. This aspect will be further evaluated in a future work, where blended models could be used parting from predictions of different models, as has been done for SuperCam.⁴⁷

It is important to note that this research focused on the study of mineral phases that are relatively easy to distinguish by both techniques. As such, the advantages provided by the combination of Raman and LIBS dataset

should further improve when analyzing more complex mineralogical compositions (e.g., including the presence of polymorphs, or solid solutions). For all these reasons, the decision to utilize both techniques with an equal number of parameters for each allowed us to leverage their complementary strengths and obtain a more comprehensive and accurate understanding of the studied samples.

As this work proved that the prediction based on combined spectroscopic data enhances the mineral ratio inferred by machine learning methods, the potential application of such chemometric tools to the Raman-LIBS data the SuperCam instrument is collecting on Mars should be further explored. Beyond the Mars 2020 mission, and given the future development based on Raman spectroscopy, the use of machine learning to combine spectroscopic data of Raman with different techniques could also find a reliable application to forthcoming planetary missions, as is the case of the ESA Rosalind Franklin mission.

ACKNOWLEDGMENTS

This work was supported by the RamOnMars project funded by the Agencia Estatal de Investigación (Spain), Grant PID2022-142490OB-C32. The authors gratefully acknowledge the support of the SIGUE-Mars Consortium (Ministry of Economy and Competitiveness MINECO, grant RDE2018-102600-T). Additional support came from the European Union-NextGenerationEU that funded Margarita Salas Grants and Programa Investigo. Furthermore, the authors thank the European Social Fund and the Consejería de Educación de Castilla y León.

ORCID

Sofia Julve-Gonzalez  <https://orcid.org/0000-0001-9868-9931>

Jose A. Manrique  <https://orcid.org/0000-0002-2053-2819>

Marco Veneranda  <https://orcid.org/0000-0002-7185-2791>

Guillermo Lopez-Reyes  <https://orcid.org/0000-0003-1005-1760>

REFERENCES

- [1] S. Maurice, R. C. Wiens, M. Saccoccio, B. Barraclough, O. Gasnault, O. Forni, N. Mangold, D. Baratoux, S. Bender, G. Berger, J. Bernardin, M. Berthé, N. Bridges, D. Blaney, M. Bouyé, P. Caïs, B. Clark, S. Clegg, A. Cousin, D. Cremers, A. Cros, L. Deflores, C. Derycke, B. Dingler, G. Dromart, B. Dubois, M. Dupieux, E. Durand, L. D'Uston, C. Fabre, B. Faure, A. Gaboriaud, T. Gharsa, K. Herkenhoff, E. Kan, L. Kirkland, D. Kouach, J. L. Lacour, Y. Langevin, J. Lasue, S. le Mouélic, M. Lescure, E. Lewin, D. Limonadi, G. Manhès, P.

