Word Count: 10,464 with references, 7,406 without references

Revision 2

Estimating Modal Mineralogy using Raman Spectroscopy: Multivariate Analysis Models and Raman Cross-Section Proxies

Laura B. Breitenfeld[1], M. Darby Dyar[2,3], Timothy D. Glotch[4], A. Deanne Rogers[4], and Miriam Eleazer[5]

[1]Department of Earth and Environmental Science, Wesleyan University, Middletown, CT 06459

[2]Department of Astronomy, Mount Holyoke College, South Hadley, MA 01075

[3]Planetary Science Institute, Tucson, AZ 85719

[4]Department of Geosciences, Stony Brook University, Stony Brook, NY 11794

[5]Department of Astronomy, University of Massachusetts Amherst, Amherst, MA 01003

Abstract

Raman spectroscopy is a powerful technique in the context of planetary exploration because it provides mineral identification, chemistry, and abundance information. For Raman spectrometers with large spot sizes, multiple mineral phases can be investigated via the collection of a single Raman spectrum. There is a lack of methodology for quantifying mineral species in mixtures due to independent signal strengths of different materials in Raman spectra. Two techniques are presented in this work for the quantification of common rock-forming minerals: partial least squares multivariate analysis and a novel approach called Raman cross-

section proxies (numerical metrics associated with specific Raman features). This paper targets 20 mineral species relevant to the mineralogy of the planet Mars. Mineral end-member samples and 187 binary mineral-mineral mixtures (mixture of two distinct minerals) are used for multivariate modeling. Eighteen diamond-mineral mixtures are used to derive Raman cross-section proxies. Mineral abundance proportions are predicted for the binary mineral-mineral mixtures with known mineralogical content to evaluate the efficacy of the two quantitative methods. Technique performance is mineral-dependent. The root-mean-square error for unseen predictions (RMSE-P) using Raman cross-section proxies ranges from ±3.2-17.0 volume%. For the multivariate models, the cross-validated root-mean-square error (RMSE-CV) ranges from ±8.8 to 26.2 volume%. Although these error estimates are not directly comparable, they provide the most accurate error estimate currently available. Different scenarios may favor the use of one or the other of the two quantitative methods. This work provides fundamental groundwork that can be applied to common rock-forming minerals on terrestrial planets including Mars. Quantification of mineral abundances aids in answering critical geologic questions related to ancient primary and altered rocks as well as planetary processes and conditions.

Key Words: Modal Mineralogy, Raman Spectroscopy, Raman Cross-Section

Introduction

There are several applications for Raman mineral quantification methodologies including laboratory and field-based measurements, and planetary exploration. Planetary scientists currently utilize Raman spectroscopy for in-situ measurements of rocks, minerals, organics, and chemicals on Mars. Similarly, scientists have proposed to use this technique on other terrestrial bodies such as Venus (Sharma et al., 2011; Clegg et al., 2014) and the Moon (Wang et al., 1995).

Raman spectroscopy is gaining popularity as a tool among other sample characterization techniques within the field of planetary science. This work presents unmixing methodologies pertinent to several geologic applications but highlights its relevance to the study of Mars mineralogy. Therefore, we utilize a sample suite applicable to primary and secondary (altered) mineralogy of that planet.

Two Raman instruments were recently deployed on the Mars 2020 *Perseverance* Rover. The Scanning Habitable Environments with Raman and Luminescence for Organics & Chemicals (SHERLOC) instrument utilizes deep ultraviolet resonance for the analysis of organics, chemicals, and surface mineralogy (Bhartia et al., 2021). SuperCam analyzes surface materials with a 532 nm laser from long ranges (Wiens et al., 2021). A third Raman spectrometer, the Raman Laser Spectrometer (RLS) on the ExoMars mission, is planned to arrive in this decade. It will use a 532 nm laser to analyze powdered samples obtained by a drill (Rull et al., 2017). The beam sizes of RLS and SHERLOC (~111 μm) are small enough to primarily probe individual mineral phases, though in some cases SHERLOC has investigated mineral mixtures (Bhartia et al., 2021; Corpolongo et al., 2023). With the larger spot size of SuperCam (>1 mm) (Wiens et al., 2016) mixed-mineral spectra are produced (Beyssac et al., 2020). Spectra of fine-grained materials such as powders and dust are particularly likely to result in mixed-mineral spectra using any of these Raman instruments. The success of data interpretation from Raman instruments on Mars relies on the availability of appropriate databases (e.g., Wang et al., 1994; Lafuente et al., 2015; Veneranda et al., 2022) and software for mineral identification at relevant scales.

In addition to identification techniques, quantification methods must also be considered because it is desirable to know not only which minerals are present, but their relative

abundances. Complex Raman mixing effects due to different signal strengths of components make quantification of mixed-mineral spectra difficult (Nasdala et al., 2004). Quantitative relationships between peak area and mineral abundance are obscured by many complicating factors: the Raman cross-section of the mineral, exciting laser frequency, crystal orientation, long-range chemical/ structural ordering in the crystal lattices and experimental factors such as effective sampling volume (Haskin et al., 1997). Except for the Raman cross-section, these factors must be held constant (e.g., laser wavelength) or averaged (e.g., large spot sizes for sample averaging) to minimize their effect on quantification. These issues must be considered for interpretation of mineral mixtures through either a theoretical method like partial least squares multivariate analysis algorithms or empirical formulae (e.g., Raman cross-section proxies, a novel method presented here). There is a lack of methodology to calculate the inherent Raman signal strength of minerals and therefore quantify mineral abundances in mixtures. In fact, Raman spectra of mineral mixtures with known quantitative modal abundances do not exist in any public databases.

To begin to address this deficiency, this work uses partial least squares multivariate analysis models and introduces a new methodology called “Raman cross-section proxies”. We test and compare the two quantitative approaches. Raman cross-sections are numerical metrics for specific Raman features that rely on Raman spectra of physical mixtures between minerals of interest and a constant reference material such as diamond. This paper explicitly deals with methods for quantification rather than identification. Inspection of Raman data or a pre-classification step is recommended for mineral identification prior to quantification.

Background

Quantification of minerals in mixtures is challenging using Raman spectroscopy because of complex mixing effects caused by independent signal strengths of different minerals (Figure 1). Figure 1 shows three examples of Raman spectra of binary mineral mixtures. Despite the fact that two of the three mixtures have 20 volume% calcite (red and yellow), the diagnostic calcite features at ~1080 cm^{-1} have different peak intensities. Similarly, although forsterite is present at 80 volume% in two of the mixtures, the intensities of its doublet around ~823 and ~855 cm^{-1} are different (red and blue). This effect is mostly due to the Raman cross-section of the other phase in each mixture (e.g., forsterite versus labradorite for the calcite mixtures) while other complicating quantification factors are minimized. Raman cross-sections highlight the unmixing problem that motivates this work.

Martian Mineralogy

Part of the impetus for this work is the interpretation of Mars data from the *Perseverance* rover, including Raman data from both the SuperCam (e.g., Clave et al., 2022; Meslin et al., 2022) and SHERLOC (e.g., Scheller et al., 2022; Corpolongo et al., 2023) instruments. Accordingly, minerals that occur on Mars were chosen for the sample suite for relevance. Martian crust and surface evolution is commonly assessed via a framework based on its degree of alteration (primary versus secondary mineralogy) (Ehlmann et al., 2014). Primary minerals consist of silicates (olivine, pyroxene, and feldspar), sulfides, and certain oxide minerals (e.g., ilmenite) (e.g., Morris et al., 2006, 2008; Rogers and Christensen, 2007; Poulet et al., 2009; Vaniman et al., 2014). Secondary minerals include other oxide minerals (e.g., hematite, goethite, and ferrihydrite), hydrated silicates including phyllosilicates, carbonates, sulfates, zeolites, chlorides, and perchlorates (e.g., Ruff, 2004; Morris et al., 2006; Hecht et al., 2009; Bishop et al.,

2009; Osterloo et al., 2008; Vaniman et al., 2014; Ehlmann et al., 2008a; 2010). The sample suite used in this work is consistent with the primary and secondary mineralogy of Jezero crater (the target of the *Perseverance* Rover) and Mars generally.

Jezero crater (~45 km diameter) contains evidence of delta deposits entering an ancient paleolake (Fassett and Head, 2005; Ehlmann et al., 2008b; Schon et al., 2012). It is located within Nili Fossae (a set of concentric graben) at the northwest edge of Isidis Basin (Goudge et al., 2015). Fassett and Head (2005) identified two inlet valleys (north and west) and one outlet valley (east) linked to Jezero crater. The fluvial activity that formed the surrounding valley networks and the Jezero paleolake led to diverse geologic units in this region. Outside the crater, the north and west fan watershed are comprised of the altered basement unit (Fe/Mg-smectite rich), mottled terrain (olivine and carbonate bearing), and the pitted capping unit (volcanic or impact melt superimposed on the basement) (Goudge et al., 2015). The crater interior contains a volcanic floor unit, a light-toned floor unit, north and west fan deposits, a marginal carbonate unit, and additional mottled terrain (Goudge et al., 2015; Horgan et al., 2020; Zastrow and Glotch, 2021). The mottled terrain and light-toned unit are stratigraphically equivalent and precede fluvial activity. Next, the fan deposits were emplaced and lastly the formation of the volcanic unit occurred (Goudge et al., 2015).

Since the arrival of the *Perseverance* Rover to Jezero crater, scientific investigations have focused on two geologic units called the Máaz and the Séítah formations (Bell et al., 2022; Liu et al., 2022; Farley et al., 2022). The Máaz formation consists primarily of plagioclase and pyroxene minerals while the Séítah formation is olivine and carbonate bearing with pyroxene and plagioclase (Liu et al., 2022; Farley et al., 2022). Carbonates, hydrated silicates, sulfates,

perchlorates, and iron oxides like hematite are present throughout Jezero in low abundances (Bell et al., 2022).

Using *Perseverance*'s SHERLOC Raman measurements in Jezero crater, Scheller et al. (2022) and Corpolongo et al. (2023) evaluated the mineralogical differences between Máaz and Séítah, described evidence of olivine carbonation, and identified sulfate and perchlorate salt bearing units. SuperCam Raman measurements also identified materials like carbonate, olivine, anhydrous Na-perchlorate, and amorphous silica in Jezero crater (Clave et al., 2022; Meslin et al., 2022).

Raman Quantification

Previous workers utilized many different techniques to analyze materials using Raman spectroscopy for terrestrial and planetary geologic applications. These methodologies include sample mineral/phase identification and classification (e.g., Griffith, 1969; Wang et al., 1994; Ishikawa and Gulick, 2013; Cochrane and Blacksberg, 2015; Lafuente et al., 2015; Carey et al., 2015a, Berlanga et al., 2019; 2022), chemical characterization (e.g., pyroxene, Wang et al., 2001; feldspar, Freeman et al., 2008; and olivine, Breitenfeld et al., 2018), and quantification (e.g., Kristova et al., 2013; Qi et al., 2023; Zarei et al., 2023). Below we provide detail on additional investigations into Raman quantification of rocks and minerals.

Haskin et al. (1997) warned of the difficulties of quantifying mineral proportions using Raman spectroscopy (e.g., crystal orientation and experimental factors). The paper presented a method that relies on the collection of multiple Raman spectra at different locations on the sample. Next, mineral identification is made for each individual spectrum. This technique is similar to point counting, a common petrographic method. Using spectra measured at 100

locations on a lunar sample (14161,7062) and a rock fragment (15273,7039), Haskin et al. (1997) identified major, minor, and accessory minerals. Ling et al. (2011) also utilized the technique for similar purposes. This location-based quantification approach is similar to using spectral rasters that are often acquired on Mars. This method is the most direct way to quantify mineral proportions in most cases.

Stopar et al. (2005) examined Raman efficiencies of natural rocks and minerals. In this work, the authors calibrated absolute radiance and measured the laser power at the target. Using the ratio of irradiance to Raman radiance, it is possible to predict the signal for a given material. This work has implications to Raman cross-section values.

Lopez-Reyes et al. (2014) examined Ca-, Fe-, Na-, and Mg-sulfate salts with specific application to the RLS instrument on ExoMars. The paper used principal component analysis, partial least squares, and artificial neural networks to test identification and quantification performance. Seventeen spectra of sulfates and artificially mixed spectra were computed as combinations of the sulfates. Tests of model performance on real samples (mixtures of anhydrite + thenardite and thenardite + $MgSO_4$, n=14) demonstrated success. This work relied on artificial spectral mixtures rather than independently measured physical samples. The high accuracy of the predictions is likely due to the fact that the different sulfate samples in the project have comparable Raman cross-sections. However, when mixing components with different Raman cross-sections, more complicated mixing behaviors would be revealed.

Carey et al. (2015b) suggested an alternate approach to this problem based on whole spectrum matching (WSM) techniques and specifically the generalized match score (MS) formulation. The algorithm describes the match score as $\mathrm{MS}(\mathrm{y}_A, \mathrm{y}_B) = (1 - \mathrm{w})\mathrm{y}_A\mathrm{y}_B - \mathrm{w}|\mathrm{y}_A - \mathrm{y}_B|$, where w is a blending parameter between 0 and 1. The overall distance between

two spectra is computed by summing MS($y_A y_B$) for all corresponding intensities, y, in the pair (Carey et al., 2015b). The procedure for mineral unmixing follows several steps including spectral preprocessing to account for inconsistent factors like the spectral background, WSM, creating a new query spectrum (zeroing intensities where the MS is below a threshold), repeating WSM for additional components, and ordering results by composite distances. This work highlights the potentially confounding effects of particle size, baseline removal, and Raman cross-section, all of which affect peak intensities.

