RESEARCH ARTICLE

VOLCANICA

Evaluating and optimizing Raman analytical protocols for CO_2 fluids

Penny E. Wieser *[α], Charlotte L. DeVitre[α,β], and Isabelle Susman[α]

[α] Earth and Planetary Sciences, University of California, Berkeley, CA 94270, USA.
[β] Department of Earth and Environmental Sciences, University of Ottawa, Ottawa, ON K1N 6N5, Canada.

Submitted: 2025-08-14 | Accepted: 2026-08-07 | Published: 2026-08-24

Editor: W. Li | Production: A. Ryan

Abstract

Raman spectroscopy is widely used to quantify CO_2 density in geological fluids, yet acquisition parameters, spectral processing methodologies and densimeter calibrations vary between laboratories. The precision of Raman CO_2 measurements, and the broader applicability of densimeters generated for a specific instrument, are still debated. This study describes a series of experiments to investigate the effect of different instrument and acquisition parameters on CO_2 densities and their errors that can be performed on different Raman instruments, allowing users to determine analytical best practices. First, we discuss the effect of spectral non-linearity, demonstrating that shifts in the selected acquisition window and the choice of Ne lines used for drift correction can generate a diversity of densimeters on a single instrument. However, our repeat measurements of natural fluid inclusion (FI) standards over ~3 years yield a 1σ variation of <0.01 g cm^{-3}, indicating that when consistent methodologies are applied to calibration data and unknowns, Raman analyses are highly reproducible over many years. We then investigate the effect of parameters such as acquisition time, laser power, and sample positioning on peak fitting errors, rebutting recent suggestions that Raman analyses are associated with large uncertainties in Fermi-diad splitting. We show that although high laser powers greatly improve signal:noise ratios (and thus reduce peak fitting errors), natural fluid inclusions hosted in olivine show changes in calculated density outside analytical error with increasing powers (density changing by ~0 to 0.09 g cm^{-3} per 10 mW of increased power). The amount of heating of the trapped fluid, and thus the change in density, varies drastically on a sample-by-sample basis, with larger amounts of heating in materials hosting the fluid with high absorption coefficients (melt inclusions, and low Fo olivines). We recommend that users assess the degree of laser heating experienced by their samples and select a laser power that limits density changes to a level that will not compromise the objectives of their study. We also offer practical strategies to optimize signal without raising laser power, noting that small changes in focus and x-y position can greatly increase signal intensity.

Keywords: Raman Spectroscopy; Melt Inclusion; Fluid Inclusion; Volatiles.

1 Introduction

Traditionally, the density of CO_2-rich fluids in geological samples have been measured using microthermometry, a technique in which phase changes are observed during heating and cooling of trapped inclusions. Starting in the 1930s, it was recognised that CO_2 fluid/gas/vapour has a distinct signal when analysed by Raman spectroscopy, with two strong peaks at ~1285 and ~1388 cm^{-1} [Fermi 1931; McLennan and Smith 1932]. These peaks are collectively called the Fermi diad. For brevity, we refer to the lower wavenumber peak of CO_2 as diad1 (also v^-), and the higher wavenumber peak as diad2 (also v^+). In the 1970s, the distance between the two peaks—referred to as the Fermi diad separation, splitting, peak splitting or diad splitting—was shown to be related to the density of CO_2 [Wright and Wang 1973]. These findings opened up the possibility of using Raman spectroscopy to quantify the density of CO_2 fluids, an important parameter for many geological research questions [Rosso and Bodnar 1995].

The relationship between diad splitting and CO_2 density on a given Raman instrument is described using a densimeter, which is typically modelled as a polynomial function of Nth degree. A number of groups have developed experimental apparatus to determine densimeters for different instruments [Rosso and Bodnar 1995; Kawakami et al. 2003; Yamamoto and Kagi 2006; Song et al. 2009; Fall et al. 2011; Wang et al. 2011; Lamadrid et al. 2017; Wang et al. 2019; DeVitre et al. 2021, see Supplementary Material 1 Table S1]. Not only do densimeters vary widely between Raman instruments [e.g. Lamadrid et al. 2017; DeVitre et al. 2021]; changes in hardware on a single instrument have also been shown to affect the densimeter. For example, Lamadrid et al. [2017] show that the densimeter on a single instrument can be shifted by changing the laser wavelength (e.g. a ~0.02 g cm^{-3} shift from 514 nm to 632 nm using the 2400 grooves mm^{-1} grating) or the instrument grating (e.g. a 0.02–0.04 g cm^{-3} shift for the 600 to 2400 grating at 632 nm). They conclude that the same wavelength and grating must be used for calibration and analysis of unknowns.

In contrast, recent work has argued that densimeters may be more universally applicable than previously thought. Remigi et al. [2021] generate a calibration using natural fluid inclusion (FI) standards with densities between 0.37-1.08 g cm^{-3} and then compare their densimeter to published densimeters using bootstrapping techniques to better account for uncertainty in densimeter regressions. They suggest that published Raman densimeters fall into two main categories—those collected on instruments with higher spectral resolution (~0.4 cm^{-1}/pixel) and those collected on instruments with lower spectral resolution (~1–1.5 cm^{-1}/pixel). For CO_2 densities >0.3-0.4 g cm^{-3}, they show that published densimeters within the high- or low-resolution category are within error of each other. They suggest that this may be because spectral resolution is related to the ability of a Raman spectrometer to discriminate overlapping bands, and changes in spectral

* penny_wieser@berkeley.edu

resolution shifts peak center positions in a systematic way. They conclude that instrument-specific densimeters may not be required if a densimeter exists for the spectral resolution of interest.

Remigi et al. [2021] use the term *spectral resolution* to describe the wavelength range recorded in each pixel. For the remainder of this paper, we call this parameter pixel resolution. Following Hagiwara et al. [2023], pixel resolution differs from the *spectral resolution,* which is the smallest width of peak that can be resolved, typically estimated from the full width half maximum (fwhm) of very sharp emission lines such as Ne or Ar, or calcite. These two resolution parameters are not equivalent on many Raman instruments, and there is no simple way to deduce one parameter from the other. For example, using a 1200 grooves mm^{-1} grating, the pixel resolution reported by Hagiwara et al. [2023] (their Table 1) is 0.482 cm^{-1}/pixel while the spectral resolution estimated from the fwhm of the Ne emission line at 1449 cm^{-1} as 1.59 cm^{-1}. However, using the 1800 grooves mm^{-1}, they report a pixel resolution of 0.284 cm^{-1}/pixel vs. a spectral resolution of 0.29 cm^{-1}. In Hagiwara et al. [2020], the 1800 grooves mm^{-1} yields a pixel resolution of 0.29 cm^{-1}/pixel but a spectral resolution of 1.1 cm^{-1} estimated from the fwhm of calcite. The vast majority of studies do not report their spectral resolution (Supplementary Material 1 Table S1), which makes it very hard to assess if spectral or pixel resolution is controlling the densimeter equation, or whether the groupings of Remigi et al. [2021] appear by chance as a function of all the high-resolution densimeters having been collected on similar HORIBA LabRam HR instruments.

It is also notable that published densimeters show the largest relative differences for analyses of gaseous CO_2 at room temperature [$\rho \sim$ 0–0.2 $g\,cm^{-3}$; Lamadrid et al. 2017], while Remigi et al. [2021] only consider densities >0.3 $g\,cm^{-3}$. At low densities, densimeters can be offset by a factor of 2. This has major implications in igneous petrology, because many melt inclusion vapor bubbles have CO_2 densities below 0.2 $g\,cm^{-3}$ [see Figure 21 of Wieser et al. 2025]. Accurate measurements of these low density vapor bubbles are essential for reconstructing the CO_2 contents of magmas, with implications for determining magma storage depths and the cause of explosive eruptions [Moore et al. 2015; Allison et al. 2021; DeVitre et al. 2023b]. However, studies published after the Lamadrid et al. [2017] paper advocating for instrument-specific densimeters still report densities for CO_2 fluids with low densities using calibrations published in other laboratories [e.g. Venugopal et al. 2020; Anderson et al. 2024; Howe et al. 2025]. In this study, we critically evaluate sources of densimeter variability, to evaluate whether generic calibrations based on a simple instrument parameters can replace instrument-specific calibrations.

Another recent discussion of the utility of Raman spectroscopy for quantifying CO_2 densities was provided by Bakker [2021]. This work assessed the uncertainties associated with various parts of the Raman method (e.g. peak fitting, instrument drift, etc) and suggested that errors in the position of each diad propagate through to very large errors in diad splitting (±0.26–0.27 cm^{-1}) and thus CO_2 density (~±0.1 $g\,cm^{-3}$). The large quoted errors of Bakker [2021] are in contrast to the very low theoretical errors for analyses of CO_2 within a high pressure optical cell suggested by Yamamoto and Hagiwara [2024] (0.0005 cm^{-1}, 0.00015 $g\,cm^{-3}$). Thus, it is imperative to evaluate these contrasting statements, and determine the analytical errors associated with analyses of natural melt and fluid inclusions.

Overall, the Bakker [2021] and Remigi et al. [2021] papers demonstrate that while a lot of progress has been made in recent years optimizing Raman analyses of CO_2 fluids [Fall et al. 2011; Wang et al. 2011; Lamadrid et al. 2017; DeVitre et al. 2021; Le et al. 2021; DeVitre et al. 2023a], there are still some fundamental disagreements in the community regarding the applicability of densimeters, and the accuracy and precision of the Raman method when applied to natural samples. In this study, we investigate the effect of acquisition and data processing strategies on the precision and accuracy of CO_2 fluid measurements using a WITec alpha300 R in the Department of Earth and Planetary Science, University of California, Berkeley (hereafter, UCB WITec). We describe a series of tests that readers can apply to their own instruments to guide the development of laboratory-specific best practices, with the aim that this leads to a greater understanding in the community of which sensitivities are instrument and hardware dependent, and which are globally applicable.

2 Methods

We acquire Raman spectra on natural FI and melt inclusions (MI) hosted in olivine using the UCB WITec. The Ultra-High-Throughput-Spectrometer (UHTS600) was calibrated using a Hg-Ar lamp using the WITec control software. The CCD chip (1024×127 pixels) was thermoelectrically cooled to −60 °C. We use a green solid-state laser with a wavelength of 532.046 nm, and a Ne lamp in the beam path to correct for instrument drift. For measurements of the CO_2 Fermi diad, the 1800 grooves mm^{-1} grating yields a pixel resolution ranging from 0.57–0.51 cm^{-1}/pixel from wavenumbers of 1042–1592 cm^{-1}. The spectral resolution estimated from the width of Ne lines is ~0.75-0.86 cm^{-1} (Supplementary Material 1 Figure S1). To convert diad splitting into CO_2 density, we apply the densimeter developed for our instrument by DeVitre and Wieser [2024] using a gas cell calibration device [adapted from DeVitre et al. 2021]. This calibration line is valid for CO_2 densities spanning 0.0096–1.04 $g\,cm^{-3}$.

The laser was focused on the sample with a variety of objectives depending on the sample size and heating set up; an Olympus 40× objective (0.6 NA, 2.7–4 mm working distance), a Zeiss 50× objective (x0.55NA, 9.1 mm working distance), or a Zeiss 100× objective (0.95 NA, 4 mm working distance). Laser power was controlled using the WITec TruePower module, which measures the laser power in the optical fiber, and allows the user to enter values between 0–50 mW with ±0.1 mW precision. We independently verify the laser power after it has left the fibre optic and passed through the optical head using a ThorLabs S121C sensor attached to a PM100D2 optical power meter. There is a strong linear correlation between the measured power through different objectives, and the instrument TruePower, with intercepts close to 0 (Supple-