- Mauchien, C. McKay, P. Y. Meslin, Y. Michel, E. Miller, H. E. Newsom, G. Orttner, A. Paillet, L. Parès, Y. Parot, R. Pérez, P. Pinet, F. Poitrasson, B. Quertier, B. Sallé, C. Sotin, V. Sautter, H. Séran, J. J. Simmonds, J. B. Sirven, R. Stiglich, N. Striebig, J. J. Thocaven, M. J. Toplis, D. Vaniman, *Space Sci. Rev.* **2012**, *170*, 95.
- [2] A. Corpolongo, R. S. Jakubek, S. A. Asher, D. Baker, L. W. Beegle, E. L. Berger, R. Bhartia, A. J. Brown, A. S. Burton, S. V. Bykov, E. Cardarelli, E. A. Cloutis, P. Conrad, A. D. Czaja, L. Deflores, D. Flannery, T. Fornaro, M. Fries, N. C. Haney, K. Hickman-lewis, L. Kah, C. Lee, *J. Geophys. Res. Planets* **2023**, *128*(3), e2022JE007455.
- [3] R. Bhartia, L. W. Beegle, L. DeFlores, W. Abbey, J. R. Hollis, K. Uckert, B. Monacelli, K. S. Edgett, M. R. Kennedy, M. Sylvia, D. Aldrich, M. Anderson, S. A. Asher, Z. Bailey, K. Boyd, A. S. Burton, M. Caffrey, M. J. Calaway, R. Calvet, B. Cameron, M. A. Caplinger, B. L. Carrier, N. Chen, A. Chen, M. J. Clark, S. Clegg, P. G. Conrad, M. Cooper, K. N. Davis, B. Ehlmann, L. Facto, M. D. Fries, D. H. Garrison, D. Gasway, F. T. Ghaemi, T. G. Graff, K. P. Hand, C. Harris, J. D. Hein, N. Heinz, H. Herzog, E. Hochberg, A. Houck, W. F. Hug, E. H. Jensen, L. C. Kah, J. Kennedy, R. Krylo, J. Lam, M. Lindeman, J. McGlown, J. Michel, E. Miller, Z. Mills, M. E. Minitti, F. Mok, J. Moore, K. H. Neelson, A. Nelson, R. Newell, B. E. Nixon, D. A. Nordman, D. Nuding, S. Orellana, M. Pauken, G. Peterson, R. Pollock, H. Quinn, C. Quinto, M. A. Ravine, R. D. Reid, J. Riendeau, A. J. Ross, J. Sackos, J. A. Schaffner, M. Schwochert, M. O. Shelton, R. Simon, C. L. Smith, P. Sobron, K. Steadman, A. Steele, D. Thiessen, V. D. Tran, T. Tsai, M. Tuite, E. Tung, R. Wehbe, R. Weinberg, R. H. Weiner, R. C. Wiens, K. Williford, C. Wollonciej, Y. H. Wu, R. A. Yingst, J. Zan, *Space Sci. Rev.* **2021**, *217*.
- [4] F. Rull, S. Maurice, I. Hutchinson, A. Moral, C. Perez, C. Diaz, M. Colombo, T. Belenguer, G. Lopez-Reyes, A. Sansano, O. Forni, Y. Parot, N. Striebig, S. Woodward, C. Howe, N. Tarcea, P. Rodriguez, L. Seoane, A. Santiago, J. A. Rodriguez-Prieto, J. Medina, P. Gallego, R. Canchal, P. Santamaría, G. Ramos, J. L. Vago, *Astrobiology* **2017**, *17*, 627.