Yaghoobi et al. (2016) used a related approach called fast non-negative orthogonal matching pursuit. It works by iteratively subtracting an identified spectrum (from reference data) from the mixture spectrum to determine the components. This technique proved successful for identifying components in proportions greater than 10% using Raman spectroscopy (Yaghoobi et al. 2016). The methodology is primarily useful for identification while rough quantification is possible.

Breitenfeld et al. (2018a) demonstrated mineral species quantification for olivine group minerals using two multivariate analysis methods, partial least squares (PLS) and least absolute shrinkage and selection operator (LASSO). These methodologies demonstrated success in the quantification of Mg-rich and Fe-rich olivine with Raman spectroscopy. Here, we use similar multivariate analysis models for predicting mineral abundance estimates rather than mineral species determination.

Finally, Veneranda et al. (2021) reported semi-quantitative Raman measurements of olivine, feldspar, and pyroxene binary (n=27) and ternary (n=5) mixtures relevant to future ExoMars measurements on Mars. The paper utilized intensity ratios to model different mixing

206 components and noted that additional spectral factors like area and full width at half maximum
207 may aid further model optimization.
208
209 **Factors Influencing Quantitative Relationships between Spectra and Abundances**
210 **Sample Properties.** Crystal orientation can influence the Raman active modes that
211 contribute to a collected spectrum (e.g., calcite; see Bhagavantam, 1940). This challenge can be
212 mitigated by using a large spot size and a powdered sample to create a spectrum that represents
213 an average of all of the crystal orientations, as done in this study.
214 Particle size (Foucher et al., 2013) and texture (Popp et al., 2002) are known to affect
215 Raman spectra of minerals. Breitenfeld et al. (2018b) examined Raman spectra of olivine,
216 pyroxene, and feldspar samples at 10 different particle size splits (smallest <25 μm and largest
217 710-1000 μm). This work found subtle Raman peak intensity differences for the different
218 particle size fractions. There is a minor Raman intensity increase from <25 μm to between 63
219 μm and 106 μm, followed by a steady decrease in Raman intensity for increasing particle sizes
220 out to 710-1000 μm.
221
222 **Experimental Factors.** When constructing a quantitative model, it is important to hold
223 as many experimental factors as possible constant (e.g., laser power and integration time). This
224 allows the dataset to be directly compared to itself. How these factors might scale when
225 comparing the spectra to outside data collected under different sampling conditions must be
226 considered. When possible, it is best to compare spectra collected under the same sampling
227 conditions (e.g., the SuperCam engineering instruments at Los Alamos and the SuperCam
228 instrument on the *Perseverance* Rover on Mars). For best results in comparing laboratory and

mission data, spectral data should be collected on an equivalent system or processed to match the application dataset.

Different excitation laser wavelengths can be utilized in Raman spectroscopy. For example, the SHERLOC and SuperCam instruments on the *Perseverance* Rover have different laser wavelengths (248.6 nm versus 532 nm laser) (Bhartia et al., 2021; Wiens et al., 2021). A disadvantage that particularly affects visible investigations is that fluorescence emission can arise (Long et al., 2002; Edwards et al., 2013). Use of an ultraviolet wavelength (as on the SHERLOC instrument) can help mitigate this problem. Alternatively, a pulsed laser with a gated detector like that used by SuperCam can capture the Raman signal and eliminate the fluorescence component (Kögler and Heilala, 2020). Fluorescence mitigation from sample materials must be considered for quantitative Raman spectral measurements. Fluorescence can be caused by biogenic organics (Wang et al., 2020), the presence of rare-earth elements (Panczer et al., 2012), poorly crystalline materials (Bishop et al., 2004), as well as other effects that can complicate Raman measurements. Further research is needed to quantify the contribution of fluorescence to Raman measurements of geologic materials.

Spectral pre-processing. Raman spectroscopy pre-processing steps are important for building prediction models. Here, baseline removal and normalization are performed prior to model construction. Carey et al. (2015c) found that AirPLS outperforms other baseline removal methods for predicting mineral class, type, group, and species. Additional testing may be useful for understanding the effect of baseline removal on specific Raman datasets. Other pre-processing techniques like squashing and smoothing can also be tested. It is important to test spectral pre-processing methods for each unique Raman dataset.

252

253 # Methods

254 ## Sample Suite Selection, Preparation, and Characterization

255 The sample suite consists of 20 mineral end-member samples (Table 1) representing
256 species that are relevant to the primary and secondary mineralogy of the planet Mars. They
257 include oxides (hematite, ilmenite), carbonates (calcite, magnesite, siderite), sulfates (anhydrite,
258 gypsum, rozenite, alunite), a nesosilicate (forsterite), inosilicates (enstatite, diopside, augite),
259 phyllosilicates (montmorillonite, nontronite, saponite, clinochlore), and tectosilicates
260 (labradorite, bytownite, chabazite). Sample characteristics including sample name, sieved
261 particle size range, locality, and supplier information are reported in Table 1. Use of powdered
262 samples mitigated orientation effects, allowing random mineral orientations to be investigated.
263 Additionally, due to particle size effects, we chose to make mineral mixtures from sieved
264 samples where the paired materials shared the same particle size fraction.

265 Each mineral end-member sample (except the magnesite sample) was analyzed using
266 electron microprobe analysis (EMPA) by Joe Boesenberg at Brown University to confirm its
267 composition. All sample compositions are accurately represented by their mineral name except
268 the rozenite sample that contains Cu with potentially a bonattite impurity. Mineral identification
269 was also confirmed from the Raman data. EMPA data were also used to accurately estimate
270 mineral densities of each sample by calculating their chemical formulas and interpolating them
271 between published density values. Knowing the density of each sample prior to making the
272 mixtures is important for the accurate measurement of mixture components by volume
273 percentage.

274

Mineral-Mineral Mixtures

This work relies on two types of binary mixtures: paired mineral-mineral mixtures and mixtures of each mineral with diamond powder. The binary mineral-mineral mixtures serve two purposes; they act as a training set for the PLS multivariate analysis models, and they aid in the testing of unmixing performance of unseen data from the Raman cross-section proxy method. A total of 187 binary mineral-mineral mixtures were made from pairings of the 20 mineral end-members. Samples were weighed into ratios of 50:50, 20:80, or 5:95 volume% depending on the Raman cross-sections of the paired phases. The ratios of components in the mixtures were customized for each pair to ensure that each mineral would have characteristic spectral features with sufficient peak areas for their individual detection.

Diamond-Mineral Mixtures

Binary diamond-mineral mixtures of synthetic diamond powder with 18 of the mineral end-members (all except hematite and ilmenite) were created to enable calculation of Raman cross-section proxies for quantitative modeling. Binary mixtures of pure minerals with a reference material (here, diamond) provide an empirical and practical method for calculating cross-section proxies that relate the Raman peak areas of minerals to a diamond standard.

Synthetic diamond (Eastwind Diamond Abrasives) was chosen as the reference material for the Raman cross-section proxies because it has a simple spectrum with a strong peak at ~1332 cm^{-1} that does not overlap with peaks in most common rock-forming minerals. When the ~1332 cm^{-1} diamond feature overlaps with diagnostic mineral features, an alternative quantification method is suggested. The particle size of the diamond powder was selected to fall

297 within the particle size range of the paired mineral sample. Synthetic diamond powder can be
298 purchased at a variety of particle size ranges at a relatively low cost.

299 To evaluate the strength of the diamond signal, we prepared several mixtures of diamond
300 and forsterite, including ratios of 20:80, 15:85, 10:90, 5:95, and 1:99 volume% (Figure 2). It is
301 apparent that the ~1332 cm^{-1} diamond peak dwarfs the Raman features in most mineral samples.
302 Therefore, this trend was used to determine a diamond to mineral volume ratio of 5:95 volume%
303 for each diamond-mineral mixture to enable easy peak-fitting for both components in each
304 mixture. In these calculations, it was generally assumed that forsterite was representative of most
305 mineral samples. This ratio resulted in spectra with the presence of Raman features for both
306 diamond and the mineral end-members.
307
308 **Raman Spectral Measurements**
309 All samples (end-member minerals, mineral-mineral mixtures, and diamond-mineral
310 mixtures) were run on a Bruker BRAVO Raman spectrometer using its dual laser system. This is
311 accomplished through the excitation of two lasers simultaneously (785 and 852 nm) with a
312 patented fluorescence mitigation strategy involving successive laser heating (Cooper et al.,
313 2014). Raman measurements utilized standard linear polarized laser light with a spot size of 2
314 mm in diameter. Each sample was scanned for 10 seconds over a range from 300-3200 cm^{-1}. The
315 laser power of the BRAVO instrument is fixed and does not exceed 100 mW, reducing the risk
316 of material alteration. Samples were stored in glass vials and were shaken prior to all Raman
317 measurement to minimize sample inhomogeneity. Each sample was run three times and the
318 spectra were averaged. Baselines were removed using the adaptive iteratively reweighted
319 penalized least-squares (AirPLS) technique (Zhang et al., 2010). AirPLS uses the sum of the

difference between the signal and the baseline to adjust weights intelligently. The smoothness parameter was set to 100. Ultimately, a single Raman spectrum associated with each physical sample (20 pure minerals, 187 binary mineral mixtures, 18 diamond-mineral mixtures) was utilized in this investigation.

Multivariate Data Analysis

We model modal mineralogy (volume%) using partial least square (PLS) multivariate modeling (Geladi and Kowalski, 1986; Wold et al., 1983, 2001). Sample characterization techniques like laser-induced breakdown spectroscopy (e.g., Tucker et al., 2010; Dyar et al., 2016a), X-ray absorption spectroscopy (e.g., Dyar et al., 2012b; 2016b), mid-infrared spectroscopy (e.g., Pan et al., 2015; Breitenfeld et al., 2021) and many other methods have utilized multivariate methods like PLS. The use of PLS in the analysis of geologic materials using Raman spectroscopy is common (e.g., Lopez-Reyes et al., 2014; Breitenfeld et al., 2018a; Bonoldi et al., 2018; Sowoidnich et al., 2023). Other workers have also explored different multivariate analysis techniques, machine learning algorithms, and the combination of multiple methods simultaneously (e.g., Ishikawa and Gulick, 2013; Carey et al., 2015a; Jahoda et al., 2021). Here, we utilize PLS and hope to explore additional avenues in future work (as in Boucher et al., 2015a; 2015b).

PLS (Geladi and Kowalski, 1986; Wold et al., 1983, 2001) predictions exploit all channels of the spectral range, assigning a coefficient to every channel associated with each metadata category (e.g., mineral volume%). For spectral data, PLS utilizes all spectral data channels and finds the channels with maximal covariance between X (feature matrix) and Y (response matrix). This is accomplished by regressing one or multiple response variables against

multiple explanatory variables (p) (Wold et al., 2001). While finding covariance, PLS reduces the p-dimensional matrix with a shrinkage penalty to q dimensions. Finally, the response on q is regressed with ordinary least squares, minimizing the residual sum of squared error (Wold et al., 2001).

The number of components for each of the PLS models were tested from 2-15. The number of components was selected that resulted in the minimum cross-validated root-mean-square error value. In this investigation, p is equivalent to the number of channels in each spectrum and q is the number of components that is chosen during the cross-validation step.

Raman Cross-Section Proxy Data Analysis

We establish the novel methodology of Raman cross-section proxies in this investigation. It is accomplished by using Raman spectral data from a suite of diamond-mineral mixtures. The goal of this approach is to use a reference material (here diamond powder) to compare all minerals of interest. Making this quantitative comparison to the reference then allows for a mixing relationship to be established for all investigated minerals. Ultimately, this permits complex mineral mixtures to be quantified.

For the Raman cross-section proxy calculations, areas were calculated for the peaks arising from the diamond feature at ~1332 cm^{-1} and also from a diagnostic feature of the other mineral end-member in the pair. Each Raman feature was peak fitted with a Lorentzian curve at the spectral resolution of the data allowing for the extraction the curve fit parameters. Resultant peak positions and areas were tabulated. The ratio of these peak areas is termed a “Raman cross-section proxy” (Equation 1). A normalization parameter is included within the equation to

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2023-9224. http://www.minsocam.org/

account for the volume percentage of diamond given that different proportions could be chosen when making a diamond-mineral mixture.

$$Raman\ Cross\ Section\ Proxy = \left(\frac{area\ of\ mineral\ feature}{area\ of\ diamond\ feature}\right) \times (\%diamond) \qquad \text{(Equation 1)}$$

A visual representation comparing the Raman spectral features of two diamond mixtures (anhydrite versus forsterite) is provided in Figure 3. With spectra normalized to the ~1332 cm^{-1} diamond feature, the diagnostic anhydrite feature (~1017 cm^{-1}) overpowers the forsterite feature (~855 cm^{-1}) resulting in a comparatively higher Raman cross-section proxy for anhydrite. Note, the areas of the diagnostic mineral feature and the diamond feature are utilized (through peak fitting) in the calculation of each Raman cross-section proxy value. There is no assumption that the area of the diamond feature is constant across different diamond-mineral mixtures.