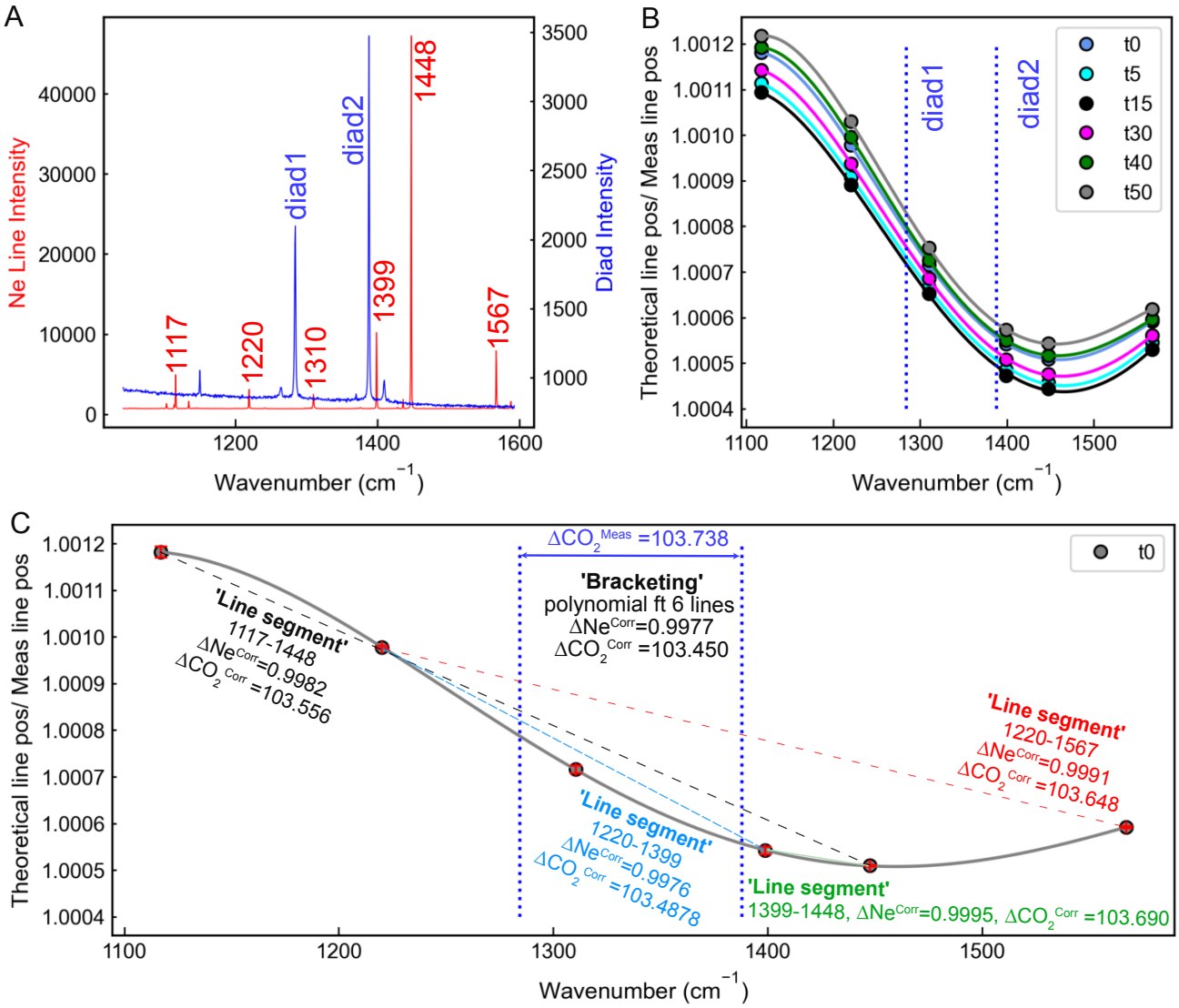


Figure 1: Examining the difference between the 'line segment' method and the 'bracketing method' when correcting instrument drift using Ne lines. [A] An example Ne line acquisition compared to a diad. Note that the intensity of the Ne (left-hand-axis) is almost 10× stronger than the diad (right-hand-axis). This is different from Figure 8 of Bakker [2021] where Ne lines show very low intensities. [B] Comparing the measured peak position of Ne to the theoretical peak position for repeated acquisitions of Ne lines through the day. For each line acquisition, a polynomial has been fitted through the data. [C] For the first acquisition (t0), we show the polynomial fit through the 6 Ne peaks. The 'bracketing technique' would correct each diad position based on the value for the polynomial at that wavenumber (where the blue dotted lines intersect the grey line). The 'Line segment' method will give similar results if the gradient drawn between two points is parallel to the intersection of the dotted blue and grey lines.

mentary Material 1 Figure S2). Placing the power meter at the opening of the objective turret with no objective screwed in yields laser powers that are 85 % of the TruePower value. For the 5-100× objective, we measure the power by focusing on the laser power meter in the normal sample position (i.e. after passing through 1 cm of air). This yields powers that are 80 % (5X), 79 % (20×), 77 % (40×), 62 % (50×) and 43 % (100×) of the TruePower value (Supplementary Material 1 Figure S2). Unlike our previous work examining the effects of laser power which report TruePower values [DeVitre et al. 2023a], in this study all subsequent mention of laser power uses the laser power at the sample, calculated using the correction factors mentioned above:

$$\text{Power}_{\text{sample}} = \text{Power}_{\text{TruePower}} \times \text{loss} \quad (1)$$

For example, if we set the TruePower value to 10 mW in the software, we report a power of 4.3 mW for 100× (10×0.43). This allows greater comparison between studies performed on different instruments.

Various methods were used to heat samples above the critical temperature of CO_2 to ensure a single homogenous fluid was being analysed [DeVitre et al. 2023a]. For most analy-

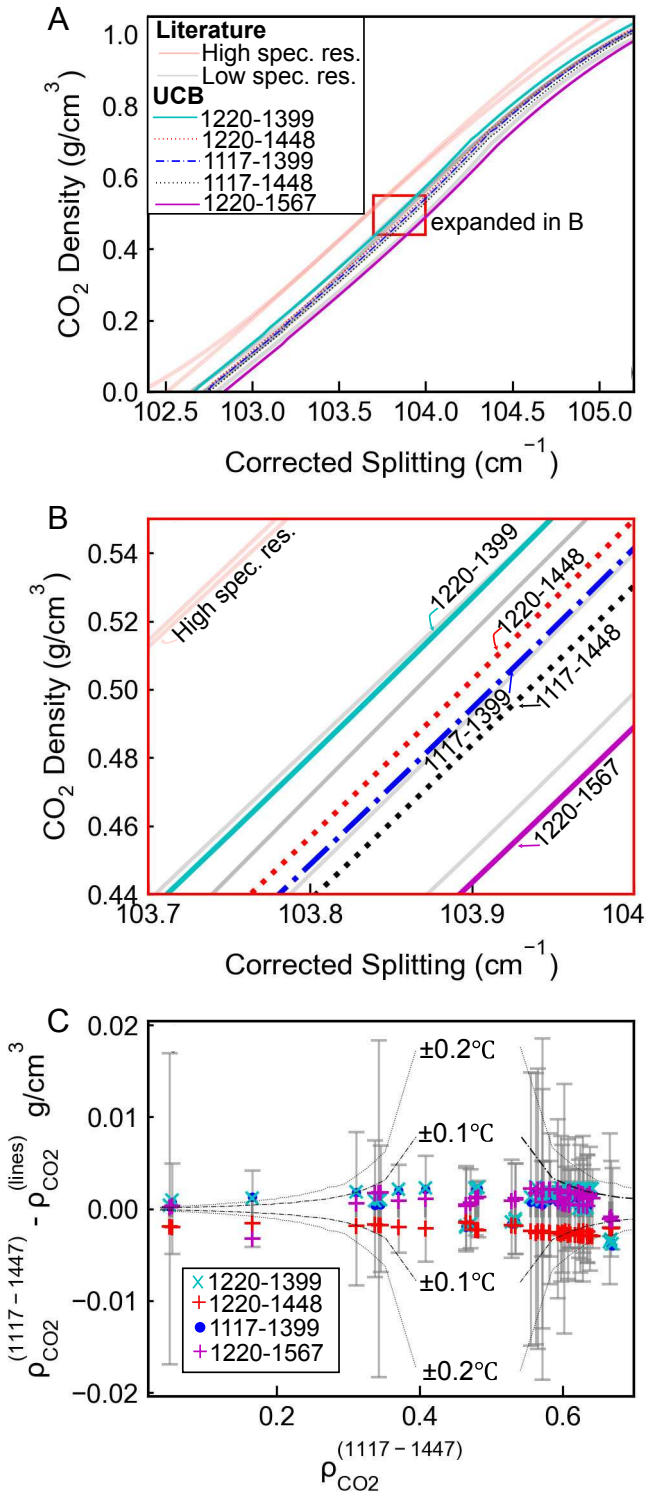


Figure 2: Assessing the effect of different Ne line reduction routines on calculated densities. [A] The relationship between corrected splitting and CO_2 density (i.e. the densimeter) was determined using 5 different pairs of Ne lines for the same UCB gas cell calibration dataset. This variation is compared to other literature densimeters (grey = densimeters calibrated using finer pixel sampling of approximately 0.30–0.43 cm^{-1}/pixel, termed 'high spectral resolution' by Remigi et al. [2021]; salmon = densimeters calibrated using coarser pixel sampling of approximately 1.50–1.66 cm^{-1}/pixel, termed 'medium spectral resolution' by Remigi et al. [2021]). [B] Inset of area in red box in [A]. [C] Natural FI data from DeVitre and Wieser [2024] was reduced using these different Ne correction routines and densimeters. The y-axis shows the difference in density for the default 1117–1447 line option compared with other data reduction strategies. The grey error bar shows the uncertainty on the 1117–1447 density. This demonstrates densities are within uncertainty for different Ne line pairs as long as spectra from unknowns and calibration data are drift-corrected in the same way. For reference, we overlay the uncertainty associated with microthermometry for ±0.1 and 0.2 °C in sample temperature.

ses, samples were placed on a glass slide in crystalbond and placed on a Peltier heating stage mounted within a custom aluminium holder with a central hole for transmitted light illumination (Supplementary Material 1 Figure S3). Temperature was monitored using a type K thermocouple taped to the slide with electrical tape. The voltage to the peltier plate was adjusted to yield a sample temperature of 37 °C, matching the gas cell calibration temperature for this instrument [DeVitre and Wieser 2024]. When sample temperature was being directly investigated (Section 3.6), crystals were double polished and placed in a Linkam THMSG600 stage (Supplementary Material 1 Figure S4), where the crystal and thus fluid temperature can be read from the thermocouple output. To achieve the long working distance required when the stage lid is in place, we initially made measurements using the LWD 40× Olympus objective with a cover slip correction collar. After doing tests with and without the top lid, we found that the stage lid cut out ~50 % of the signal, and that there was no difference in sample heating between heating series performed with and without the stage lid (Supplementary Material 1 Figure S5). Thus, later analyses were performed using the Zeiss 50× objective without the stage lid, but with the small silver cap still placed on top of the sample to help with temperature stability (Supplementary Material 1 Figure S4).

All peak fitting and data processing was performed in the open-source `Python3` tool `DiadFit` [DeVitre and Wieser 2024] using pseudovoigt peaks, except when investigating peak fit shapes. Versions <v1.0.18 of `DiadFit` prior to March 2026 use an unweighted least squares routine to fit diad peaks, resulting in each pixel contributing equal weight to the fit. As part of this work, we revise `DiadFit` (>v.1.0.19+) to use weighted least squares, with weights assigned to each pixel proportional to the inverse square root of the background-subtracted intensity, consistent with the expected errors from Poisson counting statistics. While Ne lines can now be fit with either routine in `DiadFit`, we elect to continue using unweighted least squares for Ne lines for consistency with all prior data collected in our lab, as switching Ne fitting routines causes a small systematic shift in the densimeter. However, as Section 3.1 shows, as long as consistency is used when fitting Ne lines between calibration and sample, the exact choice of correction method doesn't matter. Uncertainties in peak fit parameters are estimated from the covariance matrix, with the `lmfit` option `scale_covar=True` applied to rescale parameter uncertainties

such that the reduced chi-squared value is unity. The peak positions do not change within the original uncertainties estimated by v<1.0.18, but the errors on each peak position range from 10 to 80 % higher than estimates using unweighted least squares (Supplementary Material 1 Figure S6-9).

3 Results and Discussion

3.1 Neon line corrections and spectral non-linearity

Instrument drift during a Raman session is commonly assessed using repeated analyses of the spectral emission of neon generated by a Ne lamp [e.g. Lamadrid et al. 2017; DeVitre et al. 2021, Figure 1A]. Two emission lines bracketing the CO_2 Fermi diad are selected [the line segment method, Bakker 2021]. The theoretical distance between these lines is calculated from the published wavelength of the *observed* emission spectra of Ne in air from the NIST webbook*, converted to wavenumber using the wavelength of the laser outputted by the WITec software (a fixed value for each instrument):

$$\widetilde{\nu}_{\text{line1}} = 10^7\left(\frac{1}{\lambda_o} - \frac{1}{\lambda_{\text{line1}}}\right); \quad (2)$$

$$\widetilde{\nu}_{\text{line2}} = 10^7\left(\frac{1}{\lambda_o} - \frac{1}{\lambda_{\text{line2}}}\right); \quad (3)$$

where $\widetilde{\nu}_{\text{line1}}$ is the wavelength of Ne line 1, λ_{line1} is the Ne emission wavelength of line 1, and λ_o is the laser wavelength. No vacuum to air conversions were applied to the stated laser wavelength. We note that the choice of laser wavelength cancels out when using the line segment method, so knowing this value precisely does not affect the correction factor:

$$\Delta\widetilde{\nu}_{\text{Ne}} = \widetilde{\nu}_{\text{line2}} - \widetilde{\nu}_{\text{line1}} = 10^7\left(\frac{1}{\lambda_{\text{line1}}} - \frac{1}{\lambda_{\text{line2}}}\right). \quad (4)$$

The calculated theoretical distance between two Ne lines is then compared to the measured distance in the acquired spectra. For example, if the theoretical distance between Ne lines is 330.4776 cm^{-1}, and the measured distance is 331.0587 cm^{-1}, the correction factor is $\frac{\text{theoretical distance}(330.4776)}{\text{measured distance}(331.0587)} = 0.9982$. If the diad splitting was measured at 103 cm^{-1}, the corrected diad splitting would be $0.9982 \times 103 = 102.8192$ cm^{-1}.

It has been suggested that the common method of correcting for instrument drift using the distance between two Ne lines [e.g. Lamadrid et al. 2017] does not fully capture the complex non-linearity of Raman spectrometers [Bakker 2021]. Non-linearity arises from the mapping between detector pixel position and wavelength, which is typically 'calibrated' using numerous peaks of known position. On our WITec instrument, this calibration is performed by the WITec control software using a Hg-Ar lamp. Ideally, if spectrometer calibration was perfect, the measured wavelengths of all reference spectrum [e.g. neon and argon lamps, benzonitrile, 4-aceamidophenol etc, Liu and Hennelly 2024] would match the true wavenumbers of these peaks within peak fitting error. However, non-linearity is the norm rather than the exception, and Raman spectra show clear offsets between measured and preferred peak positions, with the offset often varying in linear or sinusoidal patterns as a function of wavenumber [Bakker 2021; Liu and Hennelly 2024, Figure 1B].

