Journal of Raman Spectroscopy

WILEY

Journal of RAMAN SPECTROSCOPY

RESEARCH ARTICLE OPEN ACCESS

A Nonlinear Quantitative Model for Measuring Concentration Ratios From Raman Intensities

Joseph Razzell Hollis[1,2]

[1]Natural History Museum, London, UK | [2]NASA Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA

Correspondence: Joseph Razzell Hollis (j.razzellhollis@gmail.com)

Received: 23 September 2025 | **Revised:** 27 January 2026 | **Accepted:** 23 February 2026

ABSTRACT

Raman spectroscopy is a valuable tool for detecting trace compounds over wide ranges of concentrations but is usually limited to qualitative analysis (e.g., identification) due to the difficulty of determining concentration from Raman peak intensity except under very controlled conditions. This study presents a new quantitative model, derived from first principles, that can be used to estimate the concentration ratio of binary mixtures from their Raman intensities over several orders of magnitude, fully accounting for any nonlinear behaviors introduced by factors such as overlapping peaks and self-absorption. By training the model on experimental data of mixtures with known concentrations, the empirical parameters describing a particular mixture can be ascertained and then used to predict concentrations in further samples. The efficacy of the model is explored using synthetic datasets representing four scenarios depending on which compounds contribute to each peak. Bootstrapped model training can be used to consider the effects of noise, determine uncertainties for future predictions, and estimate the limits of detection and quantification for any given measurement. Finally, the model's efficacy is tested on experimental data for aqueous solutions of different organic nucleotides at concentration ratios between 0.1 and 1000 ppm, showing that the model works over 4 orders of magnitude and can be used to reliably predict the concentration ratio of test samples to within 0.1 orders of magnitude. This advanced model will improve our ability to estimate and assess concentrations in a wide range of mixed samples, even when their peaks overlap significantly.

1 | Introduction

Raman spectroscopy is a well-established noncontact, nondestructive analytical technique for detecting trace compounds in complex samples, obtaining distinctive spectra for a wide range of organic and inorganic compounds that are of interest to different research communities, with Raman routinely used in pharmaceuticals, mineralogy, meteoritics, material science, and microbiology, among others [1–8]. Raman is now even being used to look for chemical evidence of past life and habitability on Mars, with two instruments operating onboard NASA's *Perseverance* rover mission [9, 10]. Raman spectroscopy detects molecules by measuring the inelastic scattering of laser light by their vibrations, with the pattern of Raman scattering peaks being highly specific to the molecule's chemical structure [11]. The additive nature of Raman scattering means that the Raman signal is received from all compounds present in the sample simultaneously, weighted by the product of their concentrations and scattering cross-sections, enabling the combined detection of different components if their spectra do not overlap considerably [11, 12]. Raman has been used in this manner to study the composition of microbial cells, mineral mixtures, meteorites, and many other samples [6, 7, 13].

Combined with resonant enhancement to increase scattering yields, it can be possible to detect the Raman signature of certain organics at concentrations below 1 part per million [14]. While Raman intensity is intrinsically linear with respect to concentration, the quantitative use of Raman to measure concentration has historically been limited by the sensitivity of measured Raman intensities to other factors, such as variable output and focusing of the laser, and attenuation

and self-absorption of scattered light within the sample, all of which can introduce unexpected fluctuations and nonlinearity to any experimentally measured trend [15, 16]. These factors are not easily accounted for except in specific samples under controlled conditions, typically requiring the presence of an internal or external standard for normalization [12, 17, 18]. When normalization is not practical, these factors can be side-stepped by instead evaluating the intensity ratio of the analyte in question to another compound present in the sample, such as the medium, provided both can be detected simultaneously. By taking the ratio of intensities for two peaks, one from the analyte and one from the medium, for a series of experimental mixtures of known concentration ratio, it becomes possible to back-calculate the analyte/medium concentration ratio for subsequent samples of the same analyte/medium mixture. This approach has been used to achieve quantitative and semiquantitative analysis of concentration ratios in binary and ternary mineral/mineral and organic/mineral mixtures [13, 19–21].

However, these methods have always used linear equations to describe the relationship between Raman intensity ratio and concentration ratio, which assumes that one measured peak has an intensity exclusively due to analyte 1 while the other is exclusively from the medium. In practice, analyte and medium may have minor vibrational modes that overlap with the peaks of the other and will contribute a non-negligible signal at very high or very low concentration ratios, leading to decidedly nonlinear behavior at extreme concentrations. This nonlinearity has not been reported in previous studies, which have typically evaluated a relatively narrow range of concentration (1–2 orders of magnitude) where linearity is expected, and very rarely examine concentration ratios below 1% (~10,000 ppm). For a quantitative approach to work over a larger range of potential concentrations, and especially at concentrations as low as 1 part per million, the traditional linear model of Raman intensity ratios must be corrected and refined to account for the nonlinearity that can be introduced by cross-contributions from any overlapping peaks of the analyte and medium.

This article presents a novel quantitative model derived from first principles to accurately describe how measured intensity ratios relate to concentration ratios for binary mixtures, accounting for all possible permutations of cross-contributions from analyte and medium. It explores the behavior and limitations of the model based on synthetic datasets and then tests it on real experimental data of aqueous organic/salt/water solutions from a previous study, spanning an organic/water concentration ratio range of 0.1 to 1000 ppm [15]. It will also describe how to use bootstrap resampling to improve the robustness of the model and enable estimation of other important parameters, such as prediction uncertainty, limits of detection, and limits of quantification.

2 | Materials and Methods

2.1 | Aqueous Mixtures

The preparation of aqueous nucleotide solutions was originally reported in Razzell Hollis et al. [15]. In brief, deoxyribose adenosine triphosphate (dATP), deoxyribose cytidine triphosphate (dCTP), deoxyribose guanosine triphosphate (dTTP), and deoxyribose thymidine triphosphate (dTTP) were used as received, as a 100-mM PCR-grade nucleotide solution (Sigma Aldrich DNTP-100), that were mixed with a stock solution of 100 mM $Na_2SO_4.10H_2O$ (Sigma Aldrich) dissolved in purified MilliQ water to obtain nucleotide concentrations of 0.01 to 50 mM. Na_2SO_4 concentration was therefore $100 - X$ mM where X is the concentration of nucleotide.

2.2 | DUV Raman Spectroscopy

The collection of experimental Raman spectra was originally reported in Razzell Hollis et al. [15]. In brief, all measurements were done using MOBIUS, a custom deep-ultraviolet (DUV) Raman spectrometer at the NASA Jet Propulsion Laboratory. MOBIUS uses a 248.6-nm pulsed NeCu laser (Photon Systems Inc) and a liquid N_2 cooled CCD detector (Horiba Symphony e2v 42-10). A slit width of 250 μm and a grating of 1800 lines/mm produced a spectral step size of 3.8 cm^{-1}. Twenty-five spectra were collected per sample, each integrated over 1200 laser pulses at 40 Hz and an energy of 2.5–3.4 μJ/pulse. Immediately prior to measurement, 50 ± 0.05 μL of solution was deposited onto a clean aluminum wafer beneath the objective lens and the sample stage z-axis position adjusted so that the wafer was in focus through the droplet; all spectra were collected within 20 min of deposition to minimize the effects of evaporation on concentration.

2.3 | Spectral Processing

Preprocessing of spectra was done in Loupe, a custom program written for MOBIUS, and included laser output normalization and cosmic ray removal [22]. Further processing was done using a set of Jupyter Notebooks [23], adapted from the publicly available OSTRI suite developed for Raman and FTIR analysis by Razzell Hollis [24]. Processing included automatic baseline subtraction by polynomial fitting, automatic peak detection, peak fitting using pseudo-Voigt functions, and recording of spectra to a standardized data format. Intensity values were taken directly from each point spectrum according to a set of pre-defined frequencies for each component molecule in the sample.

2.4 | Synthetic Data Generation

Synthetic data comprised intensity values generated for a binary mixture with given concentrations of molecules A and B. The total intensity of peak v was generated using Equation (1), based on predefined values for incident intensity I_0, the Raman scattering cross-sections of A and B for that peak, $J_{A,v}$ 3and $J_{B,v}$, their concentrations C_A and C_B, an assumed interrogation volume V, and the instrument sensitivity F_v and the attenuation factor X_v for that wavelength. Optional dark noise and shot noise were added using normally distributed random values with standard deviation N_{dark} and N_{shot}.

$$I_v = \left(I_0 \cdot F_v \cdot X_v \cdot V \cdot \left(J_{A,v}C_A + J_{B,v}C_B\right) \pm N_{dark}\right) \times \left(1 \pm N_{shot}\right) \quad (1)$$

2.5 | Model Training and Testing

The model was trained on a given set of intensity ratios for samples of known concentration ratio to determine the best-fit values for empirical parameters *FXJ*, *FXB*, and *JA*. Model training was done in log space by Nelder–Mead (downhill simplex) minimization of the cost function using the LMFIT Python package [25]. Training data were first cleaned to remove nonpositive values. A first fit was attempted; then, outlier data points were identified based on residuals >1 order of magnitude and residuals ≥2 times the standard deviation of all residuals, and if any outliers were found, a second fit was done without those points. Bootstrapping was used to obtain more representative estimates of model parameters and their uncertainties: Intensity ratio values for each input sample were resampled with replacement 1000 times; each set of resampled values was then fitted using the model to obtain 1000 sets of estimated parameters. Overall model parameters were calculated from the means and standard deviations of estimated parameters. See Appendix B (Supporting Information) for a detailed description of all data cleaning, fitting, and evaluation steps of model training.