The modal mineralogy of mixtures can theoretically be calculated for any sample that consists only of components with calculated Raman cross-section proxies. Equation 2 allows for modal mineralogy predictions where area represents Raman peak area and RCP represents Raman cross-section proxies for mixing components A, B, etc.

$$Modal\ Percentage\ of\ A = \frac{\left(\frac{area}{RCP}\right)_A}{\left(\left(\frac{area}{RCP}\right)_A + \left(\frac{area}{RCP}\right)_B + \cdots\right)} \times 100 \qquad \text{(Equation 2)}$$

Equation 1 provides the fundamental calculation for the Raman cross-section proxy, introduced in this investigation. Equation 2 is a tool that allows workers to quantify modal mineralogy of mineral mixtures with unknown abundances. The principal of the Raman cross-section proxy can be extended to other reference materials beyond diamond if necessary and at different volume ratios.

Error Analysis

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2023-9224. http://www.minsocam.org/

Root-mean-square errors (RMSE), which compare predicted versus true values by taking the square root of their summed differences, are used to evaluate accuracies of various models in three ways, using nomenclature from Dyar and Ytsma (2021). RMSE-C is the internal error of the calibrated model (typically reported in the literature). RMSE-CV is the cross-validated RMSE, in which successive randomly selected portions of the data (folds) are held out and the model is built from the remaining folds. The cross-validated accuracy is then evaluated as the average RMSE of all the fold models. The number of folds is selected to equal the square root of the total number of spectra within a given model. Finally, RMSE-P error is test RMSE of held-out data that are not used within the model (unseen data).

RMSE-C and the more robust RMSE-CV are used in this investigation to evaluate the performance of the PLS models because the datasets are not large enough to hold out test data. It is not possible to calculate an equivalent RMSE-C or RMSE-CV values for the Raman cross-section proxy method because the underlying data consists of a single Raman spectrum that cannot be internally modeled. However, RMSE-P is used in the analyses of Raman cross-section proxies by predicting the known mineral abundances of the mineral-mineral mixtures. Given that the Raman cross-section proxy values are derived from only the diamond-mineral mixtures, the mineral-mineral mixtures act as unseen data.

Dataset Size

Figure 4 summarizes the number of spectra utilized in evaluating the two mineral quantification methods. Of the total number of mineral-mineral mixture spectra (n= 187), all were used in the multivariate models and a subset was used for evaluating the performance of the Raman cross-section proxies. Differences in these datasets prohibit a perfect comparison

between these two techniques. Raman spectra without diagnostic features due to sample heating, fluorescence, or noise were eliminated. In certain cases, spectral quality was observed to diminish through the combination of mixing components (e.g., signal-to-noise ratio for samples with augite), removing the possibility of peak fitting these spectra. For the Raman cross-section proxy method, spectra with overlapping features that were unresolvable through peak fitting were also eliminated. This resulted in more spectra utilized in the multivariate models than in the Raman cross-section proxy method. Although the PLS method is tested with more spectra (more robust), a RMSE-CV value is utilized (less robust) compared to the RMSE-P value for the Raman cross-section proxy method. In the future, we plan to make additional mixtures solely for the purpose of testing these two methodologies equivalently.

Results

Raman Spectra

Figures 5 and 6 depict Raman spectra of pure non-silicate and silicate mineral end-members in order by Dana Class. Raman spectral features of the 20 mineral samples are consistent with their mineralogy based upon chemical data from the microprobe compared against published literature compositions and Raman data for those phases. For example, Raman spectral features of hematite occur at ~411 cm^{-1}, ~493 cm^{-1}, and ~612 cm^{-1} (Marshall et al., 2020). The ilmenite spectrum has a symmetric stretching feature at 685 cm^{-1} (Vennari and Williams, 2021). All of the carbonate spectra have features around ~712-739 cm^{-1} and ~1086-1095 cm^{-1} as well as additional bands (Dufresne et al., 2018). The position of the dominant sulfate feature from the symmetric stretching vibration of the SO_4 tetrahedra varies between ~989-1024 cm^{-1} (Buzgar et al., 2009; Maubec et al., 2012; Košek et al., 2020) as seen for

anhydrite, gypsum, rozenite, and alunite. The rozenite band position may be shifted due to the Cu impurity. Forsterite spectra have the diagnostic doublet at ~823 and ~855 cm^{-1} (Iishi, 1978; Chopelas, 1991; Kolesov and Geiger, 2004; McKeown et al., 2010). All of the pyroxene spectra display both the ~662-667 cm^{-1} and the ~1011-1013 cm^{-1} features (Huang et al., 2000). Phyllosilicate spectra show Raman features around ~679-706 cm^{-1} for both chlorite and smectite group samples (Wang et al., 2015). In tectosilicates, the labradorite and bytownite spectra have diagnostic features at ~508 and ~506 cm^{-1} (Freeman et al., 2008), whereas the chabazite spectrum has a feature at ~466 cm^{-1} (Tsai et al., 2021).

Although the Raman spectra of the 20 samples were collected under identical conditions, differences in the Raman intensity values for these data (e.g., carbonates versus phyllosilicates) demonstrate variations in the Raman cross-sections of the minerals. Overall, the signal intensities are greater for the non-silicate minerals compared to the silicate minerals.

Multivariate Analysis Models

Figure 7 summarizes the PLS multivariate models for each mineral using R^2, RMSE-C, and the more robust RMSE-CV values. The R^2 values derive from the calibration model. Figure 8 demonstrates the accuracy of the partial least squares multivariate predictions for individual samples for each mineral. Generally, the lower RMSE-CV values for the carbonates, sulfates, and phyllosilicates indicate better model performance compared to oxides, nesosilicate, inosilicates, and tectosilicates. However, model performance is dependent on the individual mineral species. In other words, the accuracies of the predictions of modal percentages by multivariate analysis vary between minerals within the same mineral group. For example, calcite

455 and magnesite have high accuracies with high R^2 values and low RMSE-CV values, whereas
456 siderite has a relatively lower R^2 value and higher RMSE-CV value.
457
458 **Raman Cross-Section Proxies**
459 As an alternative to the partial least squares multivariate technique, we utilize Raman
460 cross-section proxies (Table 2) for quantitative measurements of abundance based on Raman
461 spectra. Each Raman cross-section proxy corresponds to a specific Raman band feature at a
462 specific wavenumber position. The most prominent (greatest intensity) Raman feature of each
463 sample is chosen for the calculation of the Raman cross-section proxies. If that feature differs for
464 samples within the same mineral group, a different peak resulting from the same Raman mode is
465 selected. Raman cross-section proxies apply only to the Raman features from which they are
466 derived. For example, using the olivine Raman cross-section proxy from the ~855 cm^{-1} feature
467 would yield incorrect abundance if applied to the olivine doublet feature near ~815–825 cm^{-1}.
468 The highest Raman cross-section proxy values are associated with the non-silicate
469 minerals. Except for the rozenite sample, the Raman cross-section proxy values are all greater
470 than one for the non-silicate minerals. All silicate minerals have a Raman cross-section proxy
471 value of less than one. Note that Raman cross-section proxy values could not be calculated for
472 the oxide minerals hematite and ilmenite because their unusually broad Raman features (Figure
473 5) would incorrectly inflate the Raman cross-section proxy values. A more detailed discussion of
474 this issue is provided below.
475 Derived Raman cross-section proxies are utilized to unmix the modal mineralogy of the
476 mineral-mineral mixtures (Figure 9). The predictions of the mineral-mineral mixture spectra
477 provide RMSE-P values that describe the prediction accuracies for each mineral species. Calcite

478 predictions are most accurate (RMSE-P = 3.2 volume%), while siderite predictions are least
479 accurate (RMSE-P = 17.0 volume%). Although the RMSE-P values are helpful to understand the
480 prediction performance for each mineral, their accuracy is interdependent because this unmixing
481 technique relies on each mixing component (Equation 2).
482
483 **Discussion**
484 **Performance and Recommendations for Quantification Methods**
485 It is common within planetary science to use quantitative methodologies like linear-least
486 squares models (often used for mid-infrared spectroscopy) and modified gaussian models (often
487 used for visible and near-infrared spectroscopy) to quantify modal (volume) abundances of
488 planetary materials. These two techniques have accuracies of ~5–10% when applied to
489 laboratory data (e.g., Sunshine and Pieters, 1993; Feely and Christensen, 1999), respectively.
490 Depending on the mineral, accuracies of the quantitative methodologies presented in this paper
491 are comparable to or underperform compared to these popular techniques for other types of
492 spectroscopy.
493
494 **Multivariate Analysis Models.** This work uses mixtures made from 20 end-member
495 mineral samples. Although the high number of mineral end-members is useful in representing the
496 complex geology of Mars and other planetary surfaces, the size of the metadata categories then
497 requires a larger training dataset, the creation of which is time-intensive. For planetary surfaces
498 with complex mineralogy like Mars, creating multivariate models for specific geologic units
499 (e.g., Máaz and Séítah) would reduce the required number of end-members, resulting in a more

efficient model construction with likely higher accuracy. This would however also reduce the general applicability.

Figure 7 summarizes the multivariate models for each mineral using R^2, RMSE-C, and the more robust RMSE-CV values. For multivariate analysis models, performance should be evaluated using RMSE-P values or alternatively RMSE-CV values when only small datasets are available. RMSE-C values can be deceptively low and may produce overly optimistic accuracies for the model performance (e.g., calcite, magnesite, diopside, augite, and nontronite). In addition to considering RMSE values that represent absolute error, R^2 values that represent percentage error can also be considered. Here, RMSE-CV values are utilized to evaluate error compared to RMSE-P values for the Raman cross-section proxy method. RMSE-P values are more robust that RMSE-CV values. Future work (creation of an additional set of complex mineral mixtures) is needed to compare the two quantification methodologies equivalently.

Generally, there is better model performance for the carbonates, sulfates, and phyllosilicates than for the oxides, nesosilicate, inosilicates, and tectosilicates (Figure 7). Performance also varies within mineral groups, such as for calcite versus siderite or nontronite versus saponite. Although crystal structure affects where Raman features occur, spectra are also influenced by composition, which in turn affects bond strength and covalency for different minerals within the same group. Additional work is needed to investigate these relationships. Mineral mixtures from this work represent a huge advance in the number of intimate mixtures available for further study of Raman unmixing, and they may be freely loaned upon request.

Diversifying the mixtures beyond the three mixing ratios utilized in this work (50:50, 20:80, and 5:95 volume%) to fill in gaps for each mineral phase will likely improve each multivariate model. Currently, there is only one observation of each mineral pair combination

(e.g., one sample containing olivine and calcite). This prevents the model from observing multiple examples of pairs of minerals. Multiple observations of the same mineral pairings in different proportions will greatly improve model performance.

Raman Cross-Section Proxies. Raman cross-section proxies calculated in this work are empirical formulations that reflect the many factors affecting peak intensity and provide a relationship between peak area and mineral abundance in mixtures. These data form the basis for simple calculations of mineral abundances in mixtures of powdered phases and lay the groundwork for quantifying many minerals across the Dana classification system when they are present in mixtures.

Peak intensities are higher for non-silicate minerals compared to the silicate minerals as expected due to the varying contributions from ionic versus covalent bonding in their structures. The division in Raman cross-section proxy values between the non-silicate and silicate minerals (Table 2) demonstrates a relationship between Raman signal strength and the inherent mineral properties. It shows quantitatively the effect of mineral structure and chemistry on bond polarizability and in turn on the Raman cross-section. The trend of the Raman cross-section proxy values (Table 2) is likely related to the detection and quantification limits for the given minerals.

Raman cross-section proxies are not calculated for the oxide minerals hematite and ilmenite. Masking or obscuring of the diamond signal due to the oxide within the sample volume may impact relative signal strengths, and further work is needed to address this issue to improve quantification of mixing components. Additionally, hematite and ilmenite Raman cross-section proxy values would be artificially high due to the breadth of their Raman spectral features

(Figure 5). These broad features are likely due to sample heating effects (e.g., Chio et al., 2003). Future work should explore controls for the calculation of Raman cross-section proxies, such as normalizing cross-section proxies by peak full-width half maximum (FWHM), that account for broad spectral features. Additional alternative solutions include reducing the power of the laser or increasing the particle size fraction of the diamond-mineral mixture to reduce the likelihood of Raman line broadening.

Although the Raman cross-section proxies calculated here are tied to specific examples of individual mineral species, creating a more general model using intermediate compositions (e.g., mineral solid solutions) should prove useful. Accuracy will also improve when multiple Raman cross-section proxies are measured and averaged for different samples of the same mineral. Multiple Raman measurements of the diamond-mineral mixtures should also improve the accuracy of Raman cross-section proxies. These improvements will help mitigate factors that reduce the accuracy of this technique such as physical mineral clumping, sample variations, and peak fitting error.

Figure 9 provides results of unmixing the modal mineralogy of the mineral-mineral mixtures using the Raman cross-section proxies. As previously discussed, this unmixing technique relies on each mixing component (Equation 2) and therefore accuracy is interdependent. Instead of testing Raman cross-section proxies on a diverse dataset (many end-members) as is done here, testing on a greater number of samples with a fewer number of end-members will aid in the interpretation of this interdependency. As with the multivariate models, mixtures that represent specific geologic units (fewer end-members) should be considered.

Comparison of Raman Quantification Methods

Different scenarios may favor the use each of the two Raman quantitative methods compared in Figure 10. Multivariate analysis techniques like partial least squares require training data. Due to the effect of Raman cross-section on mineral mixture spectra, we argue that training sets should rely on real physical samples rather than synthetic spectra calculated through pure minerals. In the case of this work, both physical samples and spectral measurements are utilized.