We investigate spectrometer non-linearity, and the effect of this non-linearity on instrument drift and calculated densities using data collected on the UCB WITec, and two additional alpha 300R spectrometers at Cornell using data from DeVitre et al. [2021]. On the UCB instrument, we first perform 50 consecutive repeated acquisitions of the emission spectrum of Neon. We fit pseudovoigt curves to the 6 major Ne lines within the spectral window for each acquisition using `DiadFit`. For brevity, we refer to these lines using the calculated wavenumber positions rounded to the nearest integer for the 532.046 nm laser wavelength used in this study [565.66588 nm ≈ 1117 cm^{-1}, 568.98163 nm ≈ 1220 cm^{-1}, 571.92248 nm ≈ 1310 cm^{-1}, 574.82985 nm ≈ 1399 cm^{-1} 576.44188 nm ≈ 1447 cm^{-1}, and 580.44496 nm ≈ 1567 cm^{-1}, Figure 1A, Kramida and Ralchenko 1999]. For the line at 1117 cm^{-1}, we iteratively fit two pseudovoigt functions to resolve the smaller peak on the shoulder. Figure 1B shows the fitted Ne line position divided by the theoretical position calculated from the laser wavelength (532.046 nm) and the theoretical positions in air. This comparison demonstrates that the spectrometer is indeed non-linear, with the correction factor varying across the acquisition window (Figure 1B). Spectral non-linearity has been noted as a serious issue when trying to compare spectra collected between instruments, as it persists after performing the wavelength calibration methods built into Raman acquisition software [Liu and Hennelly 2024].

As Bakker [2021] correctly points out, the 'line segment' Ne correction method assumes that the correction factor at any given point between two lines is the same. They suggest that a 'bracketing technique' should be used instead, where a polynomial model is developed to describe how the position of each Ne line deviates from the theoretical position as a function of wavenumber. This polynomial allows a correction factor to be calculated for diad1 and diad2. These corrected peak positions are used to calculate the splitting, rather than applying a correction factor to the splitting as in the line-segment method.

The polynomial fit through these correction factors is shown for the 'bracketing technique' (grey curve, Figure 1C) for a single Ne acquisition. Correcting for the position of the CO_2 diad using this curve this would yield a corrected splitting for the diad position shown by blue dotted lines that would be equivalent to a line segment correction factor of 0.9977. The 'line segment' technique for drift correction can be visualized by drawing a straight line between any two datapoints (circles) on this polynomial curve (Figure 1C). If the gradient of this linear segment is the same as the gradient of a line drawn between where the x-positions of the diad intercept the polynomial, the line segment technique and bracketing techniques would give the same result. The line segment method applied to the 1220–1399 cm^{-1} lines results in a very similar correction factor to the bracketing technique (0.9976 vs. 0.9977). The disadvantage of choosing this pair is that the 1220 cm^{-1} line is relatively small, resulting in larger fitting errors, meaning

*https://webbook.nist.gov/

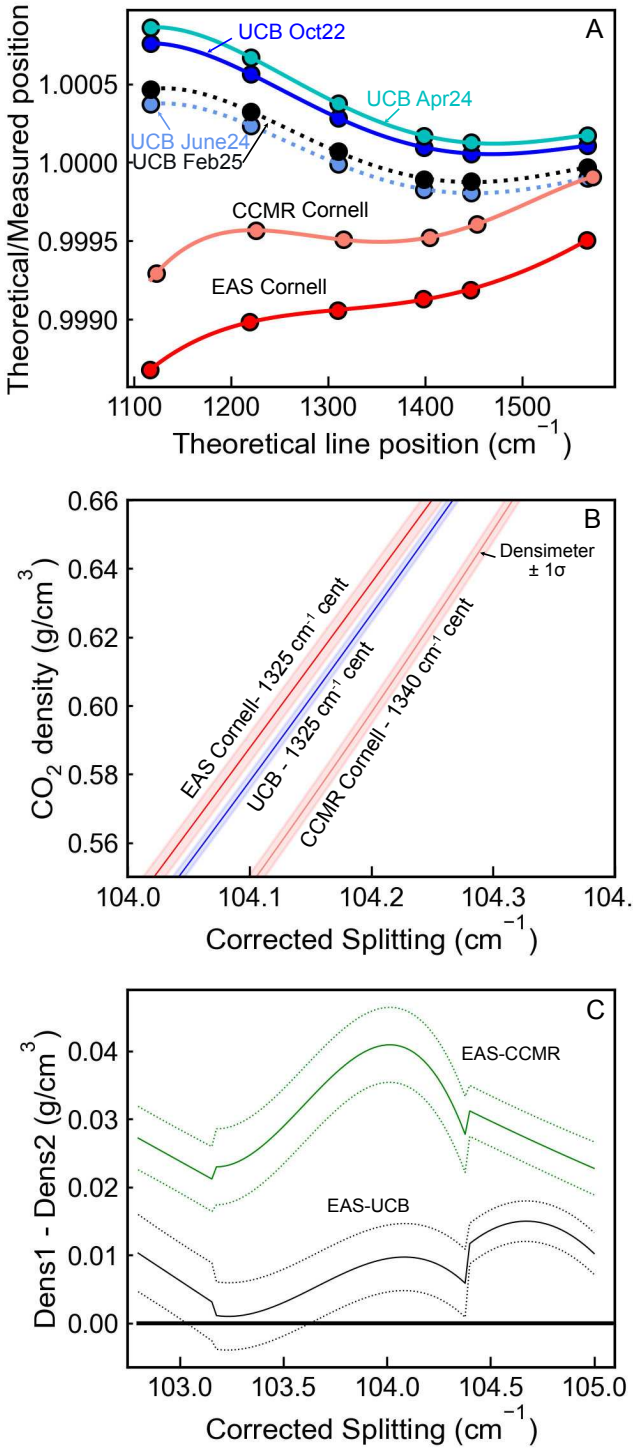


Figure 3: [A] Comparison of spectral non-linearity determined from Ne line acquisitions for three different WITec alpha300R instruments [this study and DeVitre et al. 2023a]. October 2022 and April 2024 were collected using one spectrometer calibration, and then the system was recalibrated prior to collecting the June 2024 and February 2025 lines. [B] Comparison of the densimeters across a smaller density range showing that UCB and EAS are more similar than CCMR and EAS, despite CCMR-EAS showing much more similar spectral non-linearities [A]. [C] Difference between densimeters as a function of splitting. The filled lines on b) show the 1σ errors on the densimeter fit, combined in quadrature and shown as dashed lines in [C].

that correction factors show substantially more variability in each run than using the 1117–1447 cm^{-1} pair (see error bar on Figure 1C). The 1117–1447 cm^{-1} pair used by our laboratory yields a reasonably similar correction factor to the bracketing technique (0.9982 vs. 0.9977). However, applying the line segment technique to the 1220–1567 cm^{-1} lines (red dashed segment) would result in a Ne correction factor that is far smaller than that determined using the bracketing technique (0.9991 vs. 0.9977), yielding a higher apparent diad splitting ($\Delta^{\text{Corr, brack}}_{CO_2}$ = 103.450 cm^{-1} vs. $\Delta^{\text{Corr, seg}}_{CO_2}$ = 103.648 cm^{-1}).

The differences in correction factor mean that the choice of Ne line pair does indeed affect the corrected splitting, and the relationship between CO_2 density and corrected splitting (i.e. the densimeter). To demonstrate this, we take the CO_2 and Ne spectra from the analytical sessions used to calibrate the UCB densimeter, which was originally calibrated from corrected splitting calculated using correction factors from the 1117–1447 cm^{-1} line pair [DeVitre et al. 2021; DeVitre and Wieser 2024, black dotted line, Figure 2A,B]. We reprocess the data using 4 new pairs of Ne lines: ~1220–1447 cm^{-1} (dashed red line, Figure 2A,C), ~1117–1399 cm^{-1} (dashed blue line), 1220–1398 cm^{-1} (solid cyan line), and 1220–1567 cm^{-1} (solid magenta line, Figure 2A,B). For each CO_2 spectrum collected on a fluid of known density in the calibration cell, this gives us 5 different corrected splitting values (one for each Ne line pair). Using three polynomial fits for low, medium and high densities [following DeVitre et al. 2021], we calibrate coefficients for 5 different densimeter equations using these 5 sets of corrected splitting values (Figure 2A,B). Despite being collected on a single instrument using the same calibration dataset, the densimeters differ by up to ~0.08 g cm^{-3} simply based on the Ne line pair used for data reduction. To put the magnitude of these offsets into context, we underlay the densimeters collected on a wide variety of Raman instruments [Kawakami et al. 2003; Yamamoto and Kagi 2006; Fall et al. 2011; Wang et al. 2011; Lamadrid et al. 2017], colored using the high (grey) and low (salmon) pixel resolution classifications from Remigi et al. [2021]. This demonstrates that the variability from a single instrument introduced through the choice of a Ne line pair is as large as the differences between Raman instruments, despite being collected at a single spectral and pixel resolution.

If a gas cell apparatus is available for calibration of a specific instrument, the offsets between densimeters determined using different Ne line corrections are only problematic if they yield different results when analysing unknowns. We test whether this is the case using data collected on FI from Fogo volcano collected by DeVitre and Wieser [2024], where there are independent constraints on CO_2 density from microthermometry. We process the FI data using each of the 5 different Ne line pairs (Figure 2B), generating 5 values of corrected splitting for each spectrum, and input these values into the den-

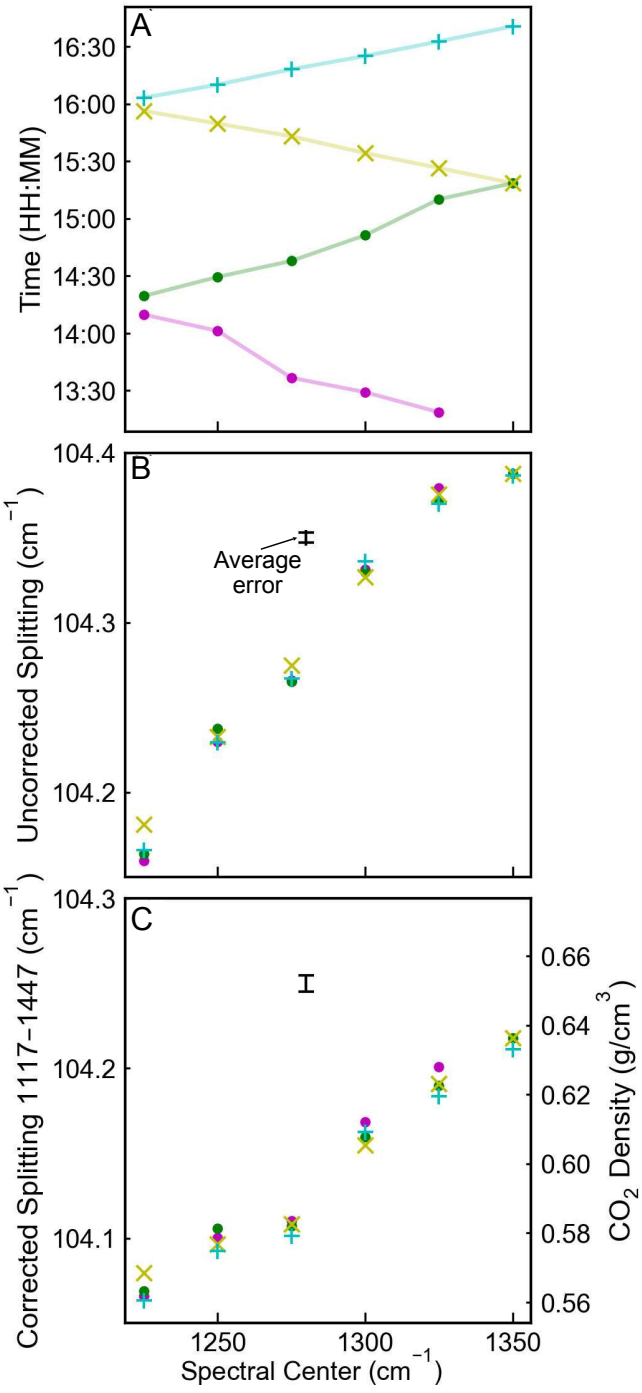


Figure 4: Change in splitting and calculated density for a single FI as the spectral center is shifted. [A] The four symbol colors show 'up' and 'down' cycles changing the spectral center as a function of time through the day. The same coloring is used for [B] and [C]. [B] Uncorrected splitting vs. spectral center and [C] corrected splitting, using Ne lines collected at the same spectral center as the diad. The black error bar shows a representative error for the splitting [B] and corrected splitting [C] respectively (offset from trend for clarity). The density-axis in [C] is calculated using the densimeter developed for a spectral center of 1325 cm^{-1}.

simeter corresponding to each Ne line pair (e.g. data reduced using correction factors for the 1220–1399 pair was fed into the 1220–1399 densimeter). To visualize the differences between these data reduction strategies, we subtract the density calculated using each line pair from that calculated using the 1117–1447 pair, which was shown to correlate very well with microthermometry [DeVitre and Wieser 2024]. The error bar on each datapoint shows the combination of the peak fitting error, drift correction error, and uncertainty on the densimeter. Densities for the 4 other line combinations are within error of the data reduced using the 1117–1447 cm^{-1} line pair. We also overlay lines showing the 1σ uncertainty associated with microthermometry assuming a 0.1 °C and 0.2 °C error in sample temperature [Bakker 2021]. This analysis demonstrates that the uncertainties in the choice of data reduction strategies are small, and comparable to the errors associated with the stage temperature when performing microthermometry.