3 | Results

3.1 | The Quantitative Model

For a binary mixture of two molecular species A and B, with two distinct and measurable Raman peaks, there will be two measured intensities, I_1 and I_2, that depend on the concentrations of A and B and how much they each contribute to each peak. The empirical relationship between concentration ratio and measured intensity ratio is given by Equation (2). For the full derivation of this equation from first principles, as well as detailed descriptions of the factors included in each term, see Appendix A in the Supporting Information. The underlying assumptions of this equation will be described in detail in the "Discussion" section.

$$\frac{I_1}{I_2} = \frac{FXJ \cdot \frac{C_A}{C_B} + FXB}{JA \cdot \frac{C_A}{C_B} + 1} \tag{2}$$

Equation (2) contains five distinct terms that describe how differently the Raman spectrometer interrogates peaks I_1 and I_2 depending on their frequencies, and how much A and B contribute to each peak. F is the ratio of the instrument's overall sensitivity to I_1 versus I_2; X is the ratio of the weighted volume interrogated by the instrument at I_1 versus I_2; J is the ratio of scattering cross-sections for A at I_1 versus B at I_2; A is the ratio of scattering cross-sections for A at I_2 versus I_1; B is the ratio of scattering cross-sections for B at I_1 versus I_2 The last two terms account for all possible cross-contributions by A and B, and the larger J is relative to A and B, the more linear the curve defined by Equation (2) will be (see Figure S1).

However, these are coupled together into just three correlated parameters *FXJ*, *FXB*, and *JA* that are needed to fully describe the behavior of $\frac{I_1}{I_2}$ in terms of $\frac{C_A}{C_B}$. These can be determined empirically by measuring the trend in $\frac{I_1}{I_2}$ versus $\frac{C_A}{C_B}$ (see Figure S1) for binary mixtures of A and B across several known concentration ratios, and fitting it with Equation (2) to get estimated values of $\widehat{FXJ}$, $\widehat{FXB}$, and $\widehat{JA}$. Once *FXJ*, *FXB*, and *JA* are known for a given combination of molecules A and B, Equation (2) can be used to predict the intensity ratio for any possible concentration ratio, and the inverse Equation (3) can be used to predict concentration ratio for any possible intensity ratio.

$$\frac{C_A}{C_B} = \frac{\frac{I_1}{I_2} - FXB}{FXJ - JA\frac{I_1}{I_2}} \tag{3}$$

Equations (2) and (3) hold true as long as $F, X, J > 0$ and $A, B \geq 0$. This leads to four possible trends for $\frac{I_1}{I_2}$ versus $\frac{C_A}{C_B}$ depending on whether $A > 0$ and/or $B > 0$, as shown in Figure 1. If $A > 0$, there is a nonzero contribution of A to I_2 that becomes significant at very high concentration ratios where A dominates, and $\frac{I_1}{I_2}$ approaches an asymptotic upper limit defined by FXJ/JA. If $B > 0$, there is a nonzero contribution of B to I_1 that becomes significant at very low concentration ratios where B dominates, and $\frac{I_{v1}}{I_{v2}}$ approaches an asymptotic lower limit defined by *FXB*. Herein, these different scenarios are referred to as Scenario 1 (where $A, B > 0$), Scenario 2 ($A = 0, B > 0$), Scenario 3 ($A > 0, B = 0$), and Scenario 4 ($A, B = 0$).

The synthetic data plotted in Figure 1a–d was generated using Equation (1) with known values of F, X, J, A, B with no added noise (see Figure S2 for absolute intensities). It shows that the linear relationship presumed in many quantitative and semi-quantitative Raman models is only strictly true when there are no cross-contributions at all ($A, B = 0$, Scenario 4). When $A > 0$ or $B > 0$ (Scenarios 2 and 3), the trend will only be approximately linear over a limited range of concentration ratios, becoming nonlinear as it approaches the asymptotic value defined by either FXJ/JA or *FXB*. This range is smallest (all else being equal) when $A, B > 0$ (Scenario 4).

When the nonlinear quantitative model was fitted to each curve in Figure 1a–d, it replicated the input data and estimated parameters $\widehat{FXJ}$, $\widehat{FXB}$, and $\widehat{JA}$ from fitting were similar to their true values (Table 1; Figure S3). For true parameters that were nonzero, estimated values were within ±0.1% of true values, indicating that the model can accurately fit simulated data derived from first principles. However, the model was only fitted over a particular range of concentration ratios and cannot correctly estimate parameter values that would only produce variations outside that range; for example, if $FXB = 0$ but the curve is only fitted to data with intensity ratios of 10^{-2} or higher, then the model has no way to distinguish $FXB = 0$ from $FXB = 10^{-13}$ (see Scenario 4 in Figure 1d). Consequently, it is important to train each model on data that represent the range of concentrations that are expected to be encountered in test samples and to understand that extrapolating beyond the range of training data may lead to inaccurate results.

3.1.1 | Effects of Noise

Spectral noise creates uncertainty in measured Raman intensities that will affect the quality of any fit and introduce uncertainty to estimated parameters. Noise comes in two main

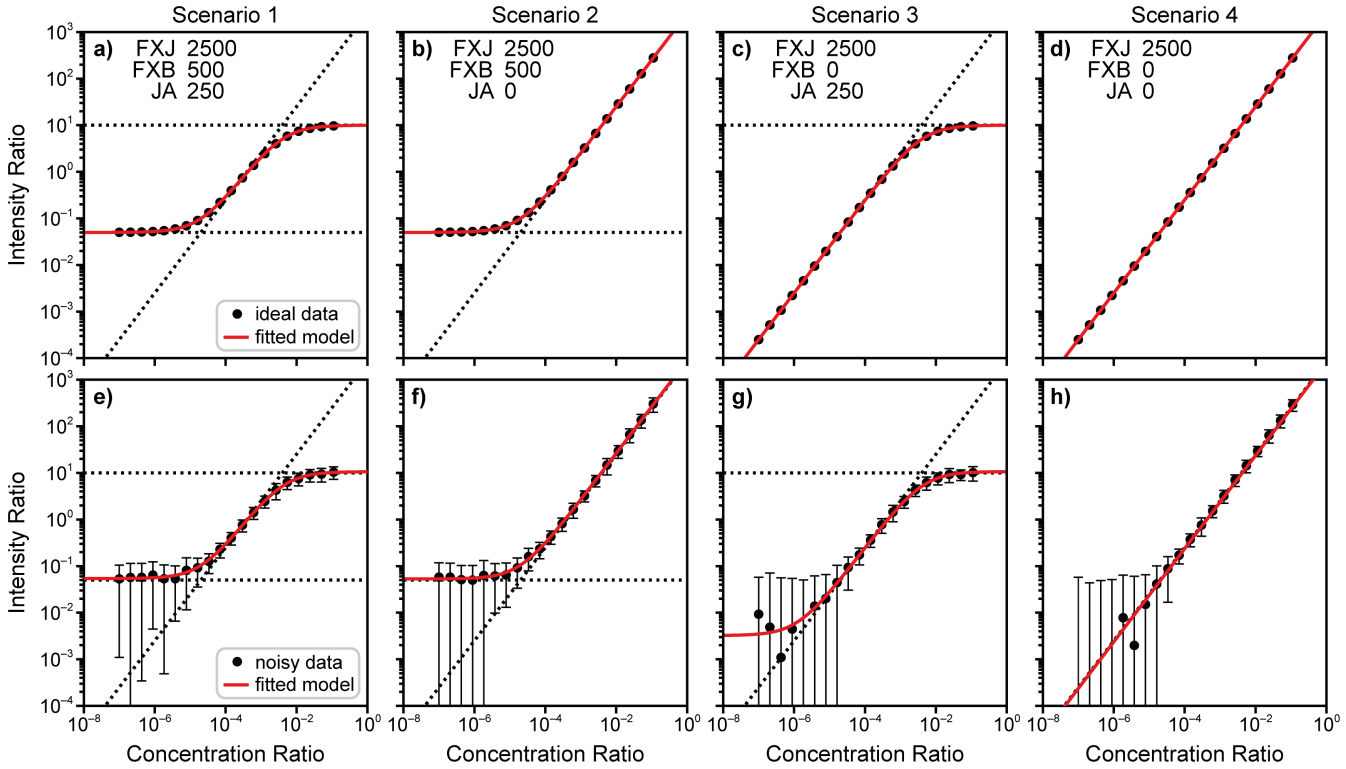

FIGURE 1 | Synthetic data showing the trends in intensity ratio vs. concentration ratio for Scenarios 1–4, where $A, B > 0$ (a), $B > 0$ (b), $A > 0$ (c), and $A, B = 0$ (d), respectively. (a–d) Intensity ratios and resulting fit with the quantitative model for idealized synthetic data with no added noise. (e–h) Intensity ratios and resulting fit with the quantitative model for idealized synthetic data with normally distributed dark noise of ±1000 and shot noise of ±0.2×. Dotted black lines indicate the linear relationship $\frac{I_{v1}}{I_{v2}} = FXJ\frac{C_A}{C_B}$ and the asymptotic limits defined by *FXB* and *FXJ* / *JA*.

TABLE 1 | True values versus fitted values for parameters *FXJ*, *FXB*, and *JA* plus asymptote *FXJ* / *JA*, evaluated for synthetic data with and without dark noise of ±1000 and shot noise of ±0.2×. Input data consisted of 100 values per evaluated concentration ratio; fitting was done after automatic trimming of nonpositive and outlier data points.