Creating multivariate analysis models is a laborious process requiring both the availability of components in appropriate grain sizes for mixing, accurate calculation of density for each phase, and arduous weighing on a high-accuracy balance before spectra can be acquired. In the context of planetary exploration, a training set must also have a wide variety of minerals to encompass unexpected materials encountered during a mission. Although multivariate models like PLS are typically quick to produce, the physical sample mixtures and spectral training set it relies on are time-consuming to create. These obstacles are diminished for the Raman cross-section proxy technique because only one physical sample is needed for every end-member mineral of interest and synthetic diamond is easily purchased in a variety of grain sizes.

Another concern relating to broader applications of the multivariate analysis method is an inability to make predictions when encountering unexpected materials (i.e., peaks that arise from minerals not in the training set). For the Raman cross-section proxy method, one new diamond-mineral mixture can be made for the unexpected material. It is possible to augment the multivariate analysis method but if it relies on physical samples, it is more arduous.

A challenge of the Raman cross-section proxy method includes the requirement of a knowledgeable worker to perform the Raman peak-fitting and mineral identification. Additionally, the mixing components are dependent on each other for quantitative predictions and overlapping spectral features can also be prohibitive (e.g., Figure 4). The multivariate

592 analysis method can also utilize a wide energy range of the spectra to leverage the information
593 within multiple diagnostic Raman peaks whereas the Raman cross-section proxy relies on a
594 single mineral feature. Major benefits of both quantification techniques are their easy, fast, and
595 non-invasive application for mineral quantification.

596 With the groundwork laid by this study, future work will expand to include additional
597 end-member samples as well as to provide representation from multiple samples of the same
598 mineral species. Additionally, we will test quantification methodologies on solid samples and
599 more complex mineral mixtures to better simulate planetary surfaces. We will also undertake a
600 direct comparison (identical test datasets) between the two quantitative methods.
601

602 **Implications**

603 Quantifying mineral abundances on planetary surfaces provides information regarding
604 ancient starting materials, altered rocks, processes, and conditions that can aid in answering
605 critical geologic questions. Additionally, the ability to estimate modal mineralogy is particularly
606 important for understanding the history of water and habitable environments for planetary
607 bodies.

608 Martian crust and surface evolution is often assessed based on the degree of alteration
609 (primary versus secondary mineralogy) (Ehlmann et al., 2014). Assessing the amount and
610 distributions of primary and secondary materials can inform the degree of alteration and the
611 spatial nature of processes on the surface of the planet. Martian mineralogy records information
612 about the planet's formation history, geologic evolution, and past atmospheric properties.
613 Quantifying the modal mineralogy of rocks on Mars provides context about past conditions (e.g.,

614 fluid and atmosphere properties) and processes (e.g., weathering, aqueous alteration) that
615 changed the planet's surface.

616 In addition to Mars applications, this work is applicable to the unmixing of common
617 rock-forming minerals present on Earth and elsewhere, and to mixtures of non-geological
618 materials. This research investigates methodologies useful in the interpretation of handheld
619 Raman spectrometer data for terrestrial field work or future crewed planetary missions. Raman
620 spectroscopy is a powerful tool for providing mineral identification and chemistry information as
621 well as abundance information through the quantitative solutions presented here. The chief
622 liability of these quantitative techniques lies in the lack of available well-characterized individual
623 pure minerals suitable for making mixtures, and in the labor-intensive nature of creating those
624 mixtures. Availability of mixtures is of paramount importance in creating the fundamental data
625 needed to improve the accuracy of these methods further.
626

627 **Acknowledgments**

628 We thank Professor Alian Wang for her feedback as a member of the dissertation
629 committee that reviewed this work. We are grateful to the reviewers whose comments improved
630 the manuscript. We thank the Massachusetts Space Grant Consortium for initial funding of this
631 project, and NASA SSERVI funding from the RIS^{4}E and RISE2 nodes for subsequent support.
632

633 **Data Availability**

634 All Raman spectral data, compositional EMPA data, and quantitative model information
635 (parameters, predictions values, and errors) are archived in an external data repository on Zenodo
636 (Breitenfeld et al., 2024).

637

638 **References**

639 Bell III, J.F., Maki, J.N., Alwmark, S., Ehlmann, B.L., Fagents, S.A., Grotzinger, J.P., ... and
640 Yingling, R. (2022). Geological, multispectral, and meteorological imaging results from
641 the Mars 2020 Perseverance rover in Jezero crater. *Science Advances*, *8*(47), eabo4856.

642

643 Berlanga, G., Acosta-Maeda, T. E., Sharma, S. K., Porter, J. N., Dera, P., Shelton, H., ... &
644 Misra, A. K. (2019). Remote Raman spectroscopy of natural rocks. *Applied*
645 *Optics*, *58*(32), 8971-8980.

646

647 Berlanga, G., Williams, Q., and Temiquel, N. (2022). Convolutional Neural Networks as a Tool
648 for Raman Spectral Mineral Classification Under Low Signal, Dusty Mars
649 Conditions. *Earth and Space Science*, *9*(10), e2021EA002125.

650

651 Bhagavantam, S. (1940). Effect of crystal orientation on the Raman spectrum of calcite.
652 In *Proceedings of the Indian Academy of Sciences-Section A* (Vol. 11, No. 1, pp. 62-71).
653 New Delhi: Springer India.

654

655 Bhartia, R., Beegle, L.W., DeFlores, L., Abbey, W., Razzell Hollis, J., Uckert, K., ... and Zan, J.
656 (2021). Perseverance's scanning habitable environments with Raman and luminescence
657 for organics and chemicals (SHERLOC) investigation. Space Science Reviews, 217(4),
658 58.

659

660 Bishop, J. L., & Murad, E. (2004). Characterization of minerals and biogeochemical markers on
661 Mars: A Raman and IR spectroscopic study of montmorillonite. *Journal of Raman*
662 *Spectroscopy*, *35*(6), 480-486.
663
664 Bishop, J.L., Parente, M., Weitz, C.M., Noe Dobrea, E.Z., Roach, L.H., Murchie, S.L., ... and
665 Mustard, J.F. (2009). Mineralogy of Juventae Chasma: Sulfates in the light-toned
666 mounds, mafic minerals in the bedrock, and hydrated silica and hydroxylated ferric
667 sulfate on the plateau. *Journal of Geophysical Research: Planets*, *114*(E2).
668
669 Bonoldi, L., Frigerio, F., Di Paolo, L., Savoini, A., Barbieri, D., & Grigo, D. (2018). Organic
670 Matter Maturity Profile of a Well Case Study by Combination of Raman Spectroscopy
671 and Principal Component Analysis–Partial Least Squares Regression (PCA–PLS)
672 Chemometric Methods. *Energy & fuels*, *32*(9), 8955-8965.
673
674 Boucher, T. F., Ozanne, M. V., Carmosino, M. L., Dyar, M. D., Mahadevan, S., Breves, E. A., ...
675 & Clegg, S. M. (2015a). A study of machine learning regression methods for major
676 elemental analysis of rocks using laser-induced breakdown spectroscopy. *Spectrochimica*
677 *Acta Part B: Atomic Spectroscopy*, *107*, 1-10.
678
679 Boucher, T., Carey, C. J., Dyar, M. D., Mahadevan, S., Clegg, S., & Wiens, R. (2015b).
680 Manifold preprocessing for laser-induced breakdown spectroscopy under Mars
681 conditions. *Journal of Chemometrics*, *29*(9), 484-491.
682

683 Breitenfeld, L.B., Dyar, M.D., Carey, C.J., Tague, T.J., Wang, P., Mullen, T., and Parente, M.
684 (2018a). Predicting olivine composition using Raman spectroscopy through band shift
685 and multivariate analyses. *American Mineralogist*, *103*(11), 1827-1836.
686
687 Breitenfeld, L.B., Dyar, M.D., and Sklute, E.C. (2018b). Effect of grain size on Raman signal of
688 silicates. *XIII GeoRaman Conference*.
689
690 Breitenfeld, L. B., Rogers, A. D., Glotch, T. D., Hamilton, V. E., Christensen, P. R., Lauretta, D.
691 S., ... & DiFrancesco, N. (2021). Machine learning mid-infrared spectral models for
692 predicting modal mineralogy of CI/CM chondritic asteroids and Bennu. *Journal of*
693 *Geophysical Research: Planets*, *126*(12), e2021JE007035.
694
695 Breitenfeld, L. B., Dyar, M. D., Glotch, T. D., Rogers, A. D., & Eleazer, M. (2024). Estimating
696 Modal Mineralogy using Raman Spectroscopy: Multivariate Analysis Models and Raman
697 Cross-Section Proxies [Data set]. Zenodo. https://doi.org/10.5281/zenodo.10578499
698
699 Buzgar, N., Buzatu, A., and Sanislav, I.V. (2009). The Raman study on certain sulfates. *An.*
700 *Stiint. U. Al. I-Mat*, *55*(1), 5-23.
701
702 Carey, C., Boucher, T., Mahadevan, S., Bartholomew, P., and Dyar, M.D. (2015a). Machine
703 learning tools for mineral recognition and classification from Raman
704 spectroscopy. *Journal of Raman Spectroscopy*, *46*(10), 894-903.
705

Carey, C.J., Boucher, T., Giguere, S., Mahadevan, S., and Dyar, M.D. (2015b). Automatic whole-spectrum matching. In *Int. Joint Conf. on AI Workshop on AI in Space*.

Carey, C., Dyar, M.D., Boucher, T.F., Giguere, S., Hoff, C.M., Breitenfeld, L.B., ... and Mahadevan, S. (2015c). Baseline removal in raman spectroscopy: Optimization techniques. In *46th Annual Lunar and Planetary Science Conference* (No. 1832, p. 2464).

Chio, C.H., Sharma, S.K., Lucey, P.G., and Muenow, D.W. (2003). Effects of particle size and laser-induced heating on the Raman spectra of alpha quartz grains. *Applied spectroscopy*, *57*(7), 774-783.

Chopelas, A.J.A.M. (1991). Single crystal Raman spectra of forsterite, fayalite, and monticellite. *American Mineralogist*, *76*(7-8), 1101-1109.

Clave, E., Benzerara, K., Meslin, P.Y., Forni, O., Royer, C., Mandon, L., ... and SuperCam team. (2022). Carbonate detection with SuperCam in igneous rocks on the floor of Jezero Crater, Mars. *Journal of Geophysical Research: Planets*, e2022JE007463.

Clegg, S.M., Wiens, R., Misra, A.K., Sharma, S.K., Lambert, J., Bender, S., ... and Maurice, S. (2014). Planetary geochemical investigations using Raman and laser-induced breakdown spectroscopy. *Applied spectroscopy*, *68*(9), 925-936.

Cochrane, C. J., & Blacksberg, J. (2015). A fast classification scheme in Raman spectroscopy for the identification of mineral mixtures using a large database with correlated predictors. *IEEE Transactions on Geoscience and Remote Sensing*, *53*(8), 4259-4274.

Cooper, J. B., Wise, K. L., Jones, R. W., and Marshall, S. (2014). Sequentially Shifted Excitation Raman Spectroscopy. *Spectroscopy*, *29*(11), 38-42.

Corpolongo, A., Jakubek, R.S., Abbey, W., Asher, S.A., Baker, D., Beegle, L.W., ... and Yanchilina, A. SHERLOC Raman Mineral Class Detections of the Mars 2020 Crater Floor Campaign. Journal of Geophysical Research: Planets, e2022JE007455.

Dufresne, W.J., Rufledt, C.J., and Marshall, C.P. (2018). Raman spectroscopy of the eight natural carbonate minerals of calcite structure. *Journal of Raman Spectroscopy*, *49*(12), 1999-2007.

Dyar, M. D., Breves, E. A., Emerson, E., Bell, S. W., Nelms, M., Ozanne, M. V., ... & Woodland, A. B. (2012). Accurate determination of ferric iron in garnets by bulk Mössbauer spectroscopy and synchrotron micro-XANES. *American Mineralogist*, *97*(10), 1726-1740.

Dyar, M. D., Fassett, C. I., Giguere, S., Lepore, K., Byrne, S., Boucher, T., ... & Mahadevan, S. (2016a). Comparison of univariate and multivariate models for prediction of major and minor elements from laser-induced breakdown spectra with and without masking. *Spectrochimica Acta Part B: Atomic Spectroscopy*, *123*, 93-104.

Dyar, M. D., Breves, E. A., Gunter, M. E., Lanzirotti, A., Tucker, J. M., Carey, C. J., ... & Delaney, J. S. (2016b). Use of multivariate analysis for synchrotron micro-XANES analysis of iron valence state in amphiboles. *American Mineralogist*, *101*(5), 1171-1189.

Dyar, M. D., & Ytsma, C. R. (2021). Effect of data set size on geochemical quantification accuracy with laser-induced breakdown spectroscopy. *Spectrochimica Acta Part B: Atomic Spectroscopy*, *177*, 106073.

Edwards, H.G., Hutchinson, I.B., Ingley, R., Parnell, J., Vítek, P., and Jehlička, J. (2013). Raman spectroscopic analysis of geological and biogeological specimens of relevance to the ExoMars mission. *Astrobiology*, *13*(6), 543-549.

Ehlmann, B.L., Mustard, J.F., Murchie, S.L., Poulet, F., Bishop, J.L., Brown, A.J., ... and Wray, J.J. (2008a). Orbital identification of carbonate-bearing rocks on Mars. Science, 322(5909), 1828-1832.

Ehlmann, B.L., Mustard, J.F., Fassett, C.I., Schon, S.C., Head III, J.W., Des Marais, D.J., ... and Murchie, S.L. (2008b). Clay minerals in delta deposits and organic preservation potential on Mars. Nature Geoscience, 1(6), 355-358.