- [5] R. Vance, Y. Cho, S. Kameda, P. Rodriguez, F. Rull, S. Ulamec, T. Usui, RAX—A Raman spectrometer for MMX onboard the rover for Phobos, **2019**.
- [6] R. C. Wiens, S. Maurice, F. R. Perez, *Spectroscopy (Santa Monica)* **2017**, *32*, 50.
- [7] R. C. Wiens, S. Maurice, S. H. Robinson, A. E. Nelson, P. Cais, P. Bernardi, R. T. Newell, S. Clegg, S. K. Sharma, S. Storms, J. Deming, D. Beckman, A. M. Ollila, O. Gasnault, R. B. Anderson, Y. André, S. M. Angel, G. Arana, E. Auden, P. Beck, J. Becker, K. Benzerara, S. Bernard, O. Beyssac, L. Borges, B. Bousquet, K. Boyd, M. Caffrey, J. Carlson, K. Castro, J. Celis, B. Chide, K. Clark, E. Cloutis, E. C. Cordoba, A. Cousin, M. Dale, L. Deflores, D. Delapp, M. Deleuze, M. Dirmyer, C. Donny, G. Dromart, M. G. Duran, M. Egan, J. Ervin, C. Fabre, A. Fau, W. Fischer, O. Forni, I. Gontijo, J. Grotzinger, X. Jacob, S. Jacquino, J. Laserna, J. Lasue, S. Le, M. Carey, L. Iv, R. Leveille, E. Lewin, G. L. Ralph, E. Lorigny, S. P. Love, B. Lucero, J. Manuel, *Space Sci. Rev.* **2021**, *217*, 4.
- [8] S. Maurice, R. C. Wiens, P. Bernardi, P. Cais, S. Robinson, T. Nelson, O. Gasnault, J. M. Reess, M. Deleuze, F. Rull, J. A. Manrique, S. Abbaki, R. B. Anderson, Y. André, S. M. Angel, G. Arana, T. Battault, P. Beck, K. Benzerara, S. Bernard, J. P. Berthias, O. Beyssac, M. Bonafous, B. Bousquet, M. Boutillier, A. Cadu, K. Castro, F. Chapron, B. Chide, K. Clark, E. Clavé, S. Clegg, E. Cloutis, C. Collin, E. C. Cordoba, A. Cousin, J. C. Dameury, W. D'Anna, Y. Daydou, A. Debus, L. Deflores, E. Dehouck, D. Delapp, G. D. L. Santos, C. Donny, A. Doressoundiram, G. Dromart, B. Dubois, A. Dufour, M. Dupieux, M. Egan, J. Ervin, C. Fabre, A. Fau, W. Fischer, O. Forni, T. Fouchet, J. Frydenvang, S. Gauffre, M. Gauthier, V. Gharakanian, O. Gilard, I. Gontijo, R. Gonzalez, D. Granena, J. Grotzinger, R. Hassen-Khodja, M. Heim, Y. Hello, G. Hervet, O. Humeau, X. Jacob, S. Jacquino, J. R. Johnson, D. Kouach, G. Lacombe, N. Lanza, L. Lapauw, J. Laserna, J. Lasue, L. le Deit, S. le Mouélic, E. le