Parameter		Scenario 1	Scenario 2	Scenario 3	Scenario 4
FXJ	True	2.50E+03	2.50E+03	2.50E+03	2.50E+03
	Fit (no noise)	2.50E+03	2.50E+03	2.50E+03	2.50E+03
	Fit (noise)	2.50E+03	2.63E+03	2.48E+03	2.36E+03
FXB	True	5.00E−02	5.00E−02	0	0
	Fit (no noise)	5.00E−02	5.00E−02	3.21E−12	4.53E−13
	Fit (noise)	5.39E−02	5.29E−02	3.23E−03	2.22E−16
JA	True	2.50E+02	0	2.50E+02	0
	Fit (no noise)	2.50E+02	1.25E−07	2.50E+02	1.06E−09
	Fit (noise)	2.36E+02	1.38E−12	2.33E+02	1.37E−13
FXJ / *JA*	True	1.00E+01	Inf	1.00E+01	inf
	Fit (no noise)	1.00E+01	2.00E+10	1.00E+01	2.35E+12
	Fit (noise)	1.06E+01	1.91E+15	1.07E+01	1.73E+16

forms in Raman spectroscopy: *dark noise*, which is additive and independent of absolute signal, and *shot noise*, which is multiplicative and increases with absolute signal (see Appendix B, Supplementary Information) [26, 27]. When noise is added to synthetic data, as shown in Figure 1e–h, dark noise tends to increase uncertainty at low concentrations, while shot noise increases uncertainty across all concentrations in proportion to their absolute signal (see Figures S5–S8). Because the model relies on assessing the ratio of two peaks, the impact of noise is compounded: normally distributed shot noise of ±0.2× leads to ±0.34× standard deviation in intensity ratio, such that 95% of measured intensity ratios are within ±0.69×.

Spectral noise has three effects on the quantitative model fit: (1) Increased variance in input intensity ratios leads to greater uncertainty in all fitted parameters and introduces random error versus their true values (see Figures S5–S8). (2) At extreme concentration ratios where one peak is very weak, even small absolute variances from dark noise can lead to outliers that deviate from expected values by several orders of magnitude, which may compromise the quality of the fit. This phenomenon is responsible for the very large negative error bars seen in Figure 1e–h; such outliers may need to be trimmed from the training dataset to avoid spurious fits (Figure S4). (3) large absolute variances can even lead to nonpositive intensity values at low concentration ratios, which must be trimmed to avoid computational failure when converting input data to log space.

Removing nonpositive values and outliers allows the fit to proceed but reduces the size of the training dataset and potentially overestimates the value of $\widehat{FXB}$, as shown for the Scenario 3 example in Figure 1g. The effects of noise can be minimized by adjusting measurement settings to maximize signal: noise ratios as much as possible for all spectra used to train the model. However, even with relatively large magnitudes of noise (N_{dark} ± 1000, N_{shot} $\pm 0.2\times$), estimated values of nonzero parameters are still within $\pm 6\%$ of their true values.

3.1.2 | Accounting for Uncertainty

It is important to understand how accurate the estimated parameters are when fitting data that includes some uncertainty from noise and measurement error [28, 29]. Many available fitting algorithms (such as SciPy's *curve_fit* and LMFIT's *minimize* functions) include functionalities for estimating the standard errors of fitted parameters as part of the fitting process [25, 30]. However, these packages typically assume normally distributed parameters with no bounds, and as parameters *FXJ*, *FXB*, and *JA* are bounded at 0 and may be non-normal, standard errors produced by these packages may not be correct.

The simplest approach is to approximate the underlying empirical distribution of observed data by bootstrapping, fitting the model N times on random samples (with replacement) of the input dataset [31]. This method (described in detail in Appendix B) is more computationally intensive, as hundreds or thousands of fits must be done to properly sample the distribution, but neatly sidesteps the need to derive the appropriate equations for propagating uncertainties to each parameter. The resulting distributions of fitted parameters $\widehat{FXJ}$, $\widehat{FXB}$, and $\widehat{JA}$ (Scenario 1 shown in Figure 2, see Figures S9–S11 for Scenarios 3 and 4) can be used to obtain median values that are robust against singularly poor fits to calculate covariance and correlation coefficients for fitted parameters (see Figures S12–15) and to estimate uncertainty of predictions made using the trained model.

Figure 2 also shows the impact of increasing shot noise on bootstrapped parameter distributions for a Scenario 1 dataset. Distributions may be normal or non-normal, depending on whether they have a lower bound (e.g., as *FXB* approaches 0). Consequently, mean and standard deviation are not always appropriate, instead parameter distributions are better described by their median and the 16th–84th percentile range ($\approx \hat{X} \pm 1\sigma$ for normally distributed values). As shot noise increases, the error between median parameters and their true values tends to increase (see Table S1) and the distributions become broader. Despite this, the variance of fitted parameters is still much smaller than the effective variance of the input data: When shot noise is $\pm 0.2\times$, the standard deviation of intensity ratio is $\pm 0.34\times$ but the equivalent spread of parameter $\widehat{FXJ}$ in Figure 2 is $(2.54–2.70)\times 10^3$ (roughly equivalent to $\pm 0.06\times$).

3.1.3 | Predicting Concentration From Intensity

Once the model has been trained on experimental data to determine the representative distribution of parameters *FXJ*, *FXB*, and *JA* for a given binary mixture of *A* and *B*, those parameters can be used with Equation (3) to predict the concentration ratio for other samples based on the observed intensity ratio. A bootstrap-trained model can provide N predictions per input test value (see Figure S16), allowing estimation of median and uncertainty that reflects the uncertainty in the model itself. For multiple test values, the resulting distribution is a convolution of uncertainty from the model and from test data (as shown in Figure 3).

Consider two hypothetical test samples of the same mixture with true concentration ratio $\frac{C_A}{C_B} = 10^{-3}$ and $\frac{C_A}{C_B} = 10^{-6}$. Each was measured 100 times to get intensity ratios $\frac{I_1}{I_2} \approx 2.1 \pm 0.28$ and $\frac{I_1}{I_2} \approx 2.1 \pm 0.28$, respectively, shown in Figure 3a,b. When a 1000-fold bootstrap-trained model was used to predict concentrations for these samples, it produced 100×1000 predicted values with a distribution that represents the convolution of the uncertainty from the model and the uncertainty of the test data itself (Figure 3c,d). For the sample where $\frac{C_A}{C_B} = 10^{-3}$, the median prediction was $\frac{C_A}{C_B} = 1.01 \times 10^{-3}$ (an error of +0.1% vs. truth) with an 16th–84th percentile range of $(0.713 - 1.45) \times 10^{-3}$. For the sample where $\frac{C_A}{C_B} = 10^{-6}$, the median prediction was an erroneously negative -0.916×10^{-6}.

Negative concentration ratio predictions occur when the evaluated intensity ratio falls outside the expected range of $FXB \leq \frac{I_1}{I_2} \leq \frac{FXJ}{JA}$. 0% of predictions were erroneous for $\frac{C_A}{C_B} = 10^{-3}$ but 56% were erroneous for $\frac{C_A}{C_B} = 10^{-6}$ as the evaluated $\frac{I_1}{I_2} \approx FXB$. When a substantial proportion of predictions are erroneous for a particular sample, the user should consider whether the measured intensity ratio falls outside the limits of quantification based on the uncertainty in the measurement and the model.

3.1.4 | Limits of Detection and Quantification

There will always be a limit to how low the concentration of a molecule can go before it becomes undetectable, or unquantifiable, in a given analytical measurement. The limit of detection is typically defined as the minimum concentration required to have an acceptably small chance of being a false positive (that the measured signal could have come from a blank sample). The limit of quantification is the minimum concentration