Ehlmann, B.L., Mustard, J.F., and Murchie, S.L. (2010). Geologic setting of serpentine deposits on Mars. Geophysical research letters, 37(6).

Ehlmann, B.L., and Edwards, C.S. (2014). Mineralogy of the Martian surface. Annual Review of Earth and Planetary Sciences, 42, 291-315.

Farley, K.A., Stack, K.M., Shuster, D.L., Horgan, B.H.N., Hurowitz, J.A., Tarnas, J.D., ... and Zorzano, M.P. (2022). Aqueously altered igneous rocks sampled on the floor of Jezero crater, Mars. *Science*, *377*(6614), eabo2196.

Fassett, C.I., and Head III, J.W. (2005). Fluvial sedimentary deposits on Mars: Ancient deltas in a crater lake in the Nili Fossae region. Geophysical Research Letters, 32(14).

Feely, K.C., and Christensen, P.R. (1999). Quantitative compositional analysis using thermal emission spectroscopy: Application to igneous and metamorphic rocks. Journal of Geophysical Research, 104(E10), 24195–24210.

Foucher, F., Lopez-Reyes, G., Bost, N., Rull-Perez, F., Rüßmann, P., and Westall, F. (2013). Effect of grain size distribution on Raman analyses and the consequences for in situ planetary missions. *Journal of Raman Spectroscopy*, *44*(6), 916-925.

Freeman, J.J., Wang, A., Kuebler, K.E., Jolliff, B.L., and Haskin, L.A. (2008). Characterization of natural feldspars by Raman spectroscopy for future planetary exploration. *The Canadian Mineralogist*, *46*(6), 1477-1500.

Geladi, P., and Kowalski, B.R. (1986). Partial least-squares regression: a tutorial. Analytica chimica acta, 185, 1-17.

Goudge, T.A., Mustard, J.F., Head, J.W., Fassett, C.I., and Wiseman, S.M. (2015). Assessing the mineralogy of the watershed and fan deposits of the Jezero crater paleolake system, Mars. Journal of Geophysical Research: Planets, 120(4), 775-808.

Griffith, W.P. (1969). Raman spectroscopy of minerals. *Nature*, *224*(5216), 264-266.

Haskin, L.A., Wang, A., Rockow, K.M., Jolliff, B.L., Korotev, R.L., and Viskupic, K.M. (1997). Raman spectroscopy for mineral identification and quantification for in situ planetary surface analysis: A point count method. Journal of Geophysical Research: Planets, 102(E8), 19293-19306.

Hecht, M.H., Kounaves, S.P., Quinn, R.C., West, S.J., Young, S.M., Ming, D.W., ... and Smith, P.H. (2009). Detection of perchlorate and the soluble chemistry of martian soil at the Phoenix lander site. Science, 325(5936), 64-67.

Horgan, B.H., Anderson, R.B., Dromart, G., Amador, E.S., and Rice, M.S. (2020). The mineral diversity of Jezero crater: Evidence for possible lacustrine carbonates on Mars. Icarus, 339, 113526.

819 Huang, E., Chen, C.H., Huang, T., Lin, E.H., and Xu, J.A. (2000). Raman spectroscopic
820 characteristics of Mg-Fe-Ca pyroxenes. *American Mineralogist*, *85*(3-4), 473-479.
821
822 Iishi, K. (1978). Lattice dynamics of forsterite. *American Mineralogist*, *63*(11-12), 1198-1208.
823
824 Ishikawa, S. T., & Gulick, V. C. (2013). An automated mineral classifier using Raman
825 spectra. *Computers & geosciences*, *54*, 259-268.
826
827 Jahoda, P., Drozdovskiy, I., Payler, S. J., Turchi, L., Bessone, L., & Sauro, F. (2021). Machine
828 learning for recognizing minerals from multispectral data. *Analyst*, *146*(1), 184-195.
829
830 Kögler, M., and Heilala, B. (2020). Time-gated Raman spectroscopy–a review. *Measurement*
831 *Science and Technology*, *32*(1), 012002.
832
833 Kolesov, B.A., and Geiger, C.A. (2004). A Raman spectroscopic study of Fe–Mg
834 olivines. *Physics and Chemistry of Minerals*, *31*, 142-154.
835
836 Košek, F., Culka, A., Fornasini, L., Vandenabeele, P., Rousaki, A., Mirao, J., ... and Jehlička, J.
837 (2020). Application of a handheld Raman spectrometer for the screening of colored
838 secondary sulfates in abandoned mining areas—The case of the São Domingos Mine
839 (Iberian Pyrite Belt). *Journal of Raman Spectroscopy*, *51*(7), 1186-1199.
840

Kristova, P., Hopkinson, L., Rutt, K., Hunter, H., & Cressey, G. (2013). Quantitative analyses of powdered multi-minerallic carbonate aggregates using a portable Raman spectrometer. *American Mineralogist*, *98*(2-3), 401-409.

Ling, Z., Wang, A., and Jolliff, B.L. (2011). A systematic spectroscopic study of four Apollo lunar soils. *Journal of Earth Science*, *22*(5), 578.

Liu, Y., Tice, M.M., Schmidt, M.E., Treiman, A.H., Kizovski, T.V., Hurowitz, J.A., ... and Zorzano, M.P. (2022). An olivine cumulate outcrop on the floor of Jezero crater, Mars. *Science*, *377*(6614), 1513-1519.

Long, D.A., and Long, D.A. (2002). *The Raman effect: a unified treatment of the theory of Raman scattering by molecules* (Vol. 8, pp. 31-48). Chichester: Wiley.

Lopez-Reyes, G., Sobron, P., Lefebvre, C., and Rull, F. (2014). 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. American Mineralogist, 99(8-9), 1570-1579.

Marshall, C.P., Dufresne, W.J., and Rufledt, C.J. (2020). Polarized Raman spectra of hematite and assignment of external modes. *Journal of Raman Spectroscopy*, *51*(9), 1522-1529.

Maubec, N., Lahfid, A., Lerouge, C., Wille, G., and Michel, K. (2012). Characterization of alunite supergroup minerals by Raman spectroscopy. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, *96*, 925-939.

Meslin, P.Y., Forni, O., Beck, P., Cousin, A., Beyssac, O., Lopez-Reyes, G., ... and Zorzano, M.P. (2022, March). Evidence for perchlorate and sulfate salts in jezero crater, mars, from supercam observations. In *Lunar and Planetary Science Conference* (Vol. 53, p. 2694).

Morris, R.V., Klingelhoefer, G., Schröder, C., Rodionov, D.S., Yen, A., Ming, D.W., ... and Arvidson, R.E. (2006). Mössbauer mineralogy of rock, soil, and dust at Gusev crater, Mars: Spirit's journey through weakly altered olivine basalt on the plains and pervasively altered basalt in the Columbia Hills. Journal of Geophysical Research: Planets, 111(E2).

Morris, R.V., Klingelhoefer, G., Schröder, C., Fleischer, I., Ming, D.W., Yen, A.S., ... and Squyres, S.W. (2008). Iron mineralogy and aqueous alteration from Husband Hill through Home Plate at Gusev crater, Mars: Results from the Mössbauer instrument on the Spirit Mars Exploration Rover. Journal of Geophysical Research: Planets, 113(E12).

Lafuente, B., Downs, R.T., Yang, H., Stone, N. (2015) The power of databases: the RRUFF project. In: Highlights in Mineralogical Crystallography, T Armbruster and R M Danisi, eds. Berlin, Germany, W. De Gruyter, pp 1-30.

McKeown, D.A., Bell, M.I., and Caracas, R. (2010). Theoretical determination of the Raman spectra of single-crystal forsterite (Mg2SiO4). *American Mineralogist*, *95*(7), 980-986.

Nasdala, L., Smith, D.C., Kaindl, R., Ziemann, M.A., Beran, A., and Libowitzky, E. (2004). Raman spectroscopy: analytical perspectives in mineralogical research. Spectroscopic methods in mineralogy, 6, 281-343.

Osterloo, M.M., Hamilton, V.E., Bandfield, J.L., Glotch, T.D., Baldridge, A.M., Christensen, P.R., ... and Anderson, F.S. (2008). Chloride-bearing materials in the southern highlands of Mars. *Science*, *319*(5870), 1651-1654.

Pan, C., Rogers, A. D., & Thorpe, M. T. (2015). Quantitative compositional analysis of sedimentary materials using thermal emission spectroscopy: 2. Application to compacted fine-grained mineral mixtures and assessment of applicability of partial least squares methods. *Journal of Geophysical Research: Planets*, *120*(11), 1984-2001.

Panczer, G., De Ligny, D., Mendoza, C., Gaft, M., Seydoux-Guillaume, A. M., Wang, X. I. A. O. C. H. U. N., ... & Rull, F. (2012). Raman and fluorescence. *EMU Notes in Mineralogy*, *12*(2), 61-82.

Popp, J., Tarcea, N., Kiefer, W., Hilchenbach, M., Thomas, N., Stuffler, T., ... and Greshake, A. (2002). The effect of surface texture on the mineralogical analysis of chondritic meteorites using Raman spectroscopy. *Planetary and Space Science*, *50*(9), 865-870.

Poulet, F., Mangold, N., Platevoet, B., Bardintzeff, J.M., Sautter, V., Mustard, J.F., ... and Aléon-Toppani, A. (2009). Quantitative compositional analysis of Martian mafic regions using the MEx/OMEGA reflectance data: 2. Petrological implications. Icarus, 201(1), 84-101.

Qi, X., Ling, Z., Liu, P., Chen, J., Cao, H., Liu, C., ... & Liu, Y. (2023). Quantitative mineralogy of planetary silicate ternary mixtures using Raman spectroscopy. *Earth and Space Science*, *10*(5), e2023EA002825.

Rogers, A.D., and Christensen, P.R. (2007). Surface mineralogy of Martian low-albedo regions from MGS-TES data: Implications for upper crustal evolution and surface alteration. Journal of Geophysical Research: Planets, 112(E1).

Ruff, S.W. (2004). Spectral evidence for zeolite in the dust on Mars. *Icarus*, *168*(1), 131-143.

Rull, F., Maurice, S., Hutchinson, I., Moral, A., Perez, C., Diaz, C., ... and Vago, J.L. (2017). The Raman laser spectrometer for the ExoMars rover mission to Mars. Astrobiology, 17(6-7).

Scheller, E.L., Razzell Hollis, J., Cardarelli, E.L., Steele, A., Beegle, L.W., Bhartia, R., ... and Zorzano, M.P. (2022). Aqueous alteration processes in Jezero crater, Mars—Implications for organic geochemistry. *Science*, *378*(6624), 1105-1110.

Schon, S.C., Head, J.W., and Fassett, C.I. (2012). An overfilled lacustrine system and progradational delta in Jezero crater, Mars: Implications for Noachian climate. Planetary and Space Science, 67(1), 28-45.

Sharma, S.K., Misra, A.K., Clegg, S.M., Barefield, J.E., Wiens, R.C., Acosta, T.E., and Bates, D.E. (2011). Remote-Raman spectroscopic study of minerals under supercritical CO2 relevant to Venus exploration. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, *80*(1), 75-81.

Sowoidnich, K., Maiwald, M., Ostermann, M., & Sumpf, B. (2023). Shifted excitation Raman difference spectroscopy for soil component identification and soil carbonate determination in the presence of strong fluorescence interference. *Journal of Raman spectroscopy*.

Stopar, J. D., Lucey, P. G., Sharma, S. K., Misra, A. K., Taylor, G. J., & Hubble, H. W. (2005). Raman efficiencies of natural rocks and minerals: Performance of a remote Raman system for planetary exploration at a distance of 10 meters. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, *61*(10), 2315-2323.

Sunshine, J.M., and Pieters, C.M. (1993). Estimating modal abundances from the spectra of natural and laboratory pyroxene mixtures using the modified Gaussian model. *Journal of Geophysical Research: Planets*, *98*(E5), 9075-9087.

Tsai, Y.L., Huang, E., Li, Y.H., Hung, H.T., Jiang, J.H., Liu, T.C., ... and Chen, H.F. (2021). Raman spectroscopic characteristics of zeolite group minerals. *Minerals*, *11*(2), 167.

Tucker, J. M., Dyar, M. D., Schaefer, M. W., Clegg, S. M., & Wiens, R. C. (2010). Optimization of laser-induced breakdown spectroscopy for rapid geochemical analysis. *Chemical Geology*, *277*(1-2), 137-148.

Vaniman, D.T., Bish, D.L., Ming, D.W., Bristow, T.F., Morris, R.V., Blake, D.F., ... and Freissinet, C. (2014). Mineralogy of a mudstone at Yellowknife Bay, Gale crater, Mars. *science*, *343*(6169), 1243480.

Veneranda, M., Manrique-Martinez, J.A., Garcia-Prieto, C., Sanz-Arranz, A., Saiz, J., Lalla, E., ... and Lopez-Reyes, G. (2021). Raman semi-quantification on Mars: ExoMars RLS system as a tool to better comprehend the geological evolution of martian crust. *Icarus*, *367*, 114542.

Veneranda, M., Sanz-Arranz, A., Manrique, J.A., Saiz, J., Garcia-Prieto, C., Pascual-Sánchez, E., ... and Lopez-Reyes, G. (2022). Analytical database of Martian minerals (ADaMM): Project synopsis and Raman data overview. *Journal of Raman Spectroscopy*, *53*(3), 364-381.