Comte, Q. M. Lee, C. Legett, R. Leveille, E. Lewin, C. Leyrat, G. Lopez-Reyes, R. Lorenz, B. Lucero, J. M. Madariaga, S. Madsen, M. Madsen, N. Mangold, F. Manni, J. F. Mariscal, J. Martinez-Frias, K. Mathieu, R. Mathon, K. P. McCabe, T. McConnochie, S. M. McLennan, J. Mekki, N. Melikechi, P. Y. Meslin, Y. Mischeau, Y. Michel, J. M. Michel, D. Mimoun, A. Misra, G. Montagnac, C. Montaron, F. Montmessin, J. Moros, V. Mousset, Y. Morizet, N. Murdoch, R. T. Newell, H. Newsom, N. N. Tuong, A. M. Ollila, G. Orttner, L. Oudda, L. Pares, J. Parisot, Y. Parot, R. Pérez, D. Pheav, L. Picot, P. Pilleri, C. Pilorget, P. Pinet, G. Pont, F. Poulet, C. Quantin-Nataf, B. Quertier, D. Rambaud, W. Rapin, P. Romano, L. Roucayrol, C. Royer, M. Ruellan, B. F. Sandoval, V. Sautter, M. J. Schoppers, S. Schröder, H. C. Seran, S. K. Sharma, P. Sobron, M. Sodki, A. Sournac, V. Sridhar, D. Standarovsky, S. Storms, N. Striebig, M. Tatat, M. Toplis, I. Torre-Fdez, N. Toulemont, C. Velasco, M. Veneranda, D. Venhaus, C. Virmondois, M. Viso, P. Willis, K. W. Wong, *Space Sci. Rev.* **2021**, *217*(3), 47.
- [9] J. A. Manrique, G. Lopez-Reyes, A. Cousin, F. Rull, S. Maurice, R. C. Wiens, M. B. Madsen, J. M. Madariaga, O. Gasnault, J. Aramendia, G. Arana, P. Beck, S. Bernard, P. Bernardi, M. H. Bernt, A. Berrocal, O. Beyssac, P. Cais, C. Castro, K. Castro, S. M. Clegg, E. Cloutis, G. Dromart, C. Drouet, B. Dubois, D. Escribano, C. Fabre, A. Fernandez, O. Forni, V. Garcia-Baonza, I. Gontijo, J. Johnson, J. Laserna, J. Lasue, S. Madsen, E. Mateo-Marti, J. Medina, P. Y. Meslin, G. Montagnac, A. Moral, J. Moros, A. M. Ollila, C. Ortega, O. Prieto-Ballesteros, J. M. Reess, S. Robinson, J. Rodriguez, J. Saiz, J. A. Sanz-Arranz, I. Sard, V. Sautter, P. Sobron, M. Toplis, M. Veneranda, *Space Sci. Rev.* **2020**, *216*, 216.
- [10] K. H. Williford, K. A. Farley, K. M. Stack, A. C. Allwood, D. Beaty, L. W. Beegle, R. Bhartia, A. J. Brown, M. de la Torre Juarez, S. E. Hamran, M. H. Hecht, J. A. Hurowitz, J. A. Rodriguez-Manfredi, S. Maurice, S. Milkovich, R. C. Wiens, *From Habitability to Life on Mars*, Elsevier **2018** 275. ISBN 9780128099353.
- [11] A. M. Zastrow, T. D. Glotch, *Geophys. Res. Lett.* **2021**, *48*, e2020GL092365. <https://doi.org/10.1029/2020GL092365>
- [12] B. H. N. Horgan, R. B. Anderson, G. Dromart, E. S. Amador, M. S. Rice, *Icarus* **2020**, *339*, 113526. <https://doi.org/10.1016/j.icarus.2019.113526>
- [13] R. C. Wiens, A. Udry, O. Beyssac, C. Quantin-Nataf, N. Mangold, A. Cousin, L. Mandon, T. Bosak, O. Forni, S. M. McLennan, V. Sautter, A. Brown, K. Benzerara, J. R. Johnson, L. Mayhew, S. Maurice, R. B. Anderson, S. M. Clegg, L.