Repeated acquisitions of Ne lines at UCB in April 2024 and October 2022 shows that the non-linearity response curve shifts up and down with time, but the shape is maintained (Figure 3A, solid lines). We also tested recalibrating the UCB spectrometer in April 2024 using a Hg-Ar lamp (WITec's default calibration routine), and the same pattern is observed. Thus, while spectrometer non-linearity is clearly significant, and likely introduces diversity in published densimeters, as long as calibration and natural data is processed in the same way, the non-linearity introduces a consistent offset through time, and choice of Ne line will not cause systematic offsets. This stability also means that the bracketing technique is not necessary, and given the additional computation steps, likely results in a larger uncertainty due to the large fitting errors associated with the weaker Ne lines (e.g. 1310 cm^{-1}). However, if a laboratory wishes to change drift correction strategy, they would need to reprocess their calibration data to use the same strategy.

This analysis demonstrates that not only does a densimeter need to be generated for each instrument, but that each densimeter is only valid for a specific detector calibration and drift reduction protocol. Supplementary Material 1 Table S1 shows that a wide variety of calibration and drift correction routines are being used by different laboratories publishing densimeters. This can be expected to generate a wide variety of non-linearity responses, and thus we speculate that a lot of the diversity in measured densimeters in the literature is a result of spectral non-linearity, both in its original effect on the diad splitting, and the compounding effect of the use of a line-segment Ne line correction and other drift correction routines. It is possible that performing a polynomial fit like that suggested by Bakker [2021] would resolve offsets between densimeters calibrated in different laboratories; extensive testing using different Raman instruments for standards with known densities would be a promising avenue for future research.

3.1.1 *The effect of changing the spectral window*

The significant non-linearity in the spectrometer response shown in Figure 1B has other implications for the densimeter equation determined for a given instrument. When setting up a Raman spectrometer, the user must define the wave-

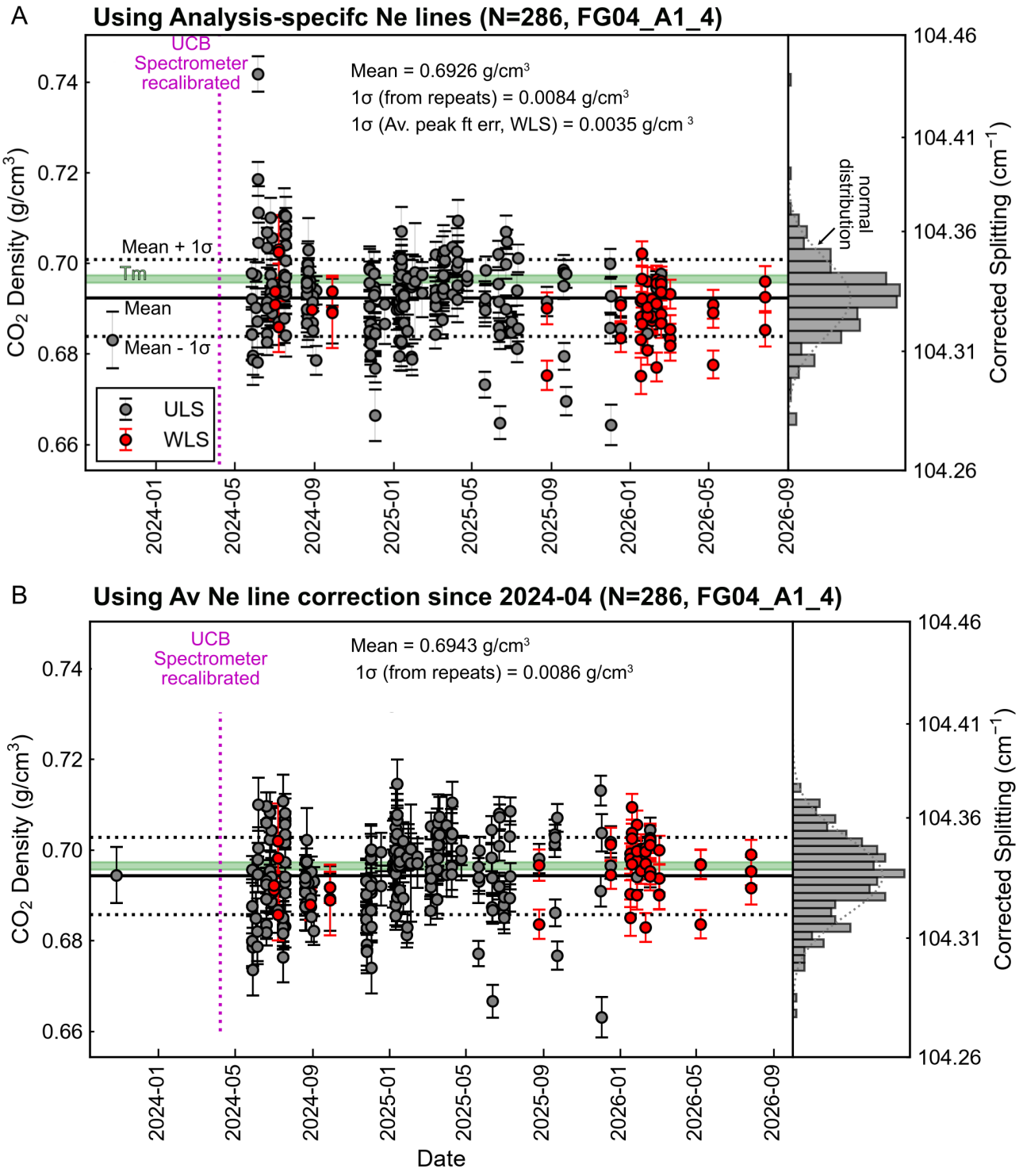


Figure 5: Long-term variability in repeat measurements of FI FG04_A1_4 at UCB. [A] Data reduced using a Ne drift correction model for each analytical session. [B] Data reduced using the average Ne line correction factor for all standard acquisitions after the spectrometer recalibration (marked with a magenta line). The green bar shows the density determined by microthermometry, accounting for the uncertainty in the temperature reading of the stage. The black solid line shows the mean Raman value for the standard, and the dotted black lines ±1σ. The distribution of densities is shown as a histogram, compared to a normal distribution based on the mean and standard deviation. Supplementary Material 1 Figure S10 shows the same plot for a second standard run less frequently at UCB. Grey symbols use unweighted least squares fits (ULS), red symbols used weighted least squares routines (WLS).

length range that is recorded on the charge-coupled device (CCD). On the three WITec instruments we use data from in this paper, this is specified by stating the spectral center—the wavenumber at the center of the acquisition window. Along with the grating, this center position determines the range of wavenumbers collected in a single spectral acquisition. Shifting the center means that the photons at a specific wavenumber are collected on a different pixel on the CCD detector within the spectrometer. On other instruments, the window edges rather than center position may be the tuneable parameter.

Given the strong spectral non-linearity shown in Figure 1B, it is worth investigating whether shifting the spectral center affects the measured splitting, because each pixel of the diad would shift to a different part of the polynomial in Figure 1B [see also Bakker 2021]. We perform repeated acquisitions on an olivine-hosted CO_2-rich FI from Fogo volcano with a bulk density of ~0.62 g cm^{-3} using different spectral centers. We started at the spectral center used for calibration of the Berkeley densimeter (1325 cm^{-1}), then decreased the spectral center (1325→1320→1275→1250→1225 cm^{-1}, magenta colors, Figure 4A). To help resolve changes from the spectral center vs. instrument drift, we increased (green colors), decreased (yellow colors) and then increased (cyan) the spectral center throughout the day (Figure 4A). We acquired a Ne line at the same spectral center as each FI measurement to correct each spectra (using the 1117–1447 cm^{-1} line pair). For each individual acquisition, we calculated the corrected splitting using the Ne line acquired at the same spectral center. Lastly, we calculated density using the UCB densimeter developed using a spectral center of 1325 cm^{-1}.

The uncorrected splitting of the diad increases by ~0.22 cm^{-1} as the spectral center is increased from 1225 to 1350 cm^{-1} (~0.0017 cm^{-1} per cm^{-1} of center shift, Figure 4B). The change in splitting for spectral center shifts of only ~25 cm^{-1} apart is far larger than the variation at a given spectral center throughout the day (e.g. instrument drift). After correcting each acquisition using the Ne line collected at the same spectral center, the change in splitting is smaller, but still notable (~0.15 cm^{-1} across the window, 0.0012 cm^{-1} per cm^{-1} of center shift, Figure 4C). Calculated CO_2 densities after a line segment Ne line correction change by 0.08 g cm^{-3} across the window (Figure 4C), which is far outside the ~0.01–0.02 g cm^{-3} uncertainty of the Raman method at the densities of interest [Dayton et al. 2023; DeVitre and Wieser 2024]. Overall, this test indicates that even just shifting the spectral window on a given instrument can generate a new densimeter, and that unknowns must be analysed at the same spectral center/spectral window at which the calibration was performed.

To further illustrate the influence of spectral center, we compare the densimeter developed on the UCB WITec alpha300R with densimeters developed in two laboratories at Cornell University: Cornell Center for Materials Research (CCMR) and Earth and Atmospheric Sciences at Cornell [EAS, DeVitre et al. 2021]. These two laboratories were calibrated with a gas cell that is very similar to that at UCB [DeVitre et al. 2021], and the Raman hardware for all three instruments is very similar: a ~532 nm laser wavelength, 1800 grating, and Ultra-High-Throughput-Spectrometer UHTS600. All data was reduced with the 565.66588–576.44188 nm Ne line pair (corresponding to 1117–1447 for the UCB 532.046 nm laser wavelength). Comparing the fitted Ne line positions to the theoretical positions (Figure 3A) shows that each spectrometer has its own unique signature of non-linearity.

Despite EAS and CCMR having more similar non-linearity responses, the CCMR densimeter shows the biggest offset in densimeter equation relative to UCB and EAS. We suggest that this is because the UCB densimeter and the EAS densimeter were both calibrated using a spectral center of 1325 cm^{-1}, while the CCMR densimeter was calibrated with a spectral center of 1340 cm^{-1}. The absolute difference between these densimeters reaches ~0.04 g cm^{-3} (Figure 3C). Again, these densimeter differences cannot be explained by differences in resolution as advocated by Remigi et al. [2021], because the pixel and spectral resolution of all three of these WITec instruments is identical.

3.2 Densimeter stability as a function of time

The above tests make it apparent that small changes in instrument hardware can shift the densimeter. It is interesting to assess whether temporal shifts occur when using identical acquisition conditions and spectrometer settings. We evaluate repeated analyses of internal fluid inclusion standards from Fogo volcano (FG04_A1_4, N = 286 and FI FG04_101_A, N = 99) collected at the start and end of analytical sessions over ~2.5 years. The standard deviation calculated from the 280 repeats on A1_4 is 0.0086 g cm^{-3} (Figure 5A) and for the 99 repeats of 101A is 0.0074 g cm^{-3}. For both inclusions, microthermometry measurements are within 0.01 g cm^{-3} of these long-term averages and lie within or extremely close to the 1σ envelope around the mean Raman measurement. Overall, this small deviation and lack of coherent trends with time indicates that at least on the WITec, changes in spectrometer non-linearity does not cause any notable drift in the densimeter, and densimeters are remarkably stable once calibrated. Repeated acquisitions of a natural fluid inclusion in a similar manner to that performed here can be used to assess the point at which a densimeter recalibration is needed.