977 Vennari, C.E., and Williams, Q. (2021). A high-pressure Raman study of FeTiO3 ilmenite:
978 Fermi resonance as a manifestation of Fe-Ti charge transfer. *Physics and Chemistry of*
979 *Minerals*, *48*(9), 34.
980
981 Wang, A., Han, J., Guo, L., Yu, J., and Zeng, P. (1994). Database of standard Raman spectra of
982 minerals and related inorganic crystals. Applied Spectroscopy, 48(8), 959-968.
983
984 Wang, A., Jolliff, B.L., and Haskin, L.A. (1995). Raman spectroscopy as a method for mineral
985 identification on lunar robotic exploration missions. *Journal of Geophysical Research:*
986 *Planets*, *100*(E10), 21189-21199.
987
988 Wang, A., Jolliff, B.L., Haskin, L.A., Kuebler, K.E., and Viskupic, K.M. (2001).
989 Characterization and comparison of structural and compositional features of planetary
990 quadrilateral pyroxenes by Raman spectroscopy. American Mineralogist, 86(7-8), 790-
991 806.
992
993 Wang, A., Freeman, J.J., and Jolliff, B.L. (2015). Understanding the Raman spectral features of
994 phyllosilicates. *Journal of Raman Spectroscopy*, *46*(10), 829-845.
995
996 Wang, A., Wei, J., and Korotev, R.L. (2020). Quantification of fluorescence emission from
997 extraterrestrial materials and its significance for planetary Raman spectroscopy. *Journal*
998 *of Raman Spectroscopy*, *51*(9), 1636-1651.
999

Wiens, R.C., Maurice, S., Robinson, S.H., Nelson, A.E., Cais, P., Bernardi, P., ... and Willis, P. (2021). The SuperCam instrument suite on the NASA Mars 2020 rover: Body unit and combined system tests. Space Science Reviews, 217, 1-87.

Wold, S., Martens, H., and Wold, H. (1983). The multivariate calibration problem in chemistry solved by the PLS method. Lecture Notes in Mathematics, 973, 286–293.

Wold, S., Sjöström, M., and Eriksson, L. (2001). PLS-regression: a basic tool of chemometrics. Chemometrics and intelligent laboratory systems, 58(2), 109-130.

Yaghoobi, M., Wu, D., Clewes, R.J., and Davies, M.E. (2016, October). Fast sparse Raman spectral unmixing for chemical fingerprinting and quantification. In Optics and Photonics for Counterterrorism, Crime Fighting, and Defence XII (Vol. 9995, p. 99950E). International Society for Optics and Photonics.

Zarei, M., Solomatova, N. V., Aghaei, H., Rothwell, A., Wiens, J., Melo, L., ... & Grant, E. (2023). Machine Learning Analysis of Raman Spectra To Quantify the Organic Constituents in Complex Organic–Mineral Mixtures. *Analytical Chemistry*, *95*(43), 15908-15916.

Zastrow, A.M., and Glotch, T.D. (2021). Distinct carbonate lithologies in Jezero crater, Mars. *Geophysical Research Letters*, *48*(9), e2020GL092365.

Zhang, Z.M., Chen, S., and Liang, Y.Z. (2010). Baseline correction using adaptive iteratively reweighted penalized least squares. Analyst, 135(5), 1138-1146.

List of Figure Captions

Figure 1. Raman spectra of three binary mixtures consisting of forsterite, calcite, labradorite, and diopside. Raman intensity variations arise because the Raman cross-sections (numerical metrics associated with inherent signal strength) of these materials all vary. Spectra were collected on Bruker's BRAVO handheld spectrometer (dual 785 and 852 nm excitation lasers) with three sample scans for 10 s of integration time over the wavenumber range of 300-3200 cm^{-1}.

Figure 2. Raman spectra of mixtures of olivine and diamond in variable volume percentages (e.g., 95 vol% olivine, 5 vol% diamond) distinguished by color.

Figure 3. Normalized Raman spectra of mixtures of anhydrite with diamond (red) and forsterite with diamond (blue) at 95:5 volume ratios. The dominant ~1017 cm^{-1} anhydrite feature compared to the ~855 cm^{-1} forsterite feature indicates a higher Raman cross-section proxy.

Figure 4. Number of spectra utilized in multivariate analysis models (green) and in assessing the Raman cross-section proxy error calculations (purple).

Figure 5. Raman spectra of non-silicate mineral end-members including oxides (top), carbonates (middle), and sulfates (bottom).

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2023-9224. http://www.minsocam.org/

Figure 6. Raman spectra of silicate mineral end-members including nesosilicates/ inosilicates (top), phyllosilicates (middle), and tectosilicates (bottom).

Figure 7. RMSE-CV (top), RMSE-C (middle), and R^2 values (bottom) for partial least squares multivariate analysis for each mineral.

Figure 8. Partial least squares multivariate predictions versus actual abundances (volume%) for each prediction variable.

Figure 9. Raman cross-section proxy RMSE-P values (top left) and predictions versus actual abundances (volume%) for each prediction variable. RMSE-P calculations were made using mineral-mineral mixtures.

Figure 10. Pros (+, green) and cons (–, red) of the Raman cross-section proxy and the multivariate analysis technique.

Tables

Table 1. Sample suite of twenty end-member minerals including sample names, particle size ranges, and localities.

Dana Class	Mineral Species	Ideal Chemical Formula	Sample Name	Particle Size (μm)	Locality	Supplier
4	hematite	Fe_2O_3	ICOSA-11	25-45	Custer County, South Dakota, USA	Ward's
	ilmenite	$(Fe,Ti)_2O_3$	ICOSA-12	25-45	Unknown	Mount Holyoke College
14	calcite	$CaCO_3$	ICOSA-5	25-45	Rossie, New York, USA	Smithsonian Institution
	magnesite	$MgCO_3$	ICOSA-16	25-45	Pomba Pit, Serra das Eguas, Brumado, Bahia, Brazil	Mount Holyoke College
	siderite	$FeCO_3$	ICOSA-21	25-45	Unknown	Mount Holyoke College
28	anhydrite	$CaSO_4$	ICOSA-2	25-45	Naica, Mexico	Metropolis Family
29	gypsum	$CaSO_4•2(H_2O)$	ICOSA-10	25-45	Naica, Mexico	Rock and Bone Mineral Shop

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2023-9224. http://www.minsocam.org/

	rozenite	$Fe^{2+}SO_4 \cdot 4(H_2O)$	ICOSA-15	25-45	Tulsequah Chief Mine, Mount Eaton, BBC, Canada	Mount Holyoke College
30	alunite	$KAl_3(SO_4)_2(OH)_6$	ICOSA-1	25-45	Synthetic	Mount Holyoke College
51	forsterite	$(Mg,Fe)_2SiO_4$	ICOSA-9	38-63	San Carlos, Arizona, USA	Mount Holyoke College
65	enstatite	$Mg_2Si_2O_6$	ICOSA-8	25-45	Kilosa, Tanzania	Mount Holyoke College
	diopside	$MgCaSi_2O_6$	ICOSA-7	38-63	Herschel, Ontario, Canada	Mount Holyoke College
	augite	$(Ca,Na)(Mg,Fe,Al,Ti)(Si,Al)_2O_6$	ICOSA-3	25-45	Harcourt, Ontario, Canada	Mount Holyoke College
71	montmorillonite	$(Na,Ca)_{0.33}(Al,Mg)_2(Si_4O_{10})(OH)_2 \cdot n(H_2O)$	ICOSA-18	25-45	SCa-2	Clay Mineral Society
	nontronite	$Na_{0.3}Fe^{3+}_2(Si,Al)_4O_{10}(OH)_2 \cdot n(H_2O)$	ICOSA-19	25-45	NAu-2	Clay Mineral Society
	saponite	$Ca_{0.25}(Mg,Fe)_3(Si,Al)_4O_{10}(OH)_2 \cdot n(H_2O)$	ICOSA-20	25-45	Bumo Creek, Arizona, USA	Janice Bishop
	clinochlore	$(Mg,Fe^{2+})_5Al(Si_3Al)O_{10}(OH)_8$	ICOSA-6	25-45	CCa-2	Clay Mineral Society
76	labradorite	$(Ca,Na)(Si,Al)_4O_8$	ICOSA-14	25-45	Chihuahua, Mexico	Mount Holyoke College
	bytownite	$(Ca,Na)(Si,Al)_4O_8$	ICOSA-4	25-45	Crystal Bay, Minnesota, USA	Mount Holyoke College
77	chabazite	$(Ca,Na_2,K_2)_2[Al_4Si_8O_{24}] \cdot 12H_2O$	ICOSA-24	25-45	Wasson Bluff, Parrsboro, Nova Scotia, Canada	Rock and Bone Mineral Shop

Table 2. Raman cross-section proxy values for each mineral sample.

Dana Class	Mineral Species	Sample Name	Diamond Percentage	Raman Band Position (cm^{-1})	Raman Cross-Section Proxies
14	calcite	DM-25	5	1086.21 ± 0.01	2.00
	magnesite	DM-36	5	1094.50 ± 2.39e-3	2.33
	siderite	DM-41	5	1085.97 ± 4.14e-3	6.06
28	anhydrite	DM-22	5	1017.01 ± 3.83e-3	2.41
29	gypsum	DM-30	5	1007.87 ± 4.24e-3	1.69
	rozenite	DM-35	5	1024.93 ± 0.05	0.40
30	alunite	DM-21	5	989.48 ± 0.01	1.07
51	forsterite	DM-29	5	854.85 ± 0.01	0.80
65	enstatite	DM-28	5	1009.06 ± 0.02	0.19
	diopside	DM-27	5	1012.10 ± 0.01	0.49
	augite	DM-23	5	1012.25 ± 0.01	0.61
71	montmorillonite	DM-38	5	706.51 ± 0.01	0.20
	nontronite	DM-39	5	688.88 ± 0.02	0.16
	saponite	DM-40	5	683.73 ± 0.01	0.24
	clinochlore	DM-26	5	683.50 ± 0.03	0.47
76	labradorite	DM-34	5	509.02 ± 0.01	0.57
	bytownite	DM-24	5	505.00 ± 0.01	0.39
77	chabazite	DM-44	5	466.79 ± 0.05	0.61

Figures

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2023-9224. http://www.minsocam.org/

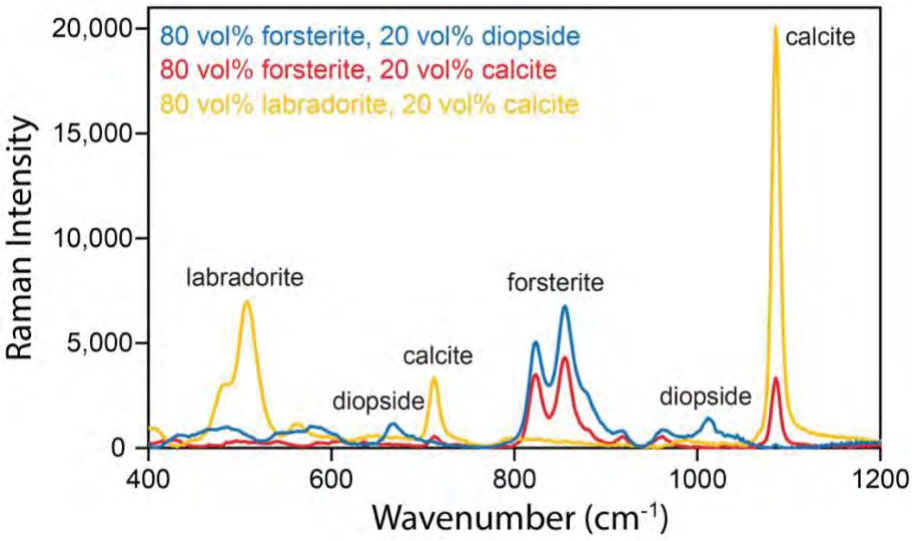


Figure 1. Raman spectra of three binary mixtures consisting of forsterite, calcite, labradorite, and diopside. Raman intensity variations arise because the Raman cross-sections (numerical metrics associated with inherent signal strength) of these materials all vary. Spectra were collected on Bruker's BRAVO handheld spectrometer (dual 785 and 852 nm excitation lasers) with three sample scans for 10 s of integration time over the wavenumber range of 300-3200 cm^{-1}.

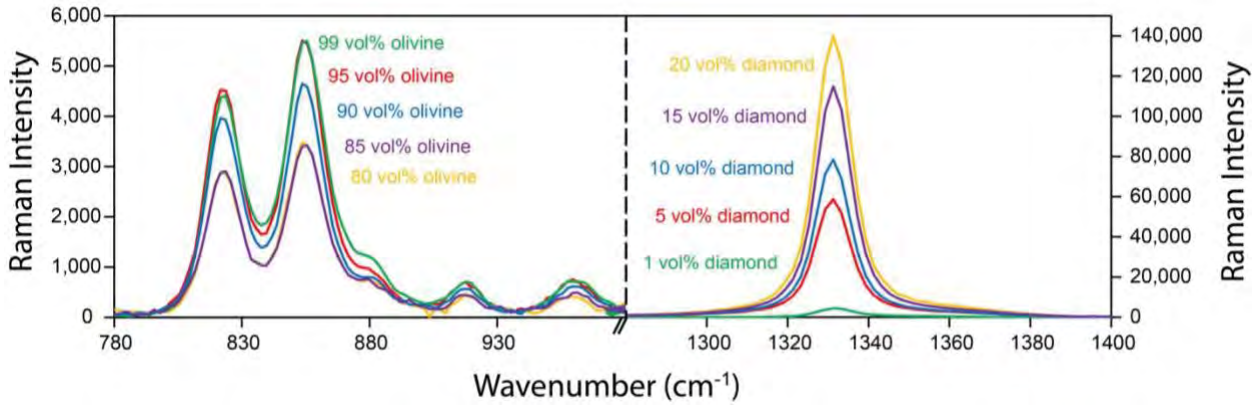


Figure 2. Raman spectra of mixtures of olivine and diamond in variable volume percentages (e.g., 95 vol% olivine, 5 vol% diamond) distinguished by color.