- Crumpler, T. S. J. Gabriel, P. Gasda, J. Hall, B. H. N. Horgan, L. Kah, C. Legett Iv, J. M. Madariaga, P.-Y. Meslin, A. M. Ollila, F. Poulet, C. Royer, S. K. Sharma, S. Siljeström, J. I. Simon, T. E. Acosta-Maeda, C. Alvarez-Llamas, S. M. Angel, G. Arana, P. Beck, S. Bernard, T. Bertrand, B. Bousquet, K. Castro, B. Chide, E. Clavé, E. Cloutis, S. Connell, E. Dehouck, G. Dromart, W. Fischer, T. Fouchet, R. Francis, J. Frydenvang, O. Gasnault, E. Gibbons, S. Gupta, E. M. Hausrath, X. Jacob, H. Kalucha, E. Kelly, E. Knutsen, N. Lanza, J. Laserna, J. Lasue, S. le Mouélic, R. Leveille, G. L. Reyes, R. Lorenz, J. A. Manrique, J. Martinez-Frias, T. Mcconnochie, N. Melikechi, D. Mimoun, F. Montmessin, J. Moros, N. Murdoch, P. Pilleri, C. Pilorget, P. Pinet, W. Rapin, F. Rull, S. Schröder, D. L. Shuster, R. J. Smith, A. E. Stott, J. Tarnas, N. Turenne, M. Veneranda, D. S. Vogt, B. P. Weiss, P. Willis, K. M. Stack, K. H. Williford, K. A. Farley, *Sci. Adv.* **2022**, *8*(34), eabo3399.
- [14] I. Melendez, K. Grice, K. Trinajstic, M. Ladjavardi, P. Greenwood, K. Thompson, *Geology* **2013**, *41*, 123.
- [15] S. Holm-Alwmark, K. M. Kinch, M. D. Hansen, S. Shahrzad, K. Svennevig, W. J. Abbey, R. B. Anderson, F. J. Calef III, S. Gupta, E. Hauber, B. H. N. Horgan, L. C. Kah, J. Knade, N. B. Miklusicak, K. M. Stack, V. Z. Sun, J. D. Tarnas, C. Quantin-Nataf, *J. Geophys. Res. Planets* **2021**, *126*, e2021JE006840.
- [16] R. E. Summons, J. P. Amend, D. Bish, R. Buick, G. D. Cody, D. J. des Marais, G. Dromart, J. L. Eigenbrode, A. H. Knoll, D. Y. Sumner, *Astrobiology* **2011**, *11*, 157.
- [17] E. Clavé, K. Benzerara, P.-Y. Meslin, O. Forni, C. Royer, L. Mandon, P. Beck, C. Quantin-Nataf, O. Beyssac, A. Cousin, B. Bousquet, R. C. Wiens, S. Maurice, E. Dehouck, S. Schröder, O. Gasnault, N. Mangold, G. Dromart, T. Bosak, S. Bernard, A. Udry, R. B. Anderson, G. Arana, A. J. Brown, K. Castro, S. M. Clegg, E. Cloutis, A. G. Fairén, D. T. Flannery, P. J. Gasda, J. R. Johnson, J. Lasue, G. Lopez-Reyes, J. M. Madariaga, J. A. Manrique, S. le Mouélic, J. I. Núñez, A. M. Ollila, P. Pilleri, C. Pilorget, P. Pinet, F. Poulet, M. Veneranda, Z. U. Wolf, *J. Geophys. Res. Planets* **2022**, *128*. <https://doi.org/10.1029/2022JE007463>
- [18] G. Lagaly, M. Ogawa, I. Dékány, *Dev. Clay Sci.* **2006**, *1*, 309.
- [19] S. Sharma, R. D. Roppel, A. E. Murphy, L. W. Beegle, R. Bhartia, A. Steele, J. R. Hollis, S. Siljeström, F. M. McCubbin, S. A. Asher, W. J. Abbey, A. C. Allwood, E. L. Berger, B. L. Bleefeld, A. S. Burton, S. V. Bykov, E. L. Cardarelli, P. G. Conrad, A. Corpolongo, A. D. Czaja, L. P. DeFlores, K. Edgett, K. A. Farley, T. Fornaro, A. C. Fox, M. D. Fries, D. Harker, K. Hickman-Lewis, J. Huggett, S. Imbeah, R. S. Jakubek, L. C. Kah, C. Lee, Y. Liu, A. Magee, M. Minitti, K. R. Moore, A. Pascuzzo, C. R. Sanchez-Vahamonde, E. L. Scheller, S. Shkolyar, K. M. Stack, K. Steadman, M. Tuite, K. Uckert, A. Werynski, R. C. Wiens, A. J. Williams, K. Winchell, M. R. Kennedy, A. Yanchilina, *Nature* **2023**, *619*, 724.
- [20] Y. Chen, Y. Sun, L. Liu, J. Shen, Y. Qu, Y. Pan, W. Lin, *Astrobiology* **2023**, *23*, 172.
- [21] D. Wacey, *Astrobiology* **2010**, *10*(4), 381.
- [22] A. E. Murphy, S. T. Wieman, J. Gross, J. C. Stern, A. Steele, M. Glamoclija, *Sediment. Geol.* **2020**, *410*, 105777.
- [23] A. Goodwin, D. Papineau, *Astrobiology* **2022**, *22*, 49.