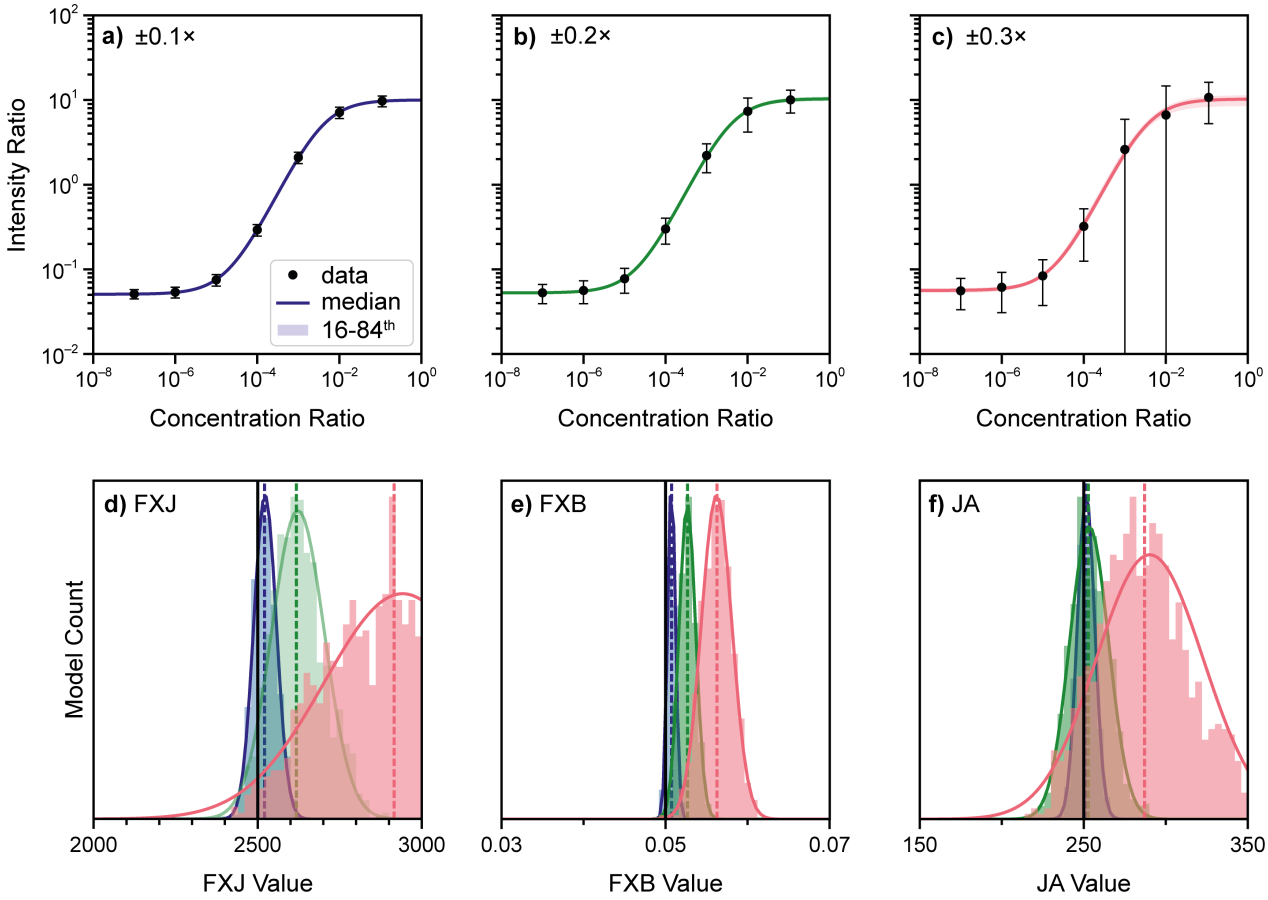

FIGURE 2 | Application of bootstrapped model training to noisy datasets. (a–c) Synthetic datasets generated using true values $FXJ = 2500, FXB = 0.05, JA = 250$ (Scenario 1) with $N_{dark} = 0$ and either N_{shot} ±0.1× (blue), ±0.2× (green), and ±0.3× (red). (d–f) Distributions of fitted parameters from 1000-fold bootstrapped model training on each dataset. Dashed vertical lines indicate median parameter values from bootstrapping, solid vertical lines indicate true values, and solid curves indicate approximate normal distributions. See Figure S9–11 for Scenarios 2–4.

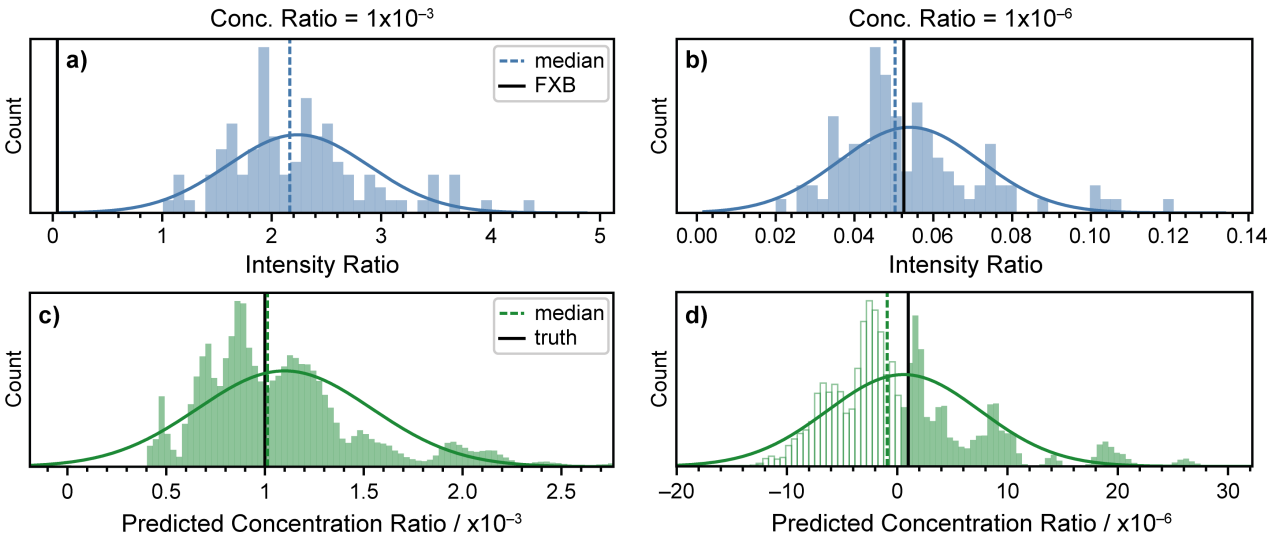

FIGURE 3 | Application of trained models to predict concentration for previously unseen synthetic test data representing samples with concentration ratio $\frac{C_A}{C_B} = 10^{-3}$ (a) and 10^{-6} (c). (a, b) The distribution of 100 synthetic intensity ratios for each test sample. (c, d) Distribution of predicted concentration ratios from evaluating test data distribution using a 1000-fold bootstrap-trained model.

required to have an acceptably small chance of being a false negative (that the measured signal is less than that at the limit of detection). Bootstrap-trained models allow for robust numerical estimation of these limits by simulating the expected intensity ratio distributions for mixtures at different concentrations and comparing them to the simulated distribution of a blank sample. In this study, the acceptable risk was set at 1%, a signal equivalent to the 99th percentile of the blank. Because

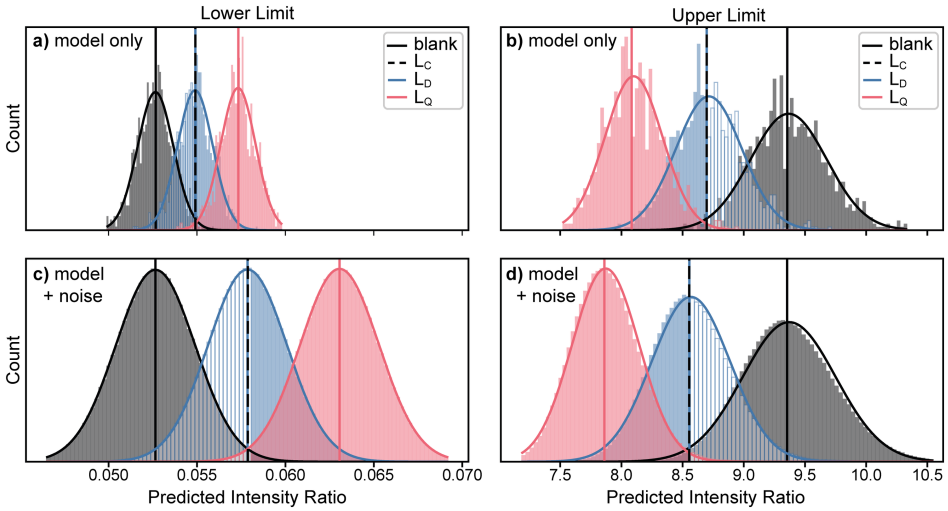

FIGURE 4 | Distribution of predicted intensity ratios at L_D (the limit of detection) and L_Q (the limit of quantification) for a trained Scenario 1 model versus the lower blank $\frac{C_A}{C_B} = 0$ (a, c) and upper blank $\frac{C_A}{C_B} = \infty$ (b, d). Results are given for the inherent model-driven limits (a, b) and noise-driven limits where either I_2 (c) or I_1 (d) has a known SNR of 500:1. Intensity ratios beyond the critical value (the 99th percentile of $\widehat{FXB}$ for lower limits, the first percentile of $\frac{\widehat{FXJ}}{\widehat{JA}}$ for upper limits) are indicated by white-filled boxes in each histogram.

the quantitative model involves two peaks I_1 and I_2, there will be upper and lower limits for both detection and quantification, with lower limits dictated by the 99th percentile of *FXB* (Figure 4a) and upper limits dictated by the 1st percentile of *FXJ* / *JA* (Figure 4b).

Figure 4 shows the intensity ratio distributions predicted by a trained model for blank samples (black), at the limit of detection (blue), and at the limit of quantification (red). In the absence of noise, the limits of detection and quantification will be determined solely by the uncertainty in $\widehat{FXB}$ and $\widehat{FXJ}/\widehat{JA}$ from model training (Figure 4a,b). Both lower and upper limits are shown, leading to an overall model-driven range of detection of $5.5 \times 10^{-7} < \frac{C_A}{C_B} < 1.4 \times 10^{-1}$, and a model-driven range of quantification of $1.1 \times 10^{-6} < \frac{C_A}{C_B} < 7.1 \times 10^{-2}$. These ranges will narrow with increasing uncertainty in $\widehat{FXB}$ and/or $\frac{\widehat{FXJ}}{\widehat{JA}}$, and no matter how accurately and precisely a test sample is measured, this particular model will never be able to robustly detect or quantify concentrations outside this range.

In practice, there will also be some background noise in a measurement that will further hinder the quantitative analysis of a given mixture. If the level of noise in a test measurement is known, it can be included in the calculation of detection/quantification limits by broadening the distributions shown in Figure 4a,b accordingly (see Appendix B) and then reevaluating L_D and L_Q. Uncertainty due to noise will, inevitably, narrow the range of concentration ratios that can be reliably detected and quantified: For the data shown in Figure 4c, where the peak I_2 has a hypothetical SNR of 500:1, the noise-driven lower limit of detection would be $1.9 \times 10^{-6} < \frac{C_A}{C_B}$ and the noise-driven lower limit of quantification would be $3.8 \times 10^{-6} < \frac{C_A}{C_B}$. These calculations enable the user, when faced with a spectrum with a well-defined peak I_2 but no clear detection of peak I_1, to estimate the maximum possible concentration ratio that could be present without I_1 being detectable. Conversely, the user can estimate the minimum possible concentration ratio when I_1 is detected and I_2 is not.

3.1.5 | Assessing Accuracy

Figure 5 shows the results obtained when four models representing Scenarios 1–4 (trained on a minimal synthetic dataset of 1 sample per order magnitude from $\frac{C_A}{C_B} = 10^{-7}$ to 10^{-1}, 100 measurements per sample) were then tested on a fresh set of synthetic data generated across the same range. Two test datasets were generated for each scenario, one with no noise and one with the same level of added noise as the training data, $N_{dark} = \pm 100$ and $N_{shot} = \pm 0.1\times$.