While the Ne line method is clearly robust, even with variations in non-linearity, it is worthwhile to determine whether it actually helps correct for drift between and within runs. Yamamoto and Hagiwara [2024] show that over short timescales (minutes) where drift is not controlled by temperature, periodicity is seen in CO_2 peak positions that is not seen in Ne line acquisitions. We reprocess the repeated FI acquisitions using the median Ne correction factor for all the standard analyses, rather than the correction factor calculated by regressing Ne correction factors as a function of time in each analytical run. The standard deviation of both FIs is extremely similar with and without an analysis-specific Ne correction. Interestingly, for FG04 one outlier is generated using a Ne line correction specific to that analysis that goes away using the median correction factor (Figure 5A), while the opposite is true for standard FG19_101 (one outlier is generated using the median Ne correction, Supplementary Material 1 Figure S10). However, it is noteworthy that after the instrument recalibration in

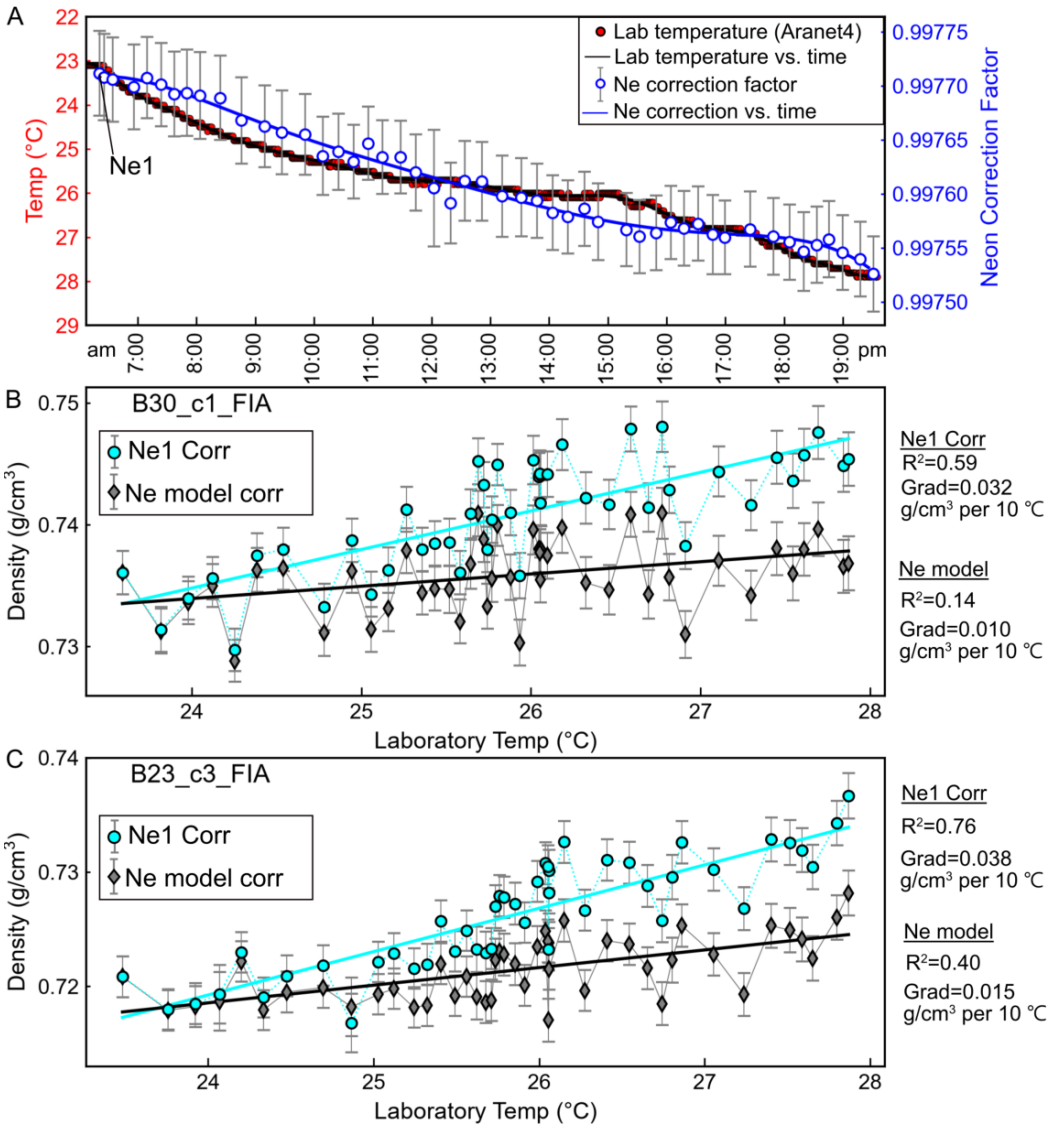


Figure 6: Testing the sensitivity of Raman acquisitions to large fluctuations in room temperature. [A] Temperature log from an Aranet4 logger (red dots) against time. The white circles with blue outlines show the Ne correction factor calculated for each Ne line acquisition. These series were fitted with a regression (black and blue lines) to calculate temperature and Ne correction factor for each FI acquisition. [B] Change in density of FI B30_c1_FIA as a function of laboratory temperature. The grey diamonds are corrected based on the Ne line correction model (blue line in [A]), while the cyan circles use the correction factor for the first Ne acquisition (Ne1, to replicate not applying a drift correction). The straight lines represent linear regressions, with the statistics shown on the right. [C] Same for B23_c3_FIA. Error bars from weighted least squares fits.

April 24, the Raman was moved to a newly renovated laboratory where the HVAC system yields temperature variations of <±1 °C per day, which means temperature variations are minimal, which could explain the small differences between the correction methods. We investigate the ability of line segment Ne line corrections to help minimise drift during larger temperature variations in Section 3.3 below.

3.3 Densimeter stability within a single analytical session

It is well know that changes in temperature can cause considerable drift in peak positions, intensities and peak widths during Raman analyses [e.g. Mestari et al. 1997; Lamadrid et al. 2017; Hagiwara et al. 2018; Jakubek and Fries 2022; Inoue et al. 2023]. For example, Fukura et al. [2006] tracked peak positions in ruby and napthalene, showing periodic fluctuations in peak positions of ~0.1 cm^{-1} with temperature fluctuations of 0.8 °C. However, they show that the multiple peaks of napthalene follow a similar trend, and the difference between two peak positions, which is of most relevance for measurements of CO_2 fluid density, is ~8X smaller than the absolute peak shift. In contrast, Mestari et al. [1997] perturb room temperature by 1-2 °C and track the relative positions of different Ne

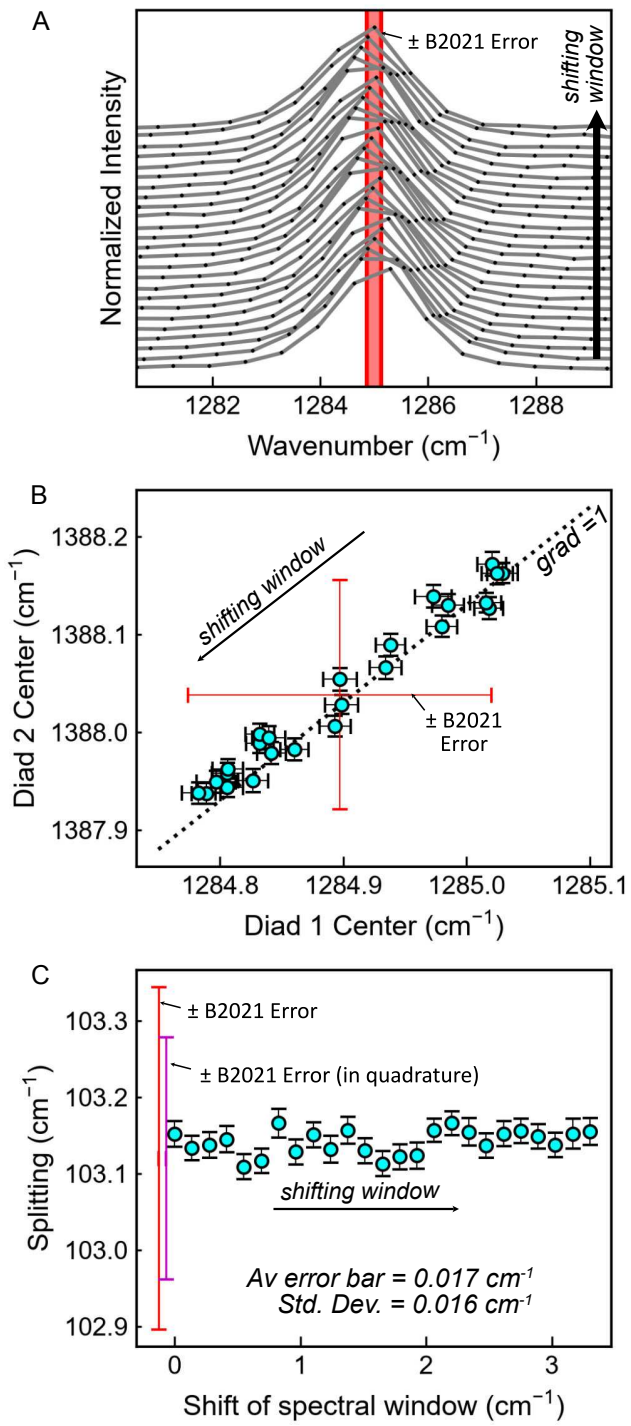


Figure 7: [A] Stacked spectra collected as the spectral window was progressively shifted by Bakker [2021]. The red band shows their stated error on the peak position, based on the wavenumber shift in the window. [B] Plot showing the changing position of diad1 and diad2 for each acquisition shown in [A]. The black dashed line has a gradient of 1 to emphasize that diad1 and diad2 shift by equal amounts. The error bar shows the 1σ error estimated by `lmfit` from the peak fit. [C] Change in splitting as a function of the shift of the window. Errors estimated by Bakker [2021] are shown in red, and purple shows their stated error on each peak combined in quadrature. The black error bar shows the uncertainty based on the `lmfit` peak positions propagated in quadrature using a weighted least squares routine.

lines and show the wavenumber shifts are not identical for different lines and diverge further as the experiment proceeds. This study also shows that amount of relative wavenumber shift between two lines differs drastically as a function of the acquisition conditions.

Hagiwara et al. [2018] investigate the effect of both room and sample temperature on CO_2 density measurements to isolate the effect of each. They show no notable change in splitting as sample temperature is perturbed by 21-26 °C, but an abrupt shift in splitting of ~0.05 cm^{-1} as the room temperature is raised by 7 °C. The splitting then rapidly shifts back to its original value again. As the authors do not plot splitting against room temperature, it is difficult to determine exactly when this perturbation happens relative to the onset of room heating. They suggest the splitting change reflects thermal expansion induced misalignment of the spectrometer, as thermal expansion of the host mineral will be negligible. The Ne line correction routines of Lamadrid et al. [2017] and Lin et al. [2007] do not mitigate the change in splitting during this perturbation. Thus, Hagiwara et al. [2018] suggest that repeated measurements of a CO_2 standard may provide a better drift correction strategy than Ne lines.

Given the differential response to temperature Mestari et al. [1997] report using different instrument settings on a single instrument, and conflicting reports of whether different peaks move together [Mestari et al. 1997 vs. Fukura et al. 2006], we investigate the effect of room temperature variations and whether the Ne line segment technique can minimise the effect on measured splitting on the UCB WITec alpha300R. We turned off the room heating and cooling system the night before a day that was forecast to be particularly warm for the San Francisco Bay area (~28-30°C). We started the test at ~6:30 am when the laboratory temperature was 23.5 °C, 30 mins after the laser and Ne lamp were turned on and the spectrometer set to cool, Two Icelandic fluid inclusions [Kahl et al. 2025] were double polished and placed in the Linkam THM600 stage with a silver lid on top of the sample to keep the sample temperature fixed at 37 ± 0.1 °C throughout the day. We do not use the top lid of the stage, as it cuts out a large proportion of the Raman signal (see Supplementary Material 1 Figure S4-5). As the room warmed up, the two FI were analysed repeatedly, with a Ne line acquisition collected after each measurement of the two FI. Room temperature was measured using an Aranet4 sensor located on the table with the Raman and Ne lamp (Figure 6A). By 3pm, the room temperature had stabilized at ~26 °C as the outside temperature peaked. To get a larger temperature rise, a space heater was switched on. At 5 pm, the temperature began to level out again, so a second heater was turned on, causing the laboratory temperature to rise to 28 °C at ~7:30 pm. Two air filters with a combined clean-air delivery rate of 500 cubic feet per minute were used to circulate air within the room to help homogenize laboratory temperature. The laser power recorded

by the WITec TruePower stayed between 5.95 and 6.2 mW with no obvious correlation with room temperature, showing similar variability to analytical days where laboratory temperature was stabilized (Supplementary Material 1 Figure S11).

The peak positions of the 1117, 1220, 1400, 1447, and 1567 lines increase by ~0.2-0.3 cm^{-1} over 5 °C, which is very similar to the 0.3 cm^{-1} shift of the peak position of GaN observed over 5 °C of room warming by Guo et al. [2020]. Interestingly, the proportional change in peak position is extremely similar for all the Ne lines (position at t=i/position at t=0), so the absolute change is largest for the 1567 line and smallest for the 1117 line (Supplementary Material 1 Figure S12). While the Ne correction factors differ for each Ne line pair, the proportional change in correction factor is within uncertainty for all Ne line pairs (Supplementary Material 1 Figure S12). This indicates that the magnitude of the drift correction using Ne lines is independent of the choice of line pair.

Given the similar response of all Ne lines and their correction factors, for simplicity, we only discuss the 1117–1447 line pair. The 1117–1447 correction factor shows a strong anticorrelation with temperature (Figure 6A, note the flipped y-axis for temperature). By parametrizing temperature vs. time and Ne correction factor vs. time using high order polynomials (Figure 6A), a laboratory temperature and correction factor was calculated for each FI measurement. We calculate the corrected splitting using the Ne correction model (blue line, Figure 6A) and compare this to corrected splitting calculated using the first Ne acquisition at the start of the day (Ne1, Figure 6A). Fitting a linear regression for calculated density vs. laboratory temperature yields a far stronger correlation when using a constant Ne correction factor from the first Ne acquisition of the day vs. the polynomial Ne correction model (R^2 = 0.57 vs. 0.14 for B30, and 0.74 vs. 0.39 for B23). This indicates that during large temperature swings, the line segment Ne line correction does help correct for instrument drift, keeping the rate of change within ~0.01 g cm^{-3} for 10 °C of laboratory heating, while not using a Ne lamp would generate ~0.03–0.04 g cm^{-3} of drift.