- [24] J. A. Manrique Martinez, M. Veneranda, G. Lopez Reyes, J. Saiz, J. A. Sanz, R. Navarro, J. Medina García, F. Rull, *GeoRaman Congress 2020, Virtual*, 14th International GeoRaman conferences Book of Abstract, Bilbao, Basque Country **2020**, 52.
- [25] M. Veneranda, J. A. Manrique, G. Lopez-Reyes, S. Julve-Gonzalez, F. Rull, C. Alvarez Llamas, T. Delgado Pérez, E. Gibbons, E. Clavé, E. Cloutis, J. Huidobro, K. Castro, J. M. Madariaga, N. Randazzo, A. Brown P. Willis, S. Maurice, R. C. Wiens, *Earth Space Sci.* **2023**, *10*, 7. <https://doi.org/10.1029/2023EA002829>
- [26] S. Se Pa, M. B. Yakoob, P. Maruthai, K. Singaravelu, N. Duraisamy, R. D. Palaniappan, J. B. Pithai, *Eng. Proc.* **2022**, *19*(1), 26. <https://doi.org/10.3390/ECP2022-12631>
- [27] R. Syah, N. Ahmadian, M. Elveny, S. M. Alizadeh, M. Hosseini, A. Khan, *Energy Rep.* **2021**, *7*, 4106.
- [28] P. A. Moreno-Sanchez, *Proceedings - 2020 IEEE international conference on big data, big data 2020*, Institute of Electrical and Electronics Engineers Inc. **2020**, 4902. CFP20BGD-POD.
- [29] A. Ghosh, S. Raha, S. Dey, K. Chatterjee, A. Roy Chowdhury, A. Barui, *Analyst* **2019**, *144*, 1309.
- [30] X. Sang, R. Zhou, Y. Li, S. Xiong, *Neural Process. Lett.* **2022**, *54*, 677.
- [31] J. A. Manrique-Martinez, G. Lopez-Reyes, A. Alvarez-Perez, T. Bozic, M. Veneranda, A. Sanz-Arranz, J. Saiz, J. Medina-Garcia, F. Rull-Perez, *J. Raman Spectrosc.* **2020**, *51*, 1702.
- [32] J. A. Manrique, M. Veneranda, Y. Merino-Lomas, F. Rull, E. Charro, M. A. Gonzalez, J. M. Lopez, E. R. Gutiez, J. A. Sanz-Arranz, S. Maurice, G. Lopez-Reyes, *J. Chemom.* **2023**, *37*(9), e3440.
- [33] S. Wang, W. Pedrycz, Q. Zhu, W. Zhu, *Knowl. Based Syst.* **2015**, *75*, 19.
- [34] H. G. M. Edwards, C. D. Moody, E. M. Newton, S. E. J. Villar, M. J. Russell, *Icarus* **2005**, *175*, 372.
- [35] T. M. Burnie, I. M. Power, C. Paulo, H. Alçiçek, L. I. Falcón, Y. Lin, S. A. Wilson, *Astrobiology* **2023**, *23*, 513.
- [36] M. E. Sanz-Montero, Ó. Cabestrero, M. Sánchez-Román, *Front. Microbiol.* **2019**, *10*, 148. <https://doi.org/10.3389/fmicb.2019.00148>
- [37] J. Saiz, G. Lopez-Reyes, M. Veneranda, J. A. Manrique, A. Guzmán, D. Moreno-Dominguez, S. Werner, F. Poulet, J. Medina, F. Rull, *Geophysical Research Abstracts*, Vol. 21, EGU **2019**, EGU2019-19704.
- [38] M. Veneranda, A. Sanz-Arranz, J. A. Manrique, J. Saiz, C. Garcia-Prieto, E. Pascual-Sánchez, J. Medina, M. Konstantinidis, E. Lalla, A. Moral, L. M. Nieto, F. Rull, G. Lopez-Reyes, *J. Raman Spectrosc.* **2021**, *53*, 1.
- [39] A. Tharwat, *Int. J. Appl. Pattern Recogn.* **2016**, *3*, 197.
- [40] M. Konstantinidis, E. A. Lalla, M. G. Daly, G. Lopez-Reyes, J. M. Stromberg, K. Cote, E. A. Cloutis, *Icarus* **2020**, *358*, 114113.
- [41] M. Veneranda, G. Lopez-Reyes, E. Pascual Sanchez, A. M. Krzesińska, J. A. Manrique-Martinez, A. Sanz-Arranz, C. Lantz, E. Lalla, A. Moral, J. Medina, F. Poulet, H. Dypvik, S. C. Werner, J. L. Vago, F. Rull, *Astrobiology* **2021**, *21*, 307.
- [42] M. Veneranda, G. Lopez-Reyes, J. A. Manrique, J. Medina, P. Ruiz-Galende, I. Torre-Fdez, K. Castro, C. Lantz, F. Poulet, H. Dypvik, S. C. Werner, F. Rull, *Astrobiology* **2020**, *20*(3), 349. ast.2019.2095.
- [43] C. Perez Canora, J. A. Rodriguez, F. Musso, A. Moral, L. Seoane, J. Zafra, P. R. Rodriguez, S. Ibarria, M. Benito, M. Veneranda, J. A. Manrique, G. Ramos, E. Charro, J. M. Lopez,