When evaluated on noiseless test data (blue lines in Figure 5), each model is limited only by how accurately it fitted the training data. Predicted concentration ratios are close to truth over a wide range, but at very low or very high concentration ratios, there is a systematic increase in error due to uncertainty in the model's estimated value of *FXB* and $\frac{FXJ}{JA}$, respectively. For Scenario 1, the estimated model-driven limits of quantification are $\sim 1 \times 10^{-6} < \frac{C_A}{C_B} < \sim 6 \times 10^{-2}$; within this range, the root-mean-square prediction error (RMSPE) for the median prediction is ~0.014 orders of magnitude; that is, the median prediction made for noiseless test data in this range will be, on average, within 0.97–1.03× of its true value. Over the same range, the risk of getting erroneous predictions ($\frac{C_A}{C_B} < 0$) or false negatives ($\frac{C_A}{C_B} < L_D$) is estimated to be zero (Figure 5i–l).

For noisy test data (red lines in Figure 5), the model's ability to accurately predict concentration ratio is further hindered by variance in input data, particularly when either peak intensity approaches the magnitude of the dark noise. This leads to increased uncertainty in predictions, especially at very low concentration ratios, until even the median prediction is no longer reliable due to random error. For Scenario

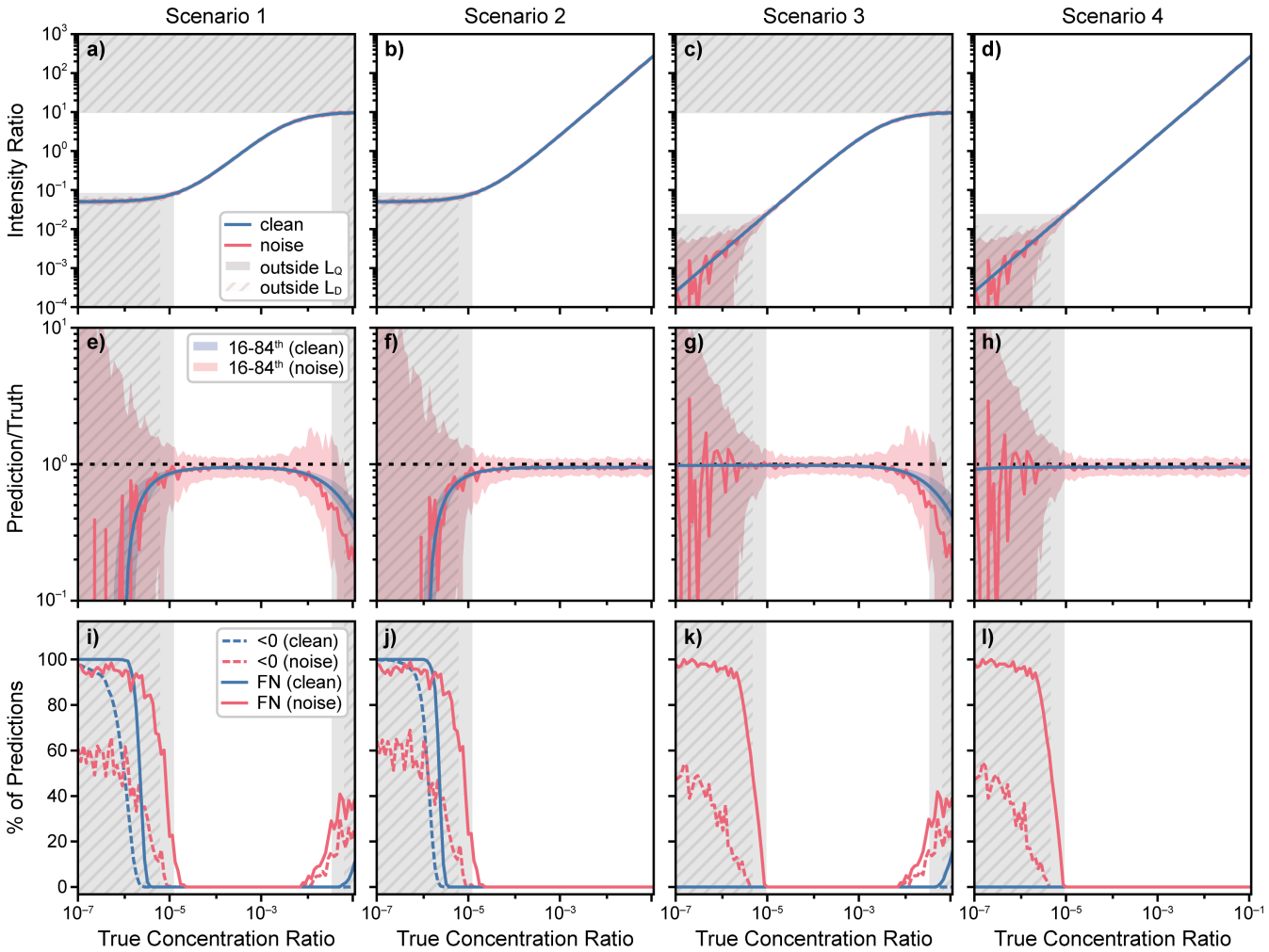

FIGURE 5 | Application of trained models to test data for Scenarios 1–4. (a–d) Intensity ratios for test datasets with noise (red) and without noise (blue). (e–h) The relative error of predicted concentration ratios versus their true values. (i–l) The corresponding rate of false negatives and nonpositive predictions. Solid lines indicate median values; colored areas show corresponding 16th–84th percentile ranges for each dataset. Solid gray areas indicate regions outside the noise-driven limits of quantification; hatched gray areas indicate regions outside the noise-driven limits of detection. Training and noisy test data were both generated with $N_{dark} = \pm 100$ and $N_{shot} = \pm 0.1\times$.

1, the estimated noise-driven limits of quantification are $\sim 1\times10^{-5} < \frac{C_A}{C_B} < \sim 6\times10^{-2}$, within which the RMSPE for the median prediction is 0.05 orders of magnitude; that is, median prediction is within 0.90–1.12× of truth. The lower limit coincides neatly with the onset of increased error and increased risk of false negatives at very low concentrations for all four scenarios, highlighting the value of estimating limits as a method for gauging the reliability of prediction for a given model and measurement (Figure 5i–l).

3.1.6 | Case Study

Aqueous Nucleotides.

An experimental case study was conducted to test the efficacy of the nonlinear quantitative model using real mixtures: solutions of an organic small molecule, deoxythymidine triphosphate (dTTP), dissolved in purified water with an internal standard, Na_2SO_4. Figure 6 shows the average normalized deep-ultraviolet Raman spectra measured for varying dTTP concentrations 0.01–50 mM (equivalent to a dTTP:H_2O concentration ratio of $10^{-7} - 10^{-3}$ and a dTTP:Na_2SO_4 concentration ratio of $10^{-4} - 10^{0}$). dTTP exhibits multiple Raman peaks between 800 and 1800 cm^{-1}, with the two strongest peaks occurring at 1350 and 1635 cm^{-1}. The internal standard Na_2SO_4 exhibits a single *v1* stretching peak of the dissolved $[SO_4]^{2-}$ anion at 975 cm^{-1} [32], while water exhibits two peaks, a strong, broad O–H stretching mode around 3400 cm^{-1} and a much weaker O–H bending peak at ~1625 cm^{-1} [33].

Peak intensities were measured using the simplest method possible: directly taking measured counts from the pixel closest to the specified peak position in each spectrum, which were then averaged across 25 measurements for each sample. Absolute peak intensities (shown in Figure S17) were highly variable and tended to decrease with increasing dTTP concentration due to strong ultraviolet self-absorption by dTTP [34], which effectively reduces the weighted interrogation volume of the sample. This is fully accounted for when intensity ratio is taken, leading to curves consistent with the quantitative model (Figure 6), but the weak 975 cm^{-1} Na_2SO_4 was so reduced in intensity that it became dominated by