In our experience with our hardware and specific analytical protocol, on a day without such extreme temperature perturbations, measuring a Ne line every 40–60 mins is generally sufficient. However, we have found that our Ne line acquisitions have helped us identify days where the HVAC system was malfunctioning and frequent Ne line acquisitions have allowed us to use the data collected on those days, whereas with less frequent acquisitions we would have had to recollect. Given the differing results reported for Raman temperature sensitivity in the literature, we suggest that analysts inspect room temperature logs to determine typical daily temperature changes in any given laboratory to decide an appropriate interval at which to measure Ne lines. In laboratories with large temperature swings, performing room temperature tests similar to that described here on their instrument can help identify the best drift correction strategy. If other Raman systems behave like ours during temperature perturbations, where very large temperature swings can be corrected using Ne lines so the analytical error only doubles (from an already small value, ~0.01 g cm^{-3}), this is excellent news for universities who want to conduct Raman measurements but do not have access to temperature-controlled spaces. However, given our instrument uses fibre optic cables to transmit photons, it is possible that instruments relying on mirrors show larger shifts, given the increased number of components susceptible to thermal expansion and contraction.

3.4 Peak fitting errors

The above analysis shows that developing a densimeter for a specific analytical conditions and drift correction methods is key to remove the large systematic offsets between different densimeters and generate accurate measurements of CO_2 density. Now, we focus our discussion to optimizing the precision of Raman measurements.

Bakker [2021] raised concerns regarding precision, noting the discrete nature of Raman spectra, where x-coordinates (wavenumber) are collected on a CCD with a spacing defined by the grating and other hardware. At each x-coordinate, an intensity value is recorded, the magnitude of which is dependent on the efficiency of each instrument. Bakker [2021] show that shifting the spectral window one pixel spacing at a time along the CCD causes a lateral shift in the maximum peak intensity (Figure 7). They suggest that the error on the peak center is equal to the magnitude of the spectral shift. For two different Raman set ups, they estimate a splitting error of ~ ±0.27 cm^{-1}. A splitting error of ±0.27 cm^{-1} yields an error of ±0.1 g cm^{-3} using the UCB densimeter. For a typical melt inclusion vapour bubble with 3 % bubble volume, this corresponds to an uncertainty of the amount of CO_2 held in the bubble of ±1111 ppm. When propagated through to calculate saturation pressure, this would introduce uncertainty of several kbar. Similarly, a FI with a density of 0.3 ± 0.1 g cm^{-3} corresponds to entrapment pressures spanning 0.5 to 1.25 kbar [Span and Wagner 1996]. Peak fitting errors of the magnitude suggested by Bakker [2021] would render Raman spectroscopy entirely unsuitable for melt and FI barometry, so it is essential to assess these statements.

We obtain the spectra shown in Figure 18 of Bakker [2021] and fit them with pseudovoigt peaks for the diads and hotbands. These spectra were collected as the spectral window was shifted ~3 cm^{-1} in steps of 0.137 cm^{-1} on a LabRamHR (600 grating, pixel resolution of 0.62 cm^{-1}/pixel). To help visualize the shift in the maximum y-pixel intensity, we vertically stack the acquisitions collected using different spectral centers after normalizing the intensity for ease of visualization (Figure 7A). The positions of diad1 and diad2 do indeed move by significant amounts as the window is shifted (Figure 7B). Bakker [2021] suggest that these shifts reflect the error on the band positions. For this dataset, both diad1 and diad2 shift by 0.226 cm^{-1} as the window is moved (red error bar, Figure 7C). However, it is crucial to recognise that diad1 and diad2 move in the same direction (Figure 7B), such that the calculated splitting varies very little and shows no systematic change as the spectral window is shifted (cyan diamonds, Figure 7C). The overall variation in calculated splitting (σ = 0.017 cm^{-1}) is very similar to the estimated splitting error combining the peak fitting errors on diad1 and diad2 from `lmfit` in quadrature within `DiadFit` [0.017 cm^{-1}; Newville et al. 2016]. Thus, the

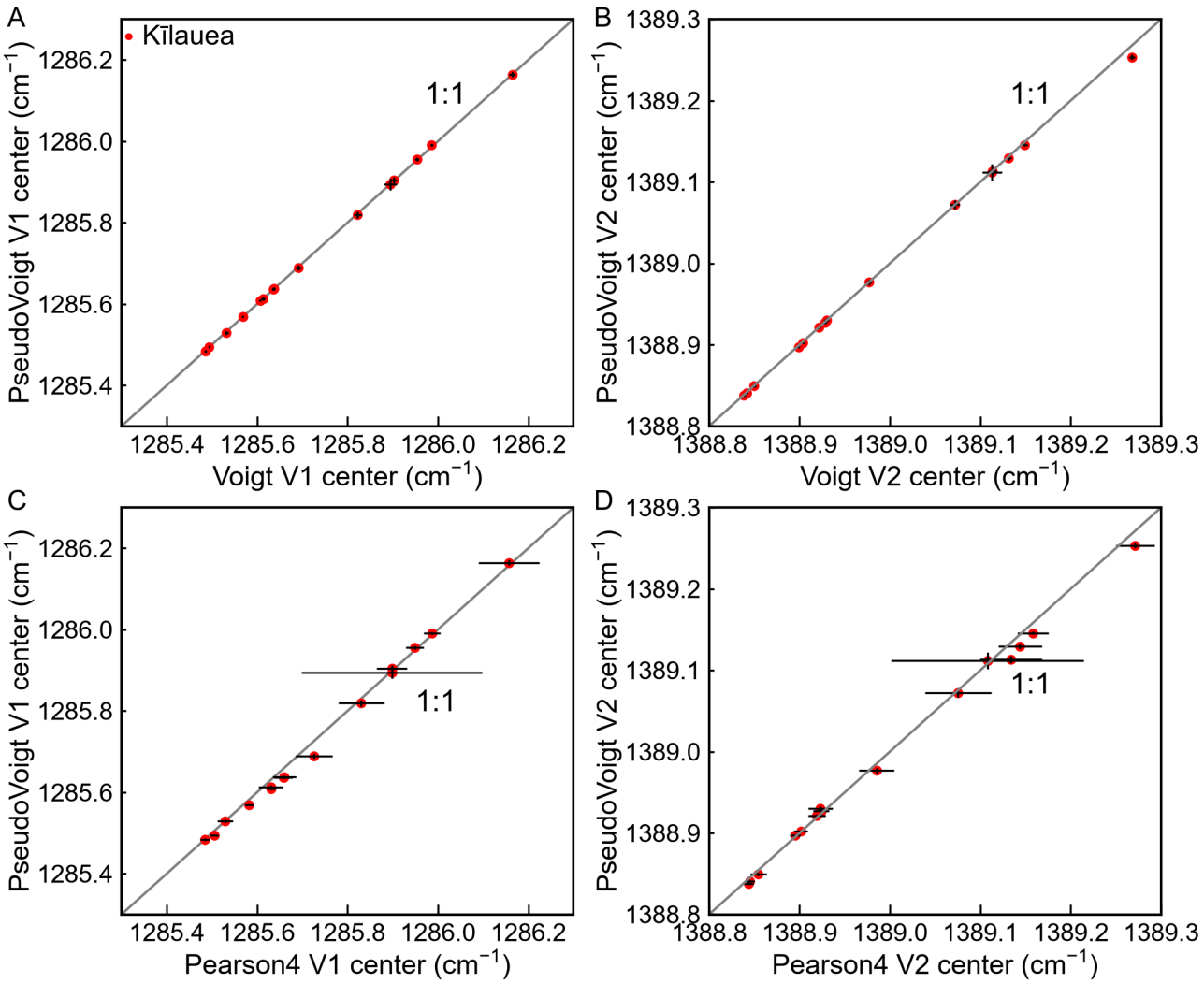


Figure 8: Comparison of peak centers obtained with Voigt (V), Pseudovoigt (PV) and Pearson4 (P4) peak fits for diad1 (V1) and diad2 (V2) for natural FIs from Kīlauea volcano. [A-B] Comparison of diad1 and diad2 position, with error bars showing the 1σ error on each peak fit from `lmfit`. Comparison of PDF fits for Fogo inclusions are shown in Supplementary Material 1 Figure S13.

splitting errors suggested by Bakker [2021] are overestimated, because shifting the window causes the position of diad1 and diad2 to move in the same direction, with no notable change in the splitting.

Bakker [2021] also suggest that the choice of different probability density functions (PDFs, e.g. Voigt, PseudoVoigt, Gaussian) return different center positions, and that the errors on these parameters are hard to estimate. Fortunately, modern peak fitting packages such as `Fityk` [Wojdyr 2010], and peak fit packages in `Python` such as `sklearn` or `lmfit` [used in `DiadFit`, Newville et al. 2016] provide an estimate of the uncertainty on the peak center using the covariance matrix obtained during the peak fitting process. We assess the difference in peak position using different fits for Raman data collected on natural FIs from Fogo volcano, Cabo Verde and Kīlauea, Hawaii [DeVitre and Wieser 2024]. We consider Voigt, PseudoVoigt, and Pearson4 functions, as these were identified as the most physically-plausible peak shapes for CO_2 fluids by Yuan and Mayanovic [2017]. Peak positions obtained from different fits overlap within the calculated uncertainty on the peak center (Figure 8, Supplementary Material 1 Figure S13). This analysis demonstrates that for a wide range of CO_2 densities and signal intensity, differences between physically plausible peak fit shapes are within the stated `lmfit` error on each model.

It is notable that the errors on the Pseudovoigt peak position are twice as large for diad1 as diad2, consistent with its lower intensity and slightly higher full width half maximum (Figure 9A). Errors are also larger at lower CO_2 densities. However, for all inclusions, peak fitting errors are still significantly smaller than the 0.1-0.3 cm^{-1} estimated by Bakker [2021]. Comparing Raman measurements with microthermometry measurements provides another opportunity to assess precision (and accuracy, Figure 9B). We use data from N = 58 FI from Fogo volcano from DeVitre and Wieser [2024] supplemented with N = 39 new microthermometry and Raman measurements from the Galápagos using the same methods. The Raman error was estimated from `DiadFit`, by com-

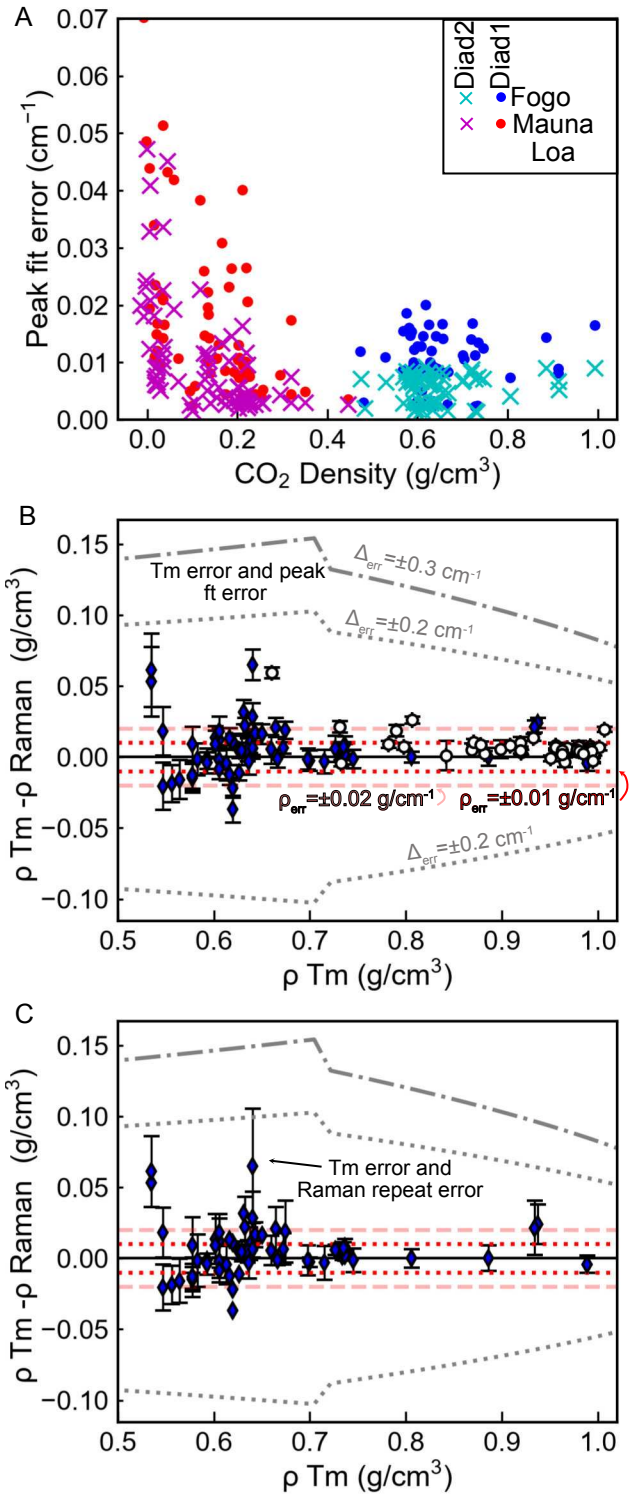


Figure 9: Peak fit errors for three different sample suites. [A] Errors on each peak position obtained from `lmfit` for low and high density FIs from DeVitre and Wieser [2024] from Fogo and Wieser et al. [2025] from Mauna Loa. [B] Comparison of densities obtained from Raman spectroscopy and microthermometry [this study - white dots; DeVitre and Wieser 2024, - blue diamonds]. The error bars show 1σ errors estimated from peak fitting, analytical drift and densimeter uncertainty for Raman spectroscopy and microthermometry combined in quadrature. [C] Same comparison but using standard errors calculated from repeated Raman measurements (only FI with repeated analyses are shown). On [B]–[C], we overlay the ±1σ error bounds that would result from the splitting errors suggested by Bakker [2021]. We also overlay red lines showing the ±0.01–0.02 g cm^{-3} errors that numerous studies comparing the two methods have stated.

bining the peak fitting errors with the uncertainty from the Ne line correction model and the densimeter. This error was combined in quadrature with the uncertainty in microthermometry homogenization temperature (±0.2 °C) to create the y-error bar shown in Figure 9B. On Figure 9C, we perform the same analysis but now show error bars estimated from multiple repeats of FI analyses, rather than the `DiadFit` peak fitting error. On both plots, the two FI with the largest offset between Raman and microthermometry have CO_2 densities very close to critical, where microthermometry errors are hard to quantify. The majority of inclusions lie within the ~0.01–0.02 g cm^{-3} error commonly quoted for the Raman method [red bars, Kobayashi et al. 2012; Remigi et al. 2021; Dayton et al. 2023; DeVitre and Wieser 2024]. In contrast, it is clear that when propagated to density, the proposed splitting errors of ~0.2-0.3 cm^{-1} (grey dotted and dashed lines) from Bakker [2021] are not consistent with this dataset.