- M. Á. González, I. Hutchinson, O. Prieto-Ballesteros, F. Rull, G. Lopez-Reyes, *J. Raman Spectrosc.* **2022**, *53*, 396.
- [44] G. Lopez-Reyes, P. Sobron, C. Lefebvre, F. Rull, *Am. Mineral.* **2014**, *99*, 1570.
- [45] S. Schröder, S. G. Pavlov, I. Rauschenbach, E. K. Jessberger, H. W. Hübers, *Icarus* **2013**, *223*, 61.
- [46] S. Maurice, S. M. Clegg, R. C. Wiens, O. Gasnault, W. Rapin, O. Forni, A. Cousin, V. Sautter, N. Mangold, L. le Deit, M. Nachon, R. B. Anderson, N. L. Lanza, C. Fabre, V. Payré, J. Lasue, P. Y. Meslin, R. J. Léveillé, B. L. Barraclough, P. Beck, S. C. Bender, G. Berger, J. C. Bridges, N. T. Bridges, G. Dromart, M. D. Dyar, R. Francis, J. Frydenvang, B. Gondet, B. L. Ehlmann, K. E. Herkenhoff, J. R. Johnson, Y. Langevin, M. B. Madsen, N. Melikechi, J. L. Lacour, S. le Mouélic, E. Lewin, H. E. Newsom, A. M. Ollila, P. Pinet, S. Schröder, J. B. Sirven, R. L. Tokar, M. J. Toplis, C. D'Uston, D. T. Vaniman, A. R. Vasavada, *J. Anal. At. Spectrom.* **2016**, *31*, 863.
- [47] R. B. Anderson, O. Forni, A. Cousin, R. C. Wiens, S. M. Clegg, J. Frydenvang, T. S. J. Gabriel, A. Ollila, S. Schröder, O. Beyssac, E. Gibbons, D. S. Vogt, E. Clavé, J.-A. Manrique, C. Leggett, P. Pilleri, R. T. Newell, J. Sarrao, S. Maurice, G. Arana, K. Benzerara, P. Bernardi, S. Bernard, B. Bousquet, A. J. Brown, C. Alvarez-Llamas, B. Chide, E. Cloutis, J. Comellas, S. Connell, E. Dehouck, D. M. Delapp, A. Essunfeld, C. Fabre, T. Fouchet, C. Garcia-Florentino, L. García-Gómez, P. Gasda, O. Gasnault, E. M. Hausrath, N. L. Lanza, J. Laserna, J. Lasue, G. Lopez, J. M. Madariaga, L. Mandon, N. Mangold, P.-Y. Meslin, A. E. Nelson, H. Newsom, A. L. Reyes-Newell, S. Robinson, F. Rull, S. Sharma, J. I. Simon, P. Sobron, I. T. Fernandez, A. Udry, D. Venhaus, S. M. McLennan, R. V. Morris, B. Ehlmann, *Spectrochim. Acta Part B at. Spectrosc.* **2022**, *188*, 106347.
- [48] M. Konstantinidis, K. Cote, E. A. Lalla, G. Zhang, M. G. Daly, X. Gao, P. Dietrich, *J. Chemom.* **2019**, *33*, e3174.

How to cite this article: S. Julve-Gonzalez, J. A. Manrique, M. Veneranda, I. Reyes-Rodríguez, E. Pascual-Sanchez, A. Sanz-Arranz, M. Konstantinidis, E. A. Lalla, M. E. Charro, E. Rodriguez-Gutierrez, J. M. Lopez-Rodríguez, J. F. Sanz-Requena, J. Delgado-Iglesias, M. A. Gonzalez, F. Rull, G. Lopez-Reyes, *J Raman Spectrosc* **2023**, *54*(11), 1353. <https://doi.org/10.1002/jrs.6611>