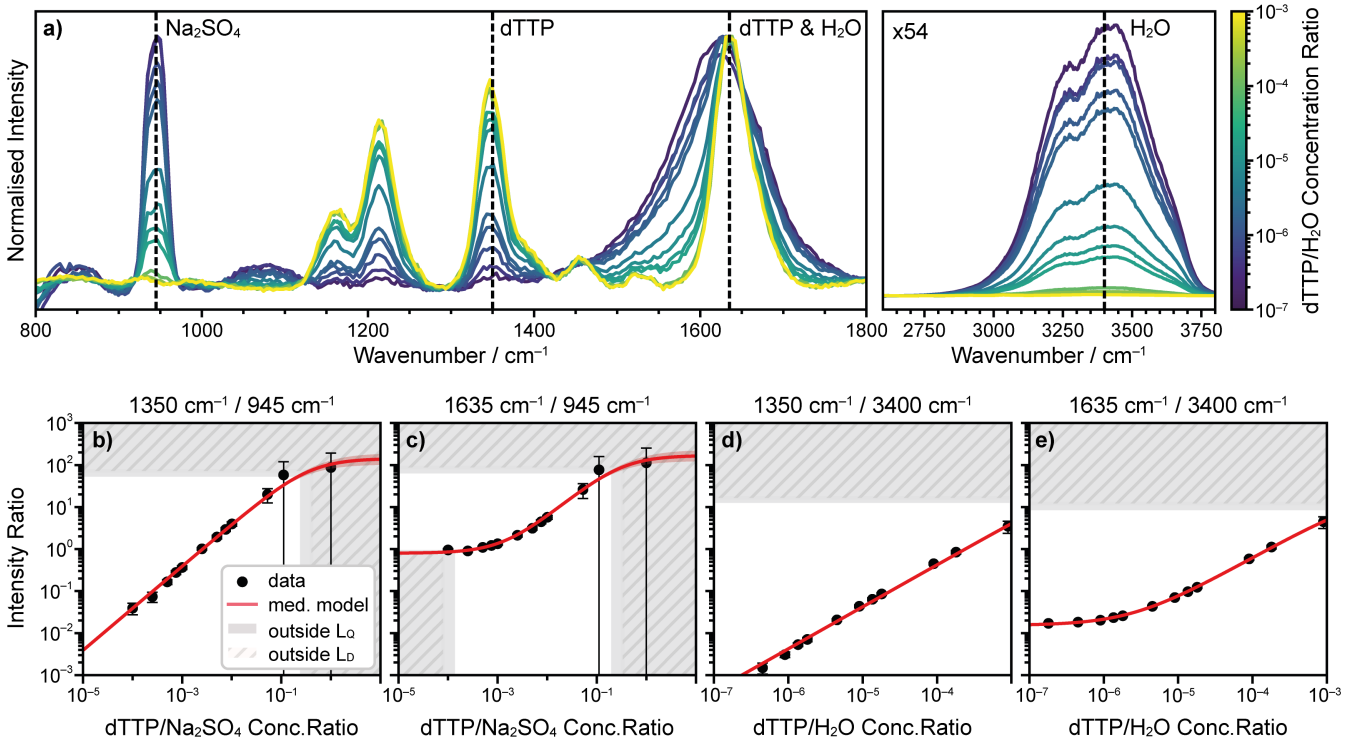

FIGURE 6 | (a) Experimentally measured spectra for aqueous mixtures of dTTP and an internal standard, Na_2SO_4, in purified H_2O [15]. Spectra were averaged over 25 measurements, baselined, normalized to the strongest peak between 800 and 1800 cm^{-1}, and colored according to dTTP/H_2O concentration ratio. (b–e) Average intensity ratios and trained model curves for different combinations of peaks from dTTP versus Na_2SO_4 (b, c) and from dTTP versus H_2O (d, e). Red areas indicate the 16the–84th percentiles for bootstrapped models, solid gray areas indicate regions outside the model-driven limits of quantification, and hatched gray areas indicate regions outside the model-driven limits of detection.

background noise (Figure S17), leading to large variance at high concentration ratios (see error bars in Figure 6b,c). By comparison, the 3400 cm^{-1} H_2O peak signal was still strong enough to be consistently measured even at very high concentration ratios (Figure 7c,d). At the other end of the range, the relative intensity of the 1635 cm^{-1} dTTP peak flattened out at low concentration ratios as it became dominated by the overlapping ~1625 cm^{-1} H_2O peak, leading to a lower limit determined by H_2O not dTTP (Figure 6c,e). Overall, the 1350 cm^{-1}/3400 cm^{-1} ratio (Figure 7d) was the most linear, with measured intensity ratio varying by 4 orders of magnitude with no apparent upper or lower limits in the evaluated range, making it most consistent with Scenario 4; the 1635 cm^{-1}/945 cm^{-1} ratio spanned just 2 orders of magnitude and exhibits both lower and upper limits, making it most consistent with Scenario 1; 1635 cm^{-1}/3400 cm^{-1} exhibited a lower limit consistent with Scenario 2; 1350 cm^{-1}/945 cm^{-1} exhibited an upper limit consistent with Scenario 3.

Figure 6 also shows the median models obtained from bootstrapped training on each of the four plotted datasets, demonstrating that each experimentally measured intensity ratio can be accurately fitted using the quantitative model, accounting for the various nonlinear behaviors observed. Estimated values and uncertainties for parameters $\widehat{FXJ}$, $\widehat{FXB}$, and $\widehat{JA}$ are reported for each model in Table 2, as well as the inherent limits of detection and quantification based on those uncertainties. The 1645 cm^{-1}/945 cm^{-1} model (dTTP/Na_2SO_4; Figure 6c) was highly nonlinear, with median parameters $\widehat{FXB} = 8.0 \times 10^{-1}$ and $\frac{\widehat{FXJ}}{\widehat{JA}} \approx 1.7 \times 10^{2}$ and model-driven limits of quantification

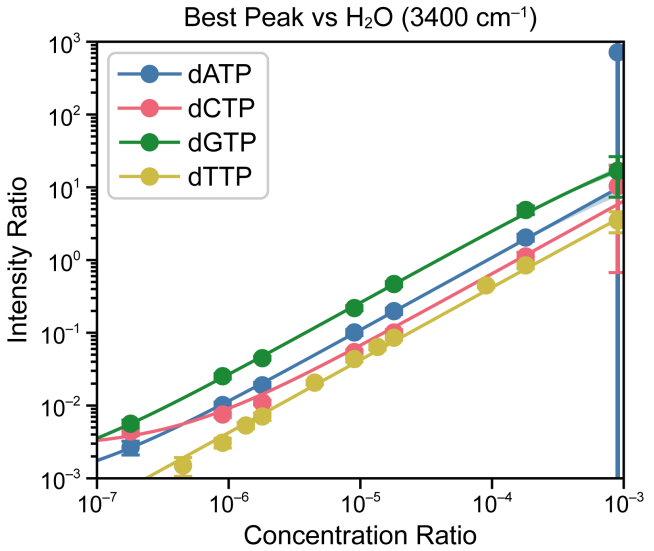

FIGURE 7 | Experimentally measured intensity ratios for mixtures of the different nucleotides dATP, dCTP, dGTP, and dTTP with Na_2SO_4 in H_2O, evaluated for the most distinctive peak of each nucleotide (1310, 1505, 1460, 1350 cm^{-1}, respectively) versus the 3400 cm^{-1} peak of H_2O. All model-driven limits of detection and quantification fall outside the plotted concentration ratio range.

of $1.3 \times 10^{-4} - 2.1 \times 10^{-1}$, meaning that it can reliably quantify dTTP/Na_2SO_4 concentration ratios spanning at most 3.2 orders of magnitude. By comparison, the 1350 cm^{-1}/3400 cm^{-1} model (dTTP/H_2O; Figure 6d) was the most effective in terms of

TABLE 2 | Estimated model parameters (median, 16th–84th percentiles) for the different peak pairs of dTTP versus Na_2SO_4 and dTTP versus H_2O. The upper and lower model-driven limits of detection and quantification are also given.

	dTTP versus Na_2SO_4		dTTP versus H_2O	
Parameter	**1350/945**	**1635/945**	**1350/3400**	**1635/3400**
FXJ	(3.9, 3.8–4.0)E2	(5.7, 5.4–6.0)E2	(4.2, 4.2–4.3)E3	(6.2, 6.2–6.3)E3
FXB	(9.1, 0.7–38)E−15	(8.0, 7.8–8.1)E−1	0	(1.5, 1.5–1.5)E−2
JA	(2.7, 2.0–3.8)E0	(3.4, 2.5–4.6)E0	(5.1, 0.0–13)E1	(2.7, 1.9–3.6)E2
FXJ / JA	(1.4, 1.0–1.9)E2	(1.7, 1.3–2.3)E2	(8.4, 3.3–12E12)E1	(2.3, 1.7–3.2)E1
$LL_D - UL_D$	0–4.1E−1	7.9E−5–3.4E−1	0–5.2E−3	3.0E−8–4.3E−3
$LL_Q - UL_Q$	0–2.6E−1	1.3E−4–2.1E−1	0–4.1E−3	6.2E−8–2.5E−3

TABLE 3 | Estimated model parameters (median, 16th–84th percentiles) for mixtures of different nucleotides in water, evaluated for each nucleotide's strongest peak vs. that of H_2O.

	dATP versus H_2O	dCTP versus H_2O	dGTP versus H_2O	dTTP versus H_2O
Parameter	1310/3400	1505/3400	1460/3400	1350/3400
FXJ	(1.6, 1.6–1.6)E+4	(6.5, 6.2–6.7)E+3	(2.6, 2.6–2.6)E+4	(4.2, 4.2–4.3)E+3
FXB	0	(2.7, 2.6–2.8)E-3	(9.6, 8.1–11)E-4	0
JA	(2.8, 0.3–3E14)E-12	(1.0, 0.1–4.9)E-13	(4.1, 2.4–5.8)E+2	(5.1, 0.0–13)E+1
FXJ / JA	(5.7, 0.0–61)E+15	(6.3, 1.3–63)E+16	(6.4, 4.5–11)E+1	(8.4, 3.3–12E12)E+1
$LL_D - UL_D$	0–1.7E−03	4.9E−8–4.2E+11	1.2E−8–2.5E−3	0–5.2E−3
$LL_Q - UL_Q$	0–1.7E−03	8.6E−8–3.7E+11	2.6E−8–1.3E−3	0–4.1E−3

linearity, with median limit parameters $\widehat{FXB} = 0$ and $\frac{\widehat{FXJ}}{\widehat{JA}} \approx 84$, indicating minimal cross-contributions from dTTP and H_2O at 3400 and 1350 cm^{-1} respectively. Because $\widehat{FXB}$ was evaluated to be 0 with 0 uncertainty, the model has a lower limit of 0. However, it should be remembered that these model-driven limits only indicate the maximum possible range over which a given model can be applied to ideal test data and that applying it to measurements with nonzero levels of background noise or uncertainty will always lead to more conservative ranges of detection/quantification.

Leave-one-out cross-validation was used to test the validity of each model and determine its effectiveness: The model was trained multiple times on $N-1$ available data points and then tested on the remaining data point. There was little variation in median model parameters during cross-validation, suggesting that the median model was well fitted to the overall dataset and not sensitive to outliers in the data (see Table S2). The predicted concentration ratios for each test sample were also relatively accurate, with an average RMSPE of 0.098 orders of magnitude (0.80–1.25 × truth). Critically, the RMSPE of the test dataset is comparable with that of the training dataset, suggesting that the models are not over-fitted to the training data.