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2023-9224. http://www.minsocam.org/

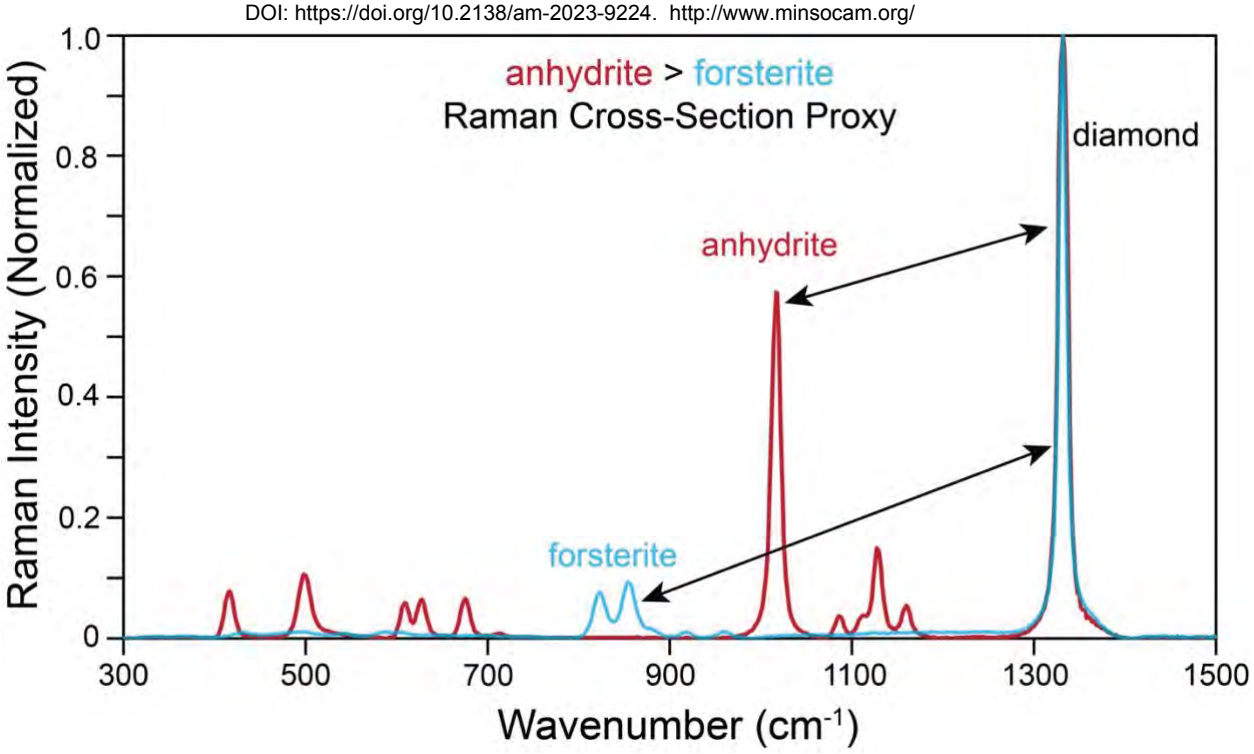


Figure 3. Normalized Raman spectra of mixtures of anhydrite with diamond (red) and forsterite with diamond (blue) at 95:5 volume ratios. The dominant ~1017 cm^{-1} anhydrite feature compared to the ~855 cm^{-1} forsterite feature indicates a higher Raman cross-section proxy.

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2023-9224. http://www.minsocam.org/

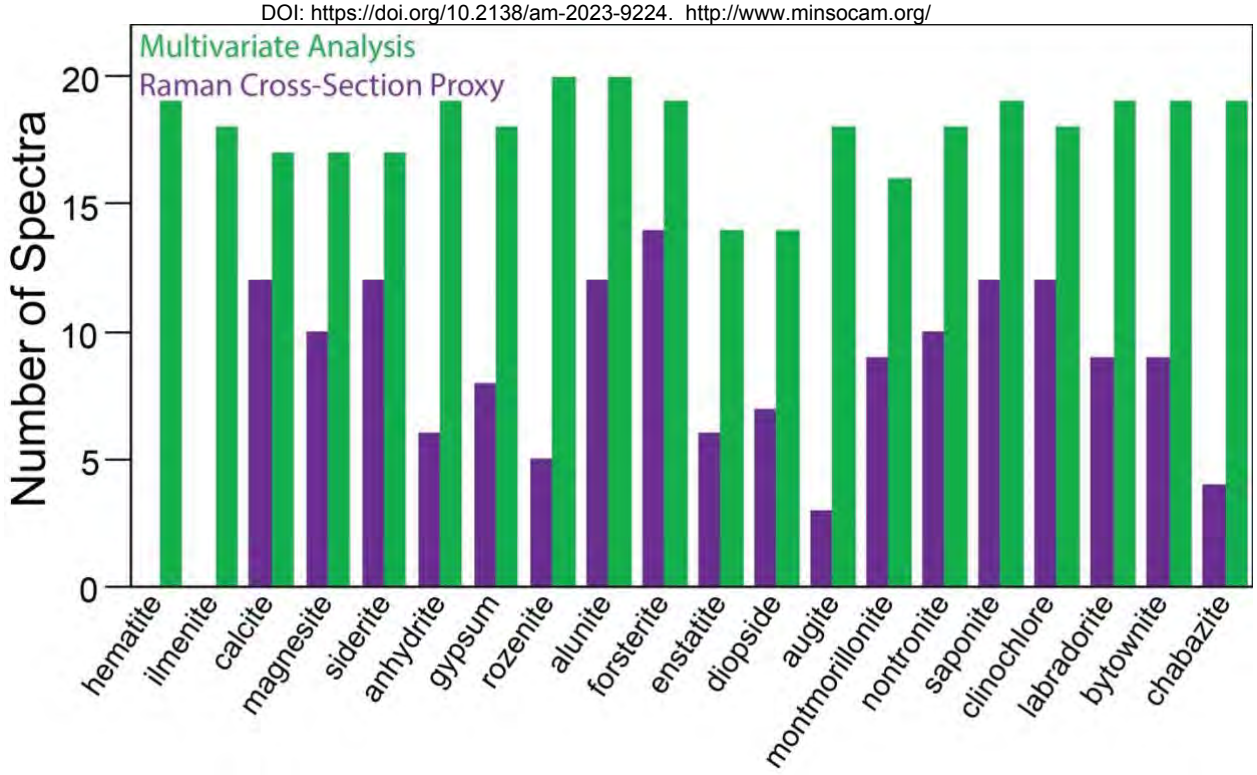


Figure 4. Number of spectra utilized in multivariate analysis models (green) and in assessing the Raman cross-section proxy error calculations (purple).

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2023-9224. http://www.minsocam.org/

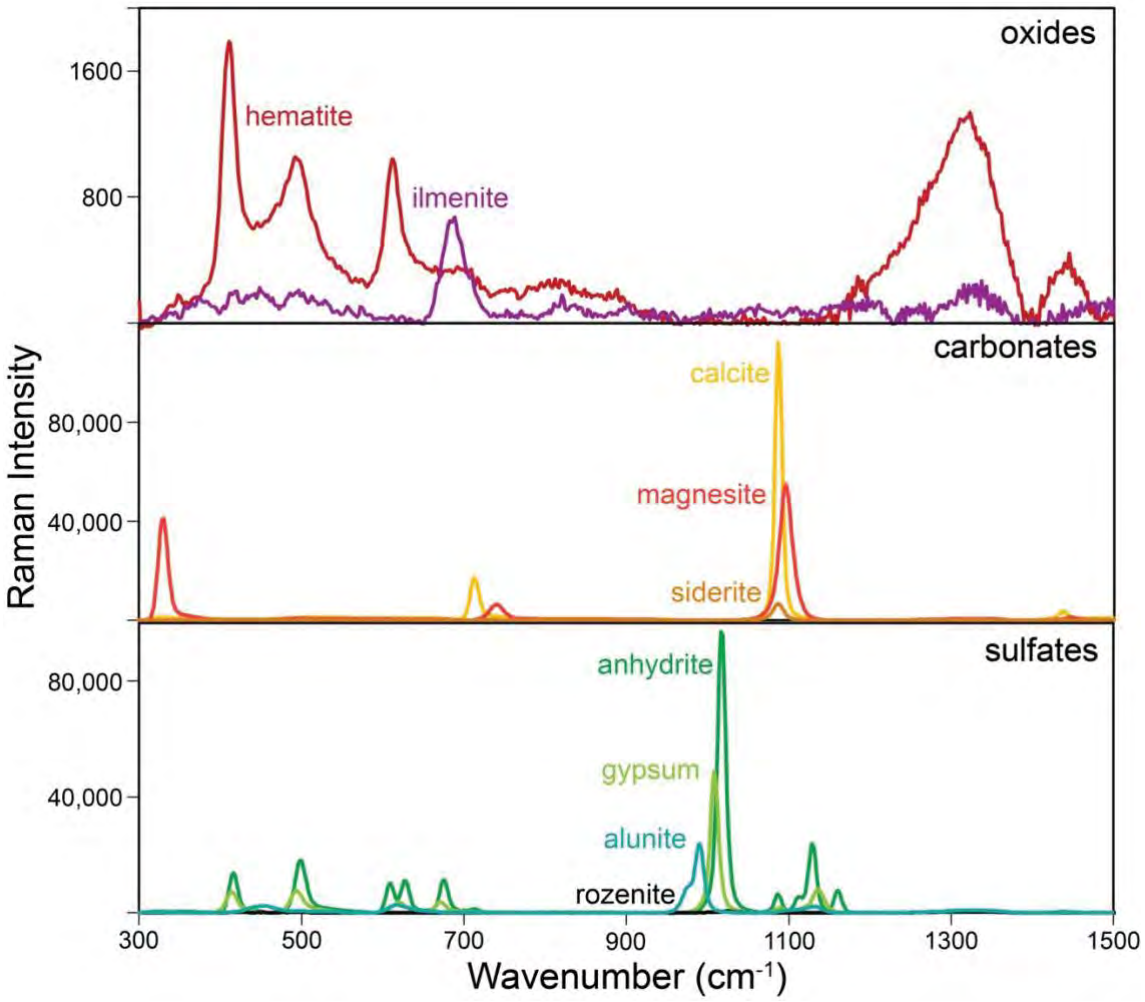
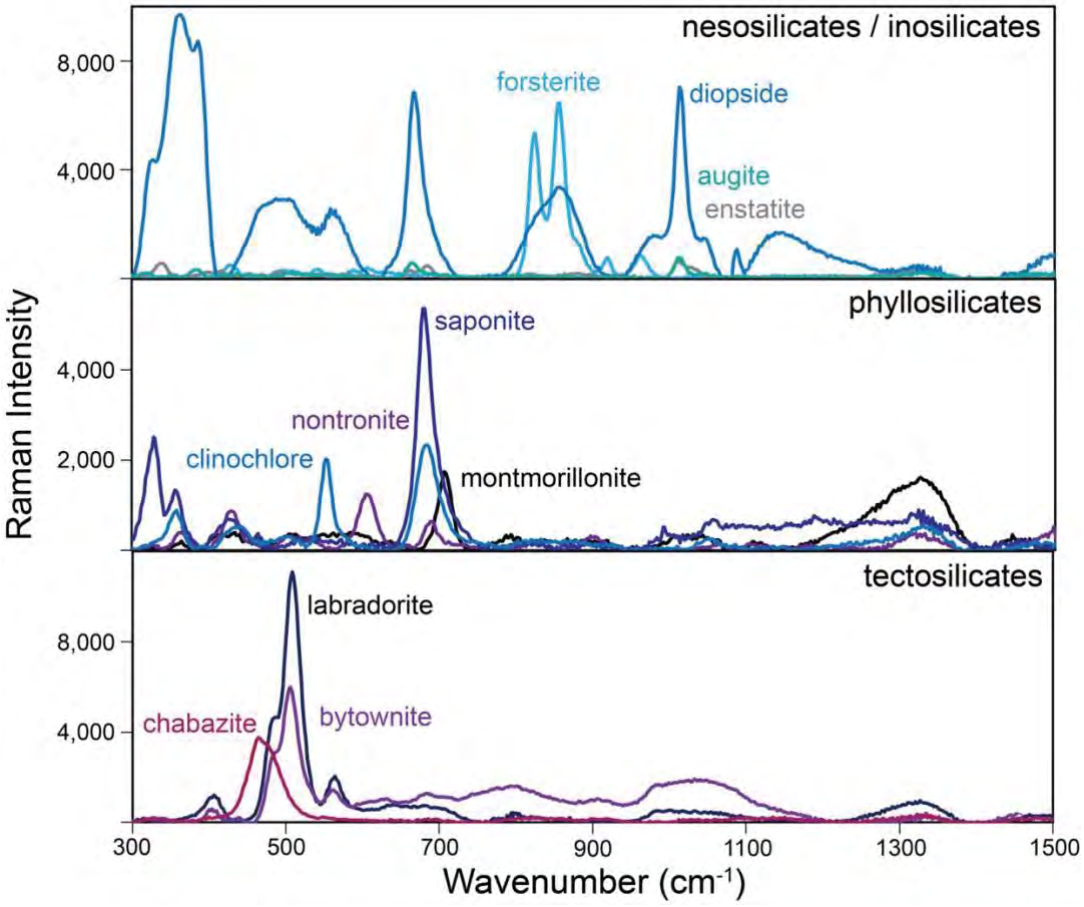


Figure 5. Raman spectra of non-silicate mineral end-members including oxides (top), carbonates (middle), and sulfates (bottom).