Although the errors estimated from comparing microthermometry and Raman (Figure 9), and repeated measurements of standards (Figure 5) are far smaller than that estimated by Bakker [2021], it is noteworthy that the standard deviation in density obtained from averaging multiple repeats is about twice as large as the error estimated from `DiadFit` (using weighted least squares or unweighted least squares, Figure 9). Similarly, the average individual fitting error of the FI standards (Figure 5) are about half the total standard deviation in repeated measurements. This indicates that the there are additional sources of uncertainty in addition to the uncertainty in peak position, Ne correction models, and the densimeter itself considered by `DiadFit`. In Section 3.6, we suggest that a reasonable fraction of the unaccounted-for uncertainty comes from the temperature of the sample, and other sources of instrumental variability.

3.5 Optimizing signals

While the above discussion indicates that peak fitting errors are far smaller than those indicated by Bakker [2021], it is still worthwhile investigating methods to reduce uncertainties as much as possible. This is particularly important for low density CO_2 fluids (i.e. melt inclusion vapour bubbles), where peak fitting errors can approach 50–100 % near the instrument detection limit [~0.01–0.02 g cm^{-3}, Bearden et al. 2026]. Peak fitting errors are highly dependent on both the instrument hardware (e.g. spectrometer efficiency, spectral resolution, read and dark noise from the CCD), analytical conditions (e.g. laser power, acquisition time, number of accumulations averaged), and sample characteristics (e.g. density of CO_2, size of inclusion).

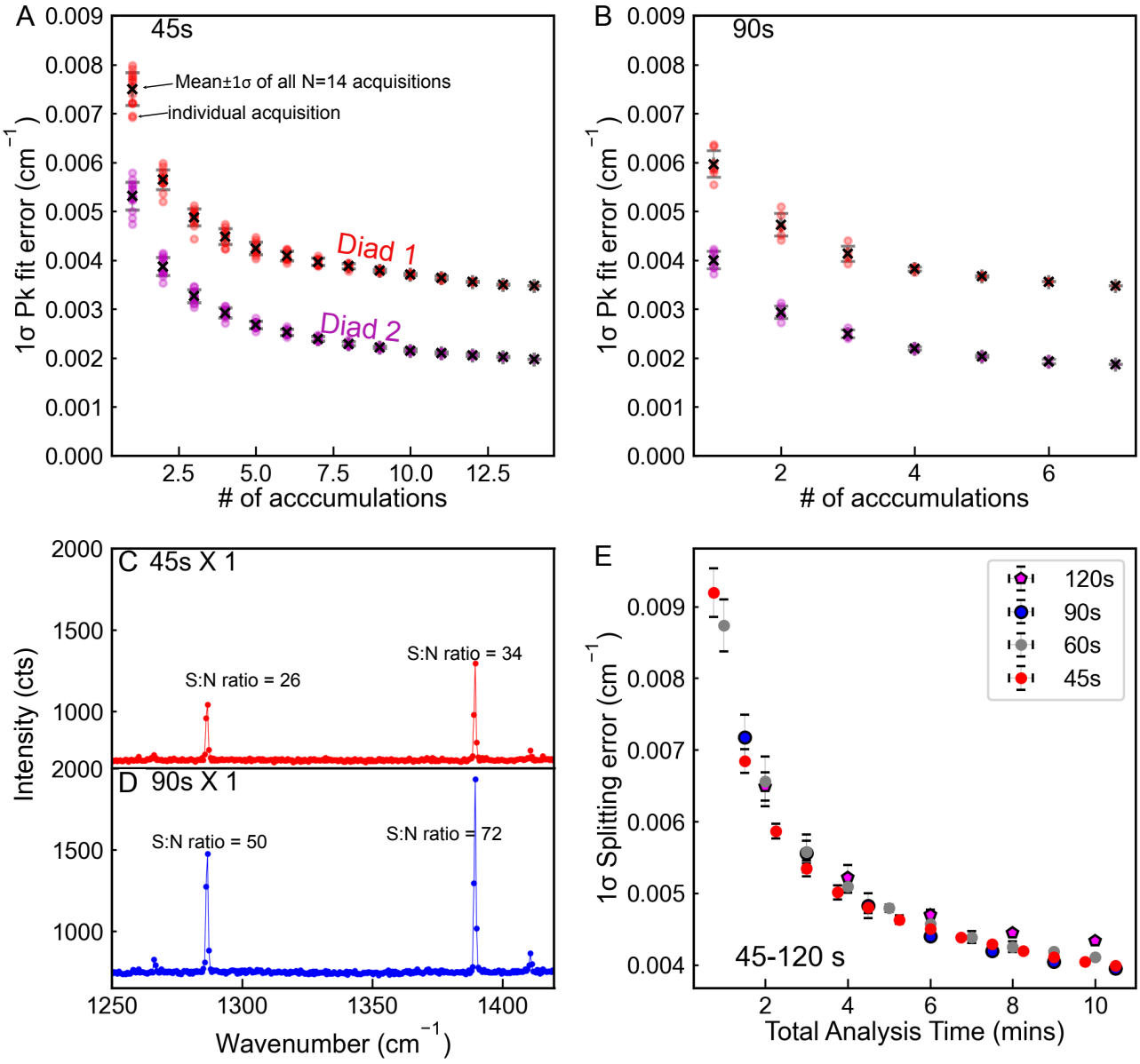


Figure 10: Investigating peak fitting error as a function of integration time and accumulations using a low density capillary tube [~0.01 g cm^{-3}, Le et al. 2021]. [A-B] Error on the peak position of diad1 and diad2 for 45 and 90 s acquisitions. The x-axis shows the number of spectra that were averaged prior to fitting the peaks (analogous to accumulation). The dots show an individual fitted spectra (e.g. spectra 1 to 14 for accumulations = 1, average of spectra 1–2, 2–3, 3–4 for n = 2 accumulations etc.), and the cross and error bar show the mean and standard deviation of averaged spectra for each number of accumulations (e.g. 14 individually fitted spectra for n = 1 accumulations). [C-D] Example spectra for 45 s × 1 and 90 s × 1. The signal to noise (S:N) ratio is calculated as the height of the highest pixel divided by the standard deviation of the region between the two diads. [E] Error on peak fitting as a function of total acquisition time. For example, five 120 s acquisitions were averaged for the 10 minute total acquisition time. The same figure for a low density melt inclusion vapour bubble is shown in Supplementary Material 1 Figure S16-17.

On most instruments, users can adjust the acquisition time (time spent collecting a single spectrum) and the number of accumulations (number of individual spectra that are collected and averaged). We investigate the interplay between these two parameters using a variety of samples with low CO_2 densities (which have the largest peak fitting errors). We collect N consecutive individual acquisitions of variable duration using the 'slow series' option in the WITec acquisition software. This mode keeps the laser on continuously, and collects and outputs spectra every Xs. This is analogous to a single accumulation when run in the regular mode. To simulate the software collecting two accumulations of time Xs, we average the 1st and 2nd collected spectra to get one averaged spectra and then fit that spectra. We then average the 2–3rd acquisi-

tion, the 3–4th and so forth. For N = 14 acquisitions collected, this yields 13 averaged spectra that we fit (shown as individual dots on Figure 10A). We also calculate the average and standard deviation of each peak parameter for the 13 averaged spectra. To simulate 3 accumulations, we average the 1–2–3rd spectra, 2–3–4th, and for 4 accumulations, the 1–2–3–4th spectra, 2–3–4–5th spectra etc. Tests were performed for different integration times (30 s, 45 s, 60 s, 90 s, 120 s and 200 s). This allows us to investigate how errors decline as count time increase through a combination of changing accumulations and acquisition times.

Figure 10 shows a CO_2 capillary tube with a pressure of 0.6 MPa, corresponding to $\rho = 0.011 \pm 0.011$ g cm^{-3} [Le et al. 2021]. Supplementary Material 1 Figure S16-17 shows the same plot for a low density melt inclusion vapour bubble from Mauna Ulu, Kīlauea [Wieser et al. 2019]. Each 'averaged spectrum' is shown as a semi-transparent dot onFigure 10A,B, and the overall average for combinations yielding N accumulations is shown as a cross with an error bar. For a single spectra, the errors are smaller for 90 s compared with 45 s, as expected given the more intense spectra, and higher signal to noise ratio (Figure 10C vs. Figure 11D). The peak fit error for diad1 is about twice as large as for diad2. The error drops quickly as the number of accumulations is increased, and then the trend flattens out (Figure 10A,B). In terms of assessing ways to optimize instrument time, it is useful to compare the errors in the context of the total acquisition time (Figure 10E). There are no clear differences in error outside of uncertainty the same total time for acquisitions 45 s or greater. For the two melt inclusion vapour bubbles, 30 s acquisitions return errors that are slightly higher than all other times.

These observations can be understood when considering the factors that influence the uncertainty in the intensity of each pixel collected on the CCD Hagiwara et al. [2023]. The error for each pixel is a function of the shot noise (i.e. the Raman photons being delivered to the detector), dark noise (arising from thermally produced electrons in the detector) and read noise (introduced during conversion of charge to voltage during readout). The variance in detected photon counts due to shot noise (σ_{shot}^2) is proportional to counts (number of detected photons), the variance in counts due to dark noise (σ_{dark}^2) is proportional to exposure time, and the variance from read noise (σ_{read}^2) is proportional to the number of times the detector is read (i.e. the number of accumulations). Although the absolute shot noise increases with the number of detected photons ($\sigma_{shot} \propto \sqrt{counts}$), the signal increases more rapidly ($\propto$ counts), so the relative contribution of shot noise and the corresponding relative fitting uncertainty, decreases with increasing acquisition time. In other words, increasing the time spent collecting each spectrum will decrease the relative contribution of σ_{shot} to the overall error, but increase the contribution from σ_{dark}, while performing more integrations will increase the relative contribution from read noise (σ_{read}).

To determine the relative contribution of each of these noise sources, we collect 20 accumulations of 2 s, 5 s, 10 s, 30 s and 60 s with the laser and room lights off using the WITec slow series option. For each pixel (wavenumber) we calculate the variance after rejection of cosmic rays and then calculate the median variance for all pixels. Plotting this variance (σ_{signal}^2) against exposure time yields a straight line (Supplementary Material 1 Figure S19). We assume that the noise in the signal is given by $\sigma_{signal}^2 = \sigma_{read}^2 + Dt$, where D is the dark current rate in counts^2 per second. We estimate these parameters from the intercept and gradient of a linear regression. For our conditions, $\sigma_{read}^2 \sim 5.8$ cts^2 per read, and $D = 0.014$ cts^2 per second (Supplementary Material 1 Figure S19). Our detector has a gain factor (K) of 2.55, meaning that one intensity count corresponds to 2.55 electrons. Shot noise can be expressed as $\sigma_{shot}^2 = N_{counts}/K$. As shot noise scales with $\sqrt{counts}$ (the noise), and Raman signals increase proportionally with counts and thus t (the signal), the signal to noise ratio from shot noise improves with time following $\sqrt{t}$, and errors decrease proportional to $1/\sqrt{t}$. This means for a theoretically perfect detector with no read or dark noise, the total acquisition time will control the signal to noise ratio, which will be the same for 10 s × 6 accumulations as 1 × 60 s. The relatively low read and dark noise from our detector relative to shot noise means that for strong signals, this is approximately true. For example, for a single 120 s accumulation, with a signal that is 1000 counts above background:

$$\begin{aligned}\sigma_{Total}^2 &= \frac{N_{counts}}{K} + \sigma_{read}^2 + Dt \\ &= \frac{1000}{2.55} + 5.8 + 0.014 \times 120 = 399.6 \text{ cts}^2.\end{aligned} \quad (5)$$

This yields a signal to noise ratio of 50 ($1000/\sqrt{399.6}$). Shot noise accounts for 98.1 % of the error, dark noise 0.4 % and read noise 1.5 %. This is an expected result on modern, cooled CCDs where shot noise dominates (Liu et al., 2018). In contrast, for the same total acquisition time, if the spectrum was collected for n = 4 accumulations of 30 s, the signal would be 4× weaker (250 cts), yielding the following errors:

$$\begin{aligned}\sigma_{Total}^2 &= \frac{N_{counts}}{nK} + \frac{\sigma_{read}^2}{n} + \frac{Dt}{n} \\ &= \frac{250}{4 \times 2.55} + \frac{5.8}{4} + \frac{0.014 \times 30}{4} = 26 \text{ cts}^2.\end{aligned} \quad (6)$$