The quantitative model was also trained on mixtures of three other organic nucleotides in water, namely, deoxyadenosine triphosphate (dATP), deoxycytidine triphosphate (dCTP), and deoxyguanosine triphosphate (dGTP), across nucleotide: H_2O concentration ratios of $10^{-7} - 10^{-3}$. Each nucleotide has its most distinctive peak at a slightly different frequency, 1310 cm^{-1} for dATP, 1505 cm^{-1} for dCTP, and 1460 cm^{-1} for dGTP (see Figures S18–S20). Figure 7 shows the resulting intensity ratios and trained models for each nucleotide's best peak when evaluated versus the 3400 cm^{-1} H_2O peak, with all four mixtures producing roughly linear models that differ only in $\widehat{FXJ}$, which varies from 4.2×10^3 for dTTP to 2.6×10^4 for dGTP (Table 3). This is consistent with experimental observations that, of the four nucleotide solutions, dGTP produces the largest intensity ratio at a given concentration ratio and dTTP the smallest.

The trained models in Figure 7 can be used to estimate the maximum possible concentration of each nucleotide that could exist in a test sample without being detected. Suppose that a test solution was measured such that the 3400 cm^{-1} H_2O peak had an intensity of 500 counts and background noise was ± 1 count. The resulting limits of detection would be 2.9×10^{-7} for dATP, 7.2×10^{-7} for dCTP, 2.6×10^{-7} for dGTP, and 1.1×10^{-6} for dTTP. Assuming no other compounds are present at significant concentrations, these concentration ratios equate to absolute nucleotide concentrations of 16, 40, 14, and 61 μmol/L, respectively.

4 | Discussion

The quantitative model provides a theoretical basis for determining concentration ratios from measured intensities of a

given binary mixture, provided that both components are detectable to Raman and the model is first trained on experimental data from samples of known concentration ratio. Analysis of experimental data from aqueous solutions demonstrates that the model can effectively describe observed trends in Raman intensity ratios, even when there are cross-contributions that introduce nonlinear behaviors. It shows that the ratio-driven model automatically accounts for effects such as interrogation volume: Optical modelling of these solutions in a previous study demonstrated that the combination of the laser's focal plane with the nucleotides' strong self-absorption led to significant decreases in interrogation volume and absolute Raman intensities at high concentrations [15]. These effects are not seen in the intensity ratio data shown in Figure 5, which follow the expected behavior predicted by the quantitative model as variations in (weighted) interrogation volume are effectively cancelled out when intensities are expressed in terms of ratios. While it is therefore implied that the model should also handle situations where the medium is the source of self-absorption, further work will be required to ascertain just how effective the model is at handling highly turbid or solid mixtures that exhibit substantial subsurface scattering, which will be addressed in a subsequent study.

4.1 | Application

As the quantitative model must be trained on experimental data for a given binary mixture of A and B, a trained model will be specific to that pair. To evaluate multiple possible mixtures, multiple models must be trained on different datasets representing each pair. Experimental data should be collected using as many samples of known concentration ratio as is practical to prepare. The concentration ratios prepared should span multiple orders of magnitude to capture a broad range of possible $\widehat{FXB}$ and $\frac{\widehat{FXJ}}{\widehat{JA}}$ values and should always encompass the range of concentration ratios that are anticipated in test samples to avoid risks associated with extrapolation. Similarly, the calibration data used to train the model should be acquired using instrument settings that provide the best possible signal:noise ratio to reduce the impact of noise on estimated parameters, provided there is no evidence of spectral degradation due to overexposure and/or photo-bleaching of molecules in the sample. If significant background noise is present in training data, it may lead to overestimation of $\widehat{FXB}$ and an underestimation of concentration ratio at very low intensity ratios. While the model can be trained on a single measured spectrum per sample, 10+ measurements per sample are recommended to permit bootstrapped model training—this provides a more accurate estimation of median $\widehat{FXJ}$, $\widehat{FXB}$, and $\widehat{JA}$ and enables calculation of uncertainties and limits of detection/quantification for the trained model. Cross-validation can then be used to evaluate the performance of the model when predicting concentration ratio without compromising the data volume used for model training.

Raman intensities should be measured the same way for both training data and any subsequent test data the model is applied to. If absolute intensities at specific frequencies are used, then the same frequencies should be evaluated in test data; if integrated peak areas are used, then the same ranges should be integrated in test data (see Appendix B). Such information should always be recorded alongside the trained model parameters to ensure correct usage.

Making predictions using a trained model will be most robust when evaluated on multiple measurements of the test sample, providing a median prediction and distribution that also reflects the uncertainty in test measurement; however, the model can still be applied to a single measurement if necessary (see Appendix B). For spectra where multiple identities of A and/or B are potentially valid, for example, peak I_1 is observed at a frequency consistent with two different molecules, the intensity ratio can be evaluated using each appropriate model to predict the concentration ratio for that binary pair. The results from each model can then be presented alongside one another as hypotheses for each possible mixture. As the model will inevitably be trained on imperfect experimental data containing some degree of noise, there will be some uncertainty and degree of error in the estimated values of FXJ, FXB, and JA. Errors can include the prediction of negative concentration ratios (which cannot be true) when a trained model is applied to intensity ratio data outside the expected range $\widehat{FXB} < \frac{I_{v1}}{I_{v2}} < \frac{\widehat{FXJ}}{\widehat{JA}}$. Consequently, any prediction of concentration ratio should always be done with consideration of the model's inherent limits of detection and quantification, L_D and L_Q.

4.2 | Implicit Assumptions

The model is built on equations that come with several implicit assumptions, the first being that the intensities I_1 and I_2 are exclusively due to Raman scattering from molecular species A and B, and not any other component C. If the test sample is a ternary mixture of A, B, and C, the model can still produce meaningful results provided C does not significantly modulate evaluated intensities by modifying X through selective attenuation of one measured peak over another. Adapting the model to directly model ternary mixtures is a priority for future work.

The second assumption is that all factors influencing measured Raman intensity are accounted for in the empirical parameters FXJ, FXB, and JA, even if the factor was not included in the derivation from first principles. This is convenient for accurate prediction of concentration from intensity, as any phenomena that systematically alter Raman intensities (e.g., solvation or mixing effects) should be represented in the experimental spectra used to train the model, and therefore, the model parameters will incorporate the effects of those phenomena. However, it does mean that the sample- or instrument-specific properties such as J, X, or F cannot be extrapolated from $\widehat{FXJ}$, $\widehat{FXB}$, $\widehat{JA}$ in the absence of additional context (e.g., an instrument response curve).

The third assumption is that molecules of A and B are distributed such that the concentration ratio $\frac{C_A}{C_B}$ is roughly constant throughout the weighted interrogation volume. This volume is defined as the region that is both illuminated by the laser (either directly or indirectly) and can be observed by the collection optics (directly or indirectly), weighted by how much incident light is received and how much scattered light can be transmitted to

the detector depending on location. This means A and B can be distributed heterogeneously provided they are always proportional to each other. For samples where A and B have dissimilar distributions, such as a pure layer of A sitting on top of a pure layer of B, molecules of A and B will experience very different attenuation environments and the resulting intensity ratio $\frac{I_{v1}}{I_{v2}}$ may deviate from what would be expected for an proportionately distributed mixture.

4.3 | Comparison With Other Studies

The model described in this study presents several advantages over previous models in the literature. First, the logarithmic regression method can handle data spanning several orders of magnitude without biasing the results to higher absolute values, whereas previous studies have typically been limited to linear regression spanning just 1–2 orders of magnitude [12, 19, 21]. Second, it can fully account for nonlinear effects and provide quantitative analysis for a greater variety of mixtures, even those that exhibit appreciable cross-contributions due to overlapping peaks. This is demonstrated in Figure 5, where the intensity ratio for the dTTP/H_2O peak pair at 1635 and 3400 cm^{-1} flattens out at very low concentrations due to overlap with the H_2O peak at 1620 cm^{-1}; evaluating this data using a conventional linear model would systematically overestimate the relative concentration of dTTP at very low concentrations, potentially by multiple orders of magnitude. Third, the use of bootstrap resampling to obtain many fitted models allows the estimation of uncertainty when making predictions, and the calculation of limits of detection and quantification for that model and mixture. For samples where a particular peak is not detected, the model can still provide valuable information on possible compositions by estimating the maximum concentration ratio that a given molecule could have based on the available data.

5 | Conclusions

The quantitative model described in this study was derived from first principles to fully describe the nonlinear behavior of Raman intensity ratios observed in binary mixtures, accounting for potential cross-contributions by overlapping peaks. This enables a more robust analysis of composition over a wider range of concentration ratios, and for a greater variety of mixtures, even ones that exhibit appreciable overlap. Combined with bootstrapped model training, the most robust model possible can be obtained even from noisy input data along with full propagation of uncertainty to any predictions made, as well as calculation of important parameters such as the limits of detection and quantification. When applied to experimental data on aqueous solutions of four different nucleotides, the trained model can correctly predict organic/water concentration ratios to within 0.1 orders of magnitude for samples ranging from 10^{-7} (0.1 ppm) to 10^{-3} (100 ppm). Overall, this represents a significant development over previous attempts at quantitative or semiquantitative Raman spectroscopy, which have historically relied on linear calibrations that do not account for cross-contributions and were evaluated over narrower concentration ranges.

The ability to estimate relative concentrations over several orders of magnitude will be invaluable to the study of many different mixtures using Raman spectroscopy, across a wide range of scientific disciplines. For truly unknown samples that may contain one or more different compounds of interest, the model can be used to assess the concentration of any potential mixture, if the appropriate models have been trained on experimental data of known samples. Even when key peaks are not observed, the corresponding limits of detection and quantification can be calculated for each potential mixture to assess the maximum possible concentrations that would not be detectable in a given measurement. All the algorithms used in this study have been made publicly available as self-contained Jupyter Notebooks with detailed documentation on their use to encourage the application of this methodology by other researchers [23]. These will continue to be updated as the quantitative model is developed further.

Acknowledgements

I would like to thank Dr. Matthew Razzelliot for valuable discussion regarding the mathematics of the model equations. No large language models were used during this study or in the creation of this manuscript.