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2023-9224. http://www.minsocam.org/



1091

1092 Figure 6. Raman spectra of silicate mineral end-members including nesosilicates/ inosilicates

1093 (top), phyllosilicates (middle), and tectosilicates (bottom).

1094

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2023-9224. http://www.minsocam.org/

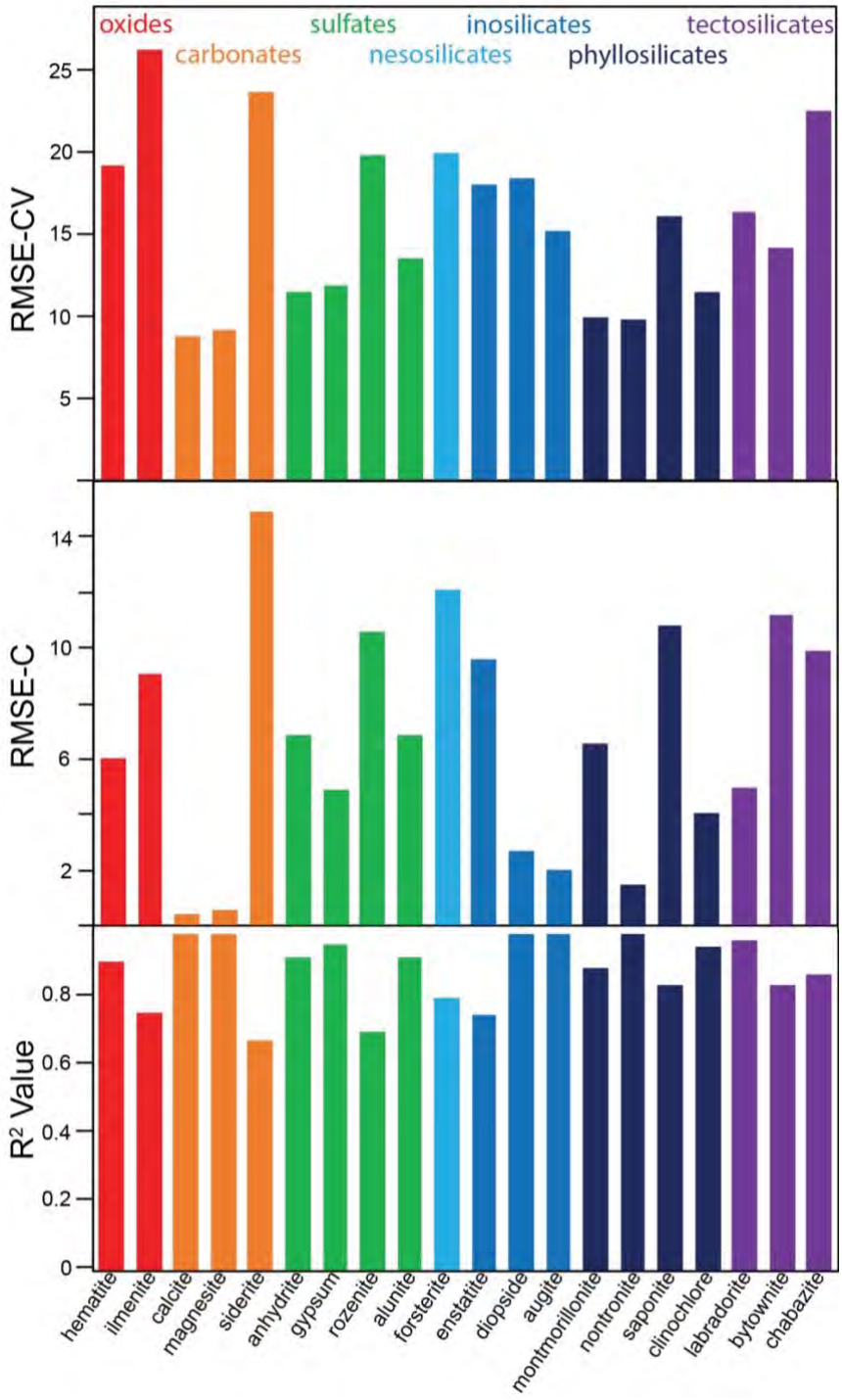


Figure 7. RMSE-CV (top), RMSE-C (middle), and R^2 values (bottom) for partial least squares multivariate analysis for each mineral.

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2023-9224. http://www.minsocam.org/

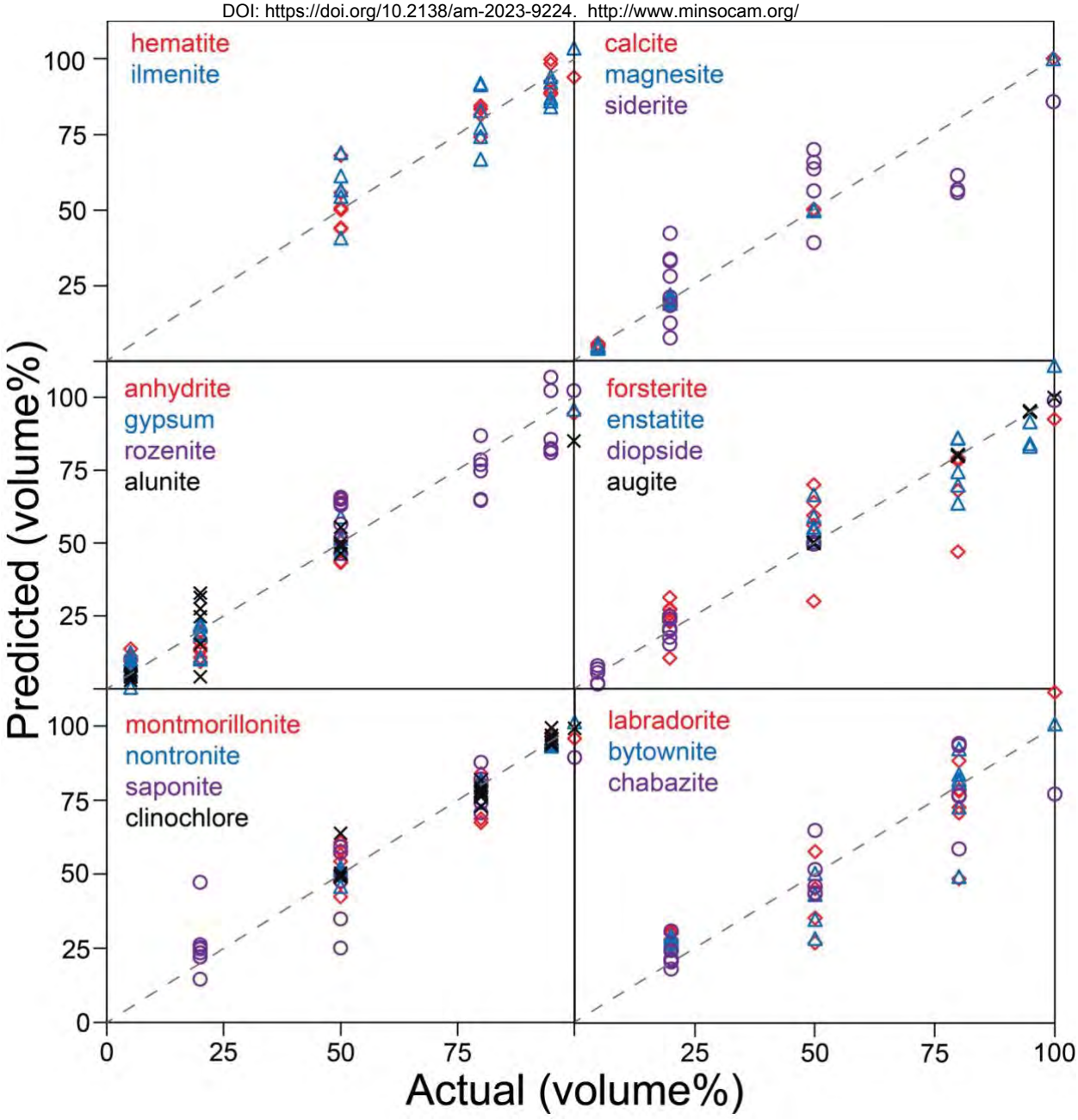
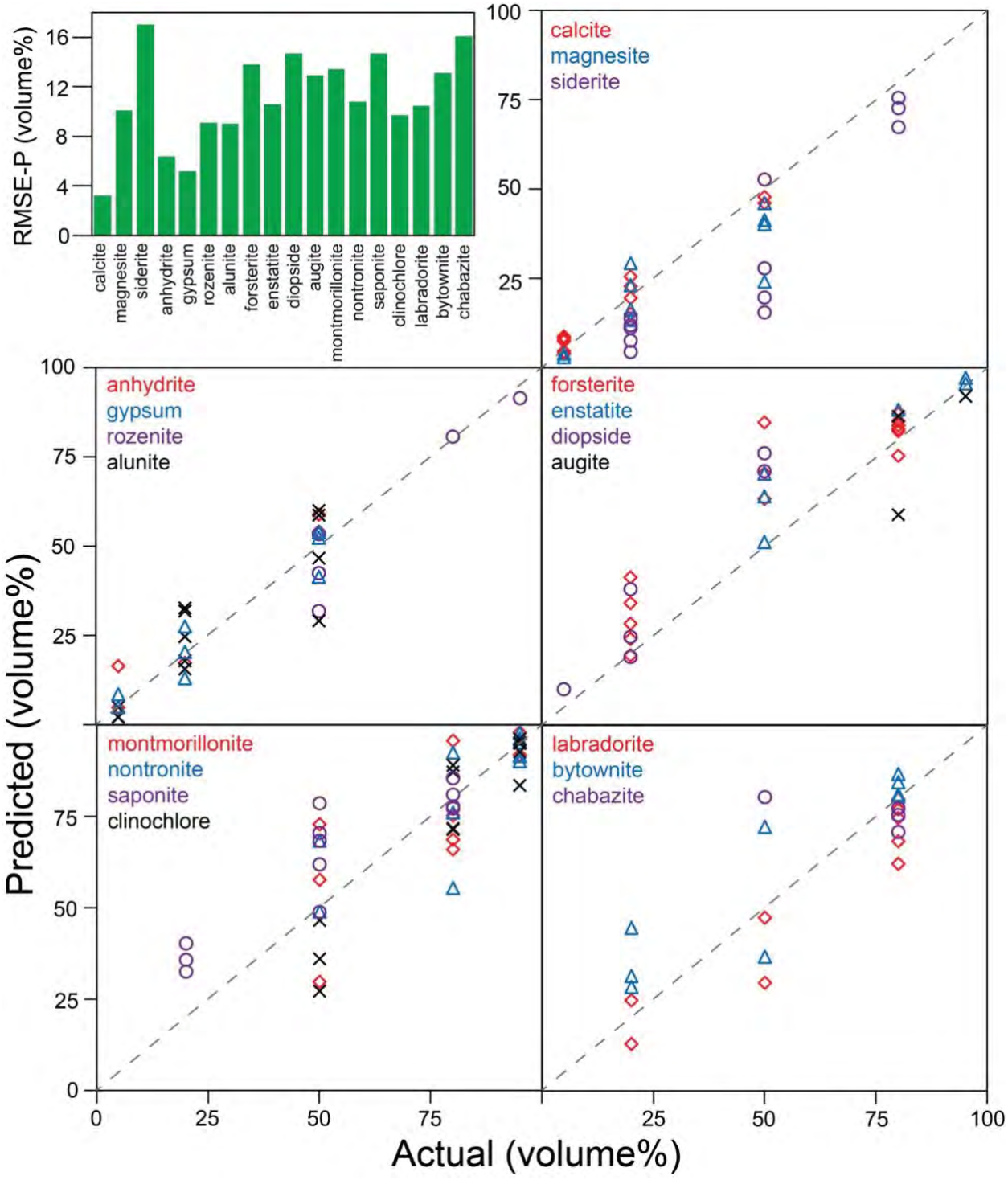


Figure 8. Partial least squares multivariate predictions versus actual abundances (volume%) for each prediction variable.

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2023-9224. http://www.minsocam.org/



1102

1103 Figure 9. Raman cross-section proxy RMSE-P values (top left) and predictions versus actual

1104 abundances (volume%) for each prediction variable. RMSE-P calculations were made using

1105 mineral-mineral mixtures.

1106

Raman Cross-Section Proxy	PLS Multivariate Analysis
(+) efficient model construction	(–) time consuming model construction
(+) small sample requirement	(–) large sample requirement
(–) mineral identification required	(+) mineral identification not required
(+) flexibility encountering unexpected materials	(–) difficulty encountering unexpected materials
(–) dependence on paired phases	(+) independence from paired phases
(–) inhibited by overlapping features	(+) uninhibited by overlapping features
(+) fast application	(+) fast application

1107

1108 Figure 10. Pros (+, green) and cons (–, red) of the Raman cross-section proxy and the

1109 multivariate analysis technique.