Funding

Experimental data were originally collected at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration, as part of a project that was funded by a NASA Postdoctoral Program fellowship awarded to Joseph Razzell Hollis, administered by the Universities Space Research Association on behalf of NASA.

Conflicts of Interest

The author declares no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are openly available in Zenodo at 10.5281/zenodo.18338662. The quantitative modelling code is openly available in Zenodo at 10.5281/zenodo.16412759.

References

1. L. Bonal, E. Quirico, G. Montagnac, et al., "The Thermal History of Ryugu Based on Raman Characterization of Hayabusa2 Samples," *Icarus* 408 (2024): 115826, https://doi.org/10.1016/j.icarus.2023.115826.

2. W. P. Griffith, "Raman Spectroscopy of Minerals," in *The Infrared Spectra of Minerals*, ed. V. C. Farmer (Mineralogical Society of Great Britain and Ireland, 1974), 119–135, https://doi.org/10.1180/mono-4.8.

3. K. Hickman-Lewis, K. R. Moore, J. J. Hollis, et al., "In Situ Identification of Paleoarchean Biosignatures Using Colocated Perseverance Rover Analyses: Perspectives for In Situ Mars Science and Sample Return," *Astrobiology* 22, no. 9 (2022): 1143–1163, https://doi.org/10.1089/ast.2022.0018.

4. J. Jehlička and H. G. M. Edwards, "Raman Spectroscopy as a Tool for the Non-Destructive Identification of Organic Minerals in the Geological Record," *Organic Geochemistry* 39, no. 4 (2008): 371–386, https://doi.org/10.1016/j.orggeochem.2008.01.005.

5. C. P. Marshall, H. G. M. Edwards, and J. Jehlicka, "Understanding the Application of Raman Spectroscopy to the Detection of Traces of

Life," *Astrobiology* 10, no. 2 (2010): 229–243, https://doi.org/10.1089/ast.2009.0344.

6. J. Razzell Hollis, K. Moore, M. Fries, et al., "Mineralogical and Chemical Mapping of Martian Meteorite SaU 008 Using Deep UV Raman and Fluorescence Spectroscopy on Earth and Mars," *JGR Planets* 130 (2025): e2024JE008826, https://doi.org/10.1029/2024JE008826.

7. H. M. Sapers, J. Razzell Hollis, R. Bhartia, L. W. Beegle, V. J. Orphan, and J. P. Amend, "The Cell and the Sum of Its Parts: Patterns of Complexity in Biosignatures as Revealed by Deep UV Raman Spectroscopy," *Frontiers in Microbiology* 10, no. March 2019 (2019): 1–15, https://doi.org/10.3389/fmicb.2019.00679.

8. A. Steele, F. M. McCubbin, and M. D. Fries, "The Provenance, Formation, and Implications of Reduced Carbon Phases in Martian Meteorites," *Meteoritics and Planetary Science* 51, no. 11 (2016): 2203–2225, https://doi.org/10.1111/maps.12670.

9. R. Bhartia, L. W. Beegle, L. Deflores, et al., "Perseverance's Scanning Habitable Environments With Raman and Luminescence for Organics and Chemicals (SHERLOC) Investigation," *Space Science Reviews* 217 (2021): 58, https://doi.org/10.1007/s11214-021-00812-z.

10. S. Maurice, R. C. Wiens, P. Bernardi, et al., "The SuperCam Instrument Suite on the Mars 2020 Rover: Science Objectives and Mast-Unit Description," *Space Science Reviews* 217, no. 3 (2021), 47, https://doi.org/10.1007/s11214-021-00807-w.

11. D. A. Long, *Raman Spectroscopy* (McGraw Hill, 1977).

12. M. J. Pelletier, "Quantitative Analysis Using Raman Spectrometry," *Applied Spectroscopy* 57, no. 1 (2003): 20A–42A, https://doi.org/10.1366/000370203321165133.

13. X. Qi, Z. Ling, P. Liu, et al., "Quantitative Mineralogy of Planetary Silicate Ternary Mixtures Using Raman Spectroscopy," *Earth and Space Science* 10, no. 5 (2023): e2023EA002825, https://doi.org/10.1029/2023EA002825.

14. S. A. Asher, "Ultraviolet Resonance Raman Spectrometry for Detection and Speciation of Trace Polycyclic Aromatic Hydrocarbons," *Analytical Chemistry* 56, no. 4 (1984): 720–724, https://doi.org/10.1021/ac00268a029.

15. J. Razzell Hollis, D. Rheingold, R. Bhartia, and L. W. Beegle, "An Optical Model for Quantitative Raman Microspectroscopy," *Applied Spectroscopy* 74, no. 6 (2020): 684–700, https://doi.org/10.1177/0003702819895299.

16. Z. Wu, C. Zhang, and P. C. Stair, "Influence of Absorption on Quantitative Analysis in Raman Spectroscopy," *Catalysis Today* 113, no. 1–2 (2006): 40–47, https://doi.org/10.1016/j.cattod.2005.11.077.

17. P. J. Aarnoutse and J. A. Westerhuis, "Quantitative Raman Reaction Monitoring Using the Solvent as Internal Standard," *Analytical Chemistry* 77, no. 5 (2005): 1228–1236, https://doi.org/10.1021/ac0401523.

18. S. C. Park, M. Kim, J. Noh, et al., "Reliable and Fast Quantitative Analysis of Active Ingredient in Pharmaceutical Suspension Using Raman Spectroscopy," *Analytica Chimica Acta* 593, no. 1 (2007): 46–53, https://doi.org/10.1016/j.aca.2007.04.056.

19. L. Demaret, I. B. Hutchinson, G. Eppe, and C. Malherbe, "Quantitative Analysis of Binary and Ternary Organo-Mineral Solid Dispersions by Raman Spectroscopy for Robotic Planetary Exploration Missions on Mars," *Analyst* 146, no. 23 (2021): 7306–7319, https://doi.org/10.1039/D1AN01514A.

20. T. Dörfer, W. Schumacher, N. Tarcea, M. Schmitt, and J. Popp, "Quantitative Mineral Analysis Using Raman Spectroscopy and Chemometric Techniques," *Journal of Raman Spectroscopy* 41, no. 6 (2010): 684–689, https://doi.org/10.1002/jrs.2503.

21. M. Veneranda, J. A. Manrique-Martinez, C. Garcia-Prieto, et al., "Raman Semi-Quantification on Mars: ExoMars RLS System as a Tool to Better Comprehend the Geological Evolution of Martian Crust," *Icarus* 367 (2021): 114542, https://doi.org/10.1016/j.icarus.2021.114542.

22. K. Uckert, R. Bhartia, and J. Michel, "A Semi-Autonomous Method to Detect Cosmic Rays in Raman Hyperspectral Data Sets," *Applied Spectroscopy* 73, no. 9 (2019): 1019–1027, https://doi.org/10.1177/0003702819850584.

23. J Razzell Hollis, "Jobium/Quantitative-Raman: Pre-Publication Release, January 2026," Zenodo, (2026), https://doi.org/10.5281/zenodo.16412759.

24. J. Razzell Hollis, "Jobium/OSTRI: v0.1.2-Alpha Release," Zenodo, (2025), https://doi.org/10.5281/ZENODO.15535393.

25. M. Newville, T. Stensitzki, D. B. Allen, and A. Ingargiola. "LMFIT: Non-Linear Least-Square Minimization and Curve-Fitting for Python," (Zenodo, 2014), https://zenodo.org/record/11813.

26. J. J. Davenport, J. Hodgkinson, J. R. Saffell, and R. P. Tatam, "Noise Analysis for CCD-Based Ultraviolet and Visible Spectrophotometry," *Applied Optics* 54, no. 27 (2015): 8135–8144, https://doi.org/10.1364/AO.54.008135.

27. J. M. Smulko, N. C. Dingari, J. S. Soares, and I. Barman, "Anatomy of Noise in Quantitative Biological Raman Spectroscopy," *Bioanalysis* 6, no. 3 (2014): 411–421, https://doi.org/10.4155/bio.13.337.

28. C. Salter, "Error Analysis Using the Variance-Covariance Matrix," *Journal of Chemical Education* 77, no. 9 (2000): 1239, https://doi.org/10.1021/ed077p1239.

29. J. Tellinghuisen, "Statistical Error Propagation," *Journal of Physical Chemistry A* 105, no. 15 (2001): 3917–3921, https://doi.org/10.1021/jp003484u.

30. P. Virtanen, R. Gommers, T. E. Oliphant, et al., "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python," *Nature Methods* 17, no. 3 (2020): 261–272, https://doi.org/10.1038/s41592-019-0686-2.

31. B. Efron, "Bootstrap Methods: Another Look at the Jackknife," *Annals of Statistics* 7, no. 1 (1979): 1–26, https://doi.org/10.1214/aos/1176344552.

32. K. Ben Mabrouk, T. H. Kauffmann, H. Aroui, and M. D. Fontana, "Raman Study of Cation Effect on Sulfate Vibration Modes in Solid State and in Aqueous Solutions," *Journal of Raman Spectroscopy* 44, no. 11 (2013): 1603–1608, https://doi.org/10.1002/jrs.4374.

33. D. M. Carey and G. M. Korenowski, "Measurement of the Raman Spectrum of Liquid Water," *Journal of Chemical Physics* 108, no. 7 (1998): 2669–2675, https://doi.org/10.1063/1.475659.

34. Z. Q. Wen and G. J. Thomas, "UV Resonance Raman Spectroscopy of DNA and Protein Constituents of Viruses: Assignments and Cross Sections for Excitations at 257, 244, 238, and 229 nm," *Biopolymers* 45, no. 3 (1998): 247–256, https://doi.org/10.1002/(SICI)1097-0282(199803)45:3<247::AID-BIP7>3.0.CO;2-R.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Appendix S1:** Appendix A: Derivation of the quantitative model from first principles. Appendix B Numerical application of the model. Additional supplementary figures and tables.