This yields a signal to noise ratio of 49. Shot noise makes up 94 % of the error, dark noise 0.4 % and read noise 5.6 %. The small increase in the relative contribution of read noise accompanying more accumulations is relatively negligible for strong signals, explaining why the # of accumulations has relatively little influence on the signal to noise ratio, and thus the overall peak fitting error in Figure 10. However, if we consider now a scenario where the 120 s accumulation has a signal only 100 cts above background (e.g. Supplementary Material 1 Figure S17), a single 120 s × 1 accumulation has a signal to noise ratio of 14.6 (83 % shot noise, 3.6 % dark noise and 12.4 % read noise) while 30 s × 4 accumulations with a signal of only 25 cts above background have a signal to noise ratio of 12.5 (61 % shot noise, 2.6 % dark noise and 36 % read noise). If there was no read noise, these signal to noise ratios would be the same. However, for very weak signals, the

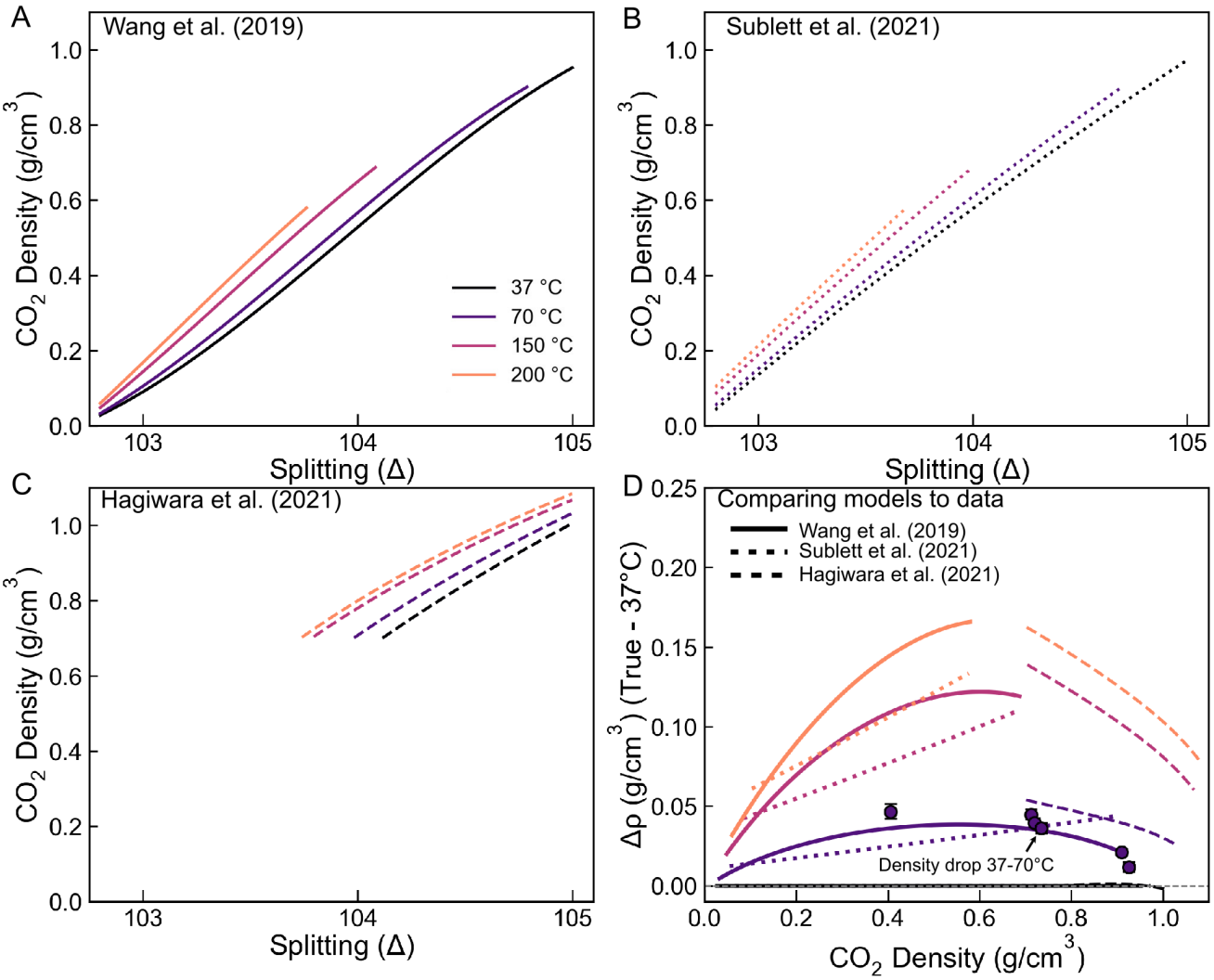


Figure 11: Models parameterizing the temperature sensitivity of the densimeter. [A-C] Relationship between splitting and CO_2 for 4 temperatures calculated using equation 3 of Wang et al. [2019] (W19, panel [A]), equation 8 of Sublett et al. [2021] (S2021, panel [B]), and equation 2 of Hagiwara et al. [2021a] (H2021, panel [C]). Lines are terminated outside the calibration range of each study. [D] Difference between true density of a FI and the density calculated using the 37 °C densimeter when analysed at t = X °C. We overlay the data from fluid inclusions we analyze at a variety of temperatures (see Figure 12).

penalty of multiple accumulations and the associated increase in read noise become significant.

Overall, these equations explain why peak fitting errors declines with $1/\sqrt{t}$ for strong signals, and there is little difference between different combinations of accumulations and acquisition times for the same total analysis time. However, for the very weakest signals, acquisition times of 30 s generate noticeably larger errors (Supplementary Material 1 Figure S16-17).

On our system, because of the decline in error following $1/\sqrt{t}$, we find that analysis times beyond about 4-5 minutes results in increasingly diminishing returns (Figure 10) and increased chance of sample drift on the stage mount causing the laser to leave the center of the inclusion. Longer acquisition times also run the risk of real drift due to changes in laboratory temperature, Raman laser wavelength, and changing sample and focus positioning. Further information on the transition from noise being dominated by photon counting vs. drift can be found in Yamamoto and Hagiwara [2024]. As errors will vary as a function of the Raman hardware and characteristics of the samples being targeted (e.g. density, fluid size etc), we suggest that performing these tests on each specific Raman instrument can help inform an analytical strategy and total acquisition time that reduce peak fitting error while remaining cost and time-effective considering the target analytical errors for any given study.

3.6 Sample temperature

It is important to investigate the effect of sample temperature on calculated densities. This has implications for how the sample should be heated during Raman analyses. Heating samples above the critical point of CO_2 (31.1 °C) ensures that a homogenous fluid can be analysed, which removes issues associated with analysis of a mixture of gaseous and liquid phases [see DeVitre et al. 2023a]. We use a low-cost heating

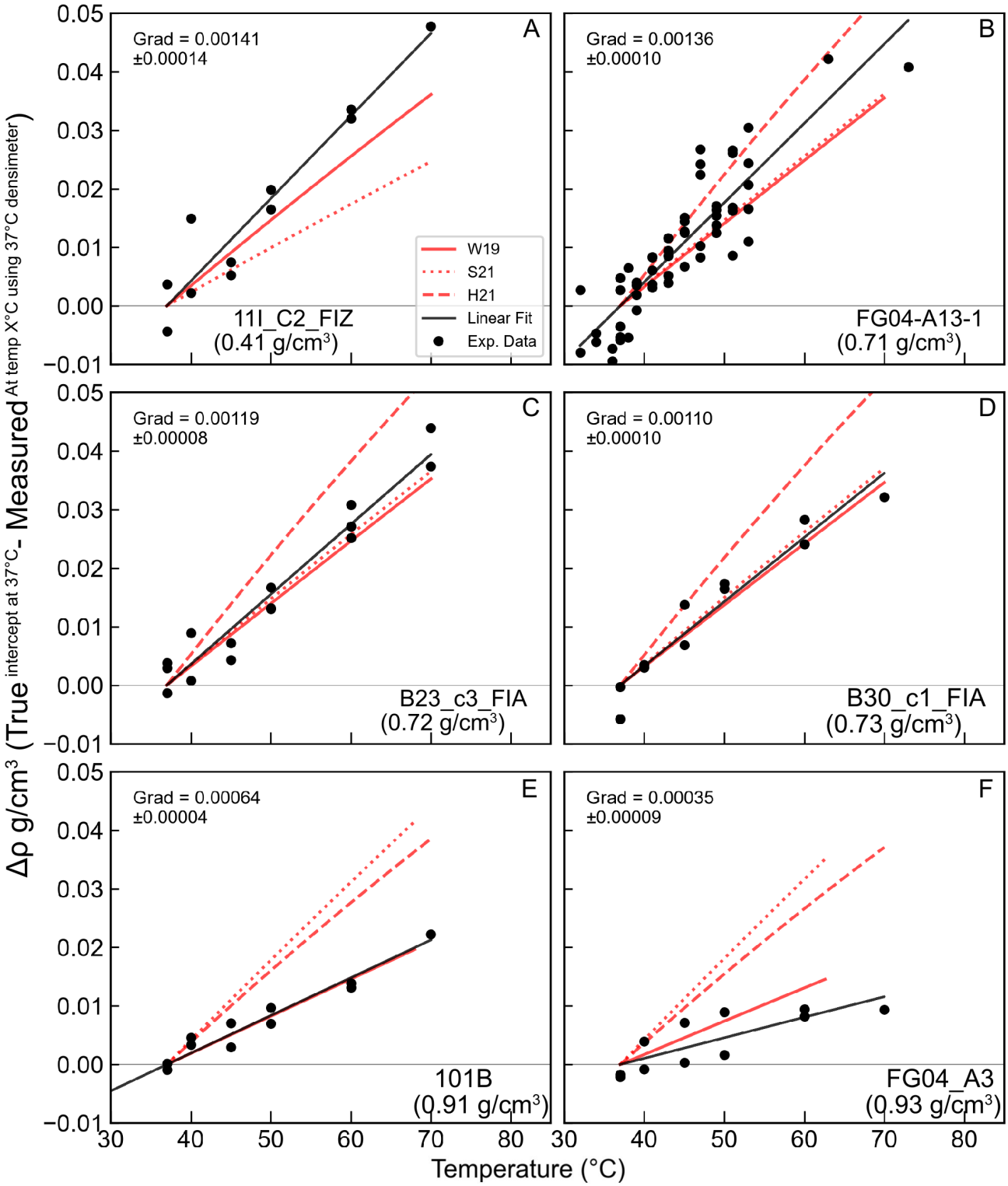


Figure 12: Changes in apparent density calculated using the 37 °C densimeter for fluid inclusions analysed at different temperatures while being heated on a Linkam stage. The 'True' density of the inclusion was determined from a regression between density and temperature, taking the value at a stage temperature of 37 °C (the temperature of our gas cell calibration). This was then compared to the calculated density at t = X °C using the 37 °C densimeter (i.e. panel [A], the density at 70 °C was 0.046 g cm^{-3} lower than the intercept at 37 °C). The solid, dotted and dashed lines show the predicted change in density using the models of W19, S21 and H21 respectively. The W19 model provides an excellent fit to all fluid inclusions.

apparatus consisting of a $60 DC power supply, a $41 Peltier plate with a hole in it to maintain the ability to use transmitted light for sample navigation, and a $24 type K thermocouple with a precision of ±2 °C attached to the surface of the sample with electrical tape (Supplementary Material 1 Figure S3). However, the limited precision of type K thermocouples means that the sample may be at 35 °C or 39 °C, even if the display reads 37 °C. Determining the error in splitting and density that results from this uncertainty in sample temperature is important to determine whether higher precision thermocouples or a more sophisticated heating apparatus is needed (i.e. a Linkam stage with a precision of ±0.1 °C).

A number of studies have investigated how sample (and thus fluid) temperature affects diad splitting. Kawakami et al. [2003] measured CO_2-rich FIs with densities of ~1 g cm^{-3} in mantle xenoliths and found that splitting stayed remarkably constant between 20–200 °C (see their Figure 10). Thus, they deduce that splitting is sensitive only to density, and is insensitive of temperature. Wang et al. [2011] investigate splitting vs. temperature for a CO_2-rich FI (0.72 g cm^{-3}) and a H_2O-CO_2 FI (see their Figure 11). Both these FI exhibit a drop in splitting of ~0.2 cm^{-1} between ~40 and 200 °C, corresponding to a drop in calculated density of 0.08 g cm^{-3}. As well as making measurements on natural FIs with specific densities, Wang et al. [2011] also investigated temperature sensitivity using a high-pressure optical apparatus where the analyte CO_2 gas was held at a controlled pressure and temperature. Their acquisitions at 21, 25, 50, 100, 150 and 200 °C indicated that Δ decreases for a given CO_2 density as temperature increases (their experiments span densities between ~0.2 and 0.7 g cm^{-3}).

More recently, Wang et al. [2019] use a high pressure optical cell to investigate the relationship between splitting and density for a pure CO_2 fluid between -20 and 360 °C, and pressures of 0.51 to 50 MPa. Their parameterization of how temperature affects splitting is shown in Figure 11A. We can also calculate the change in density that would occur if the splitting measured on a fluid heated at higher temperatures was converted into density using the 37 °C densimeter (Figure 11D). This is a scenario which may occur if a natural sample is analysed at a different temperature to the densimeter, because of a lack of heating apparatus, an issue with the thermocouple on a heating apparatus, or the heating of the fluid by the laser itself (see Section 3.7). The Wang et al. [2019] model predicts that the effect of temperature is largest for densities of ~0.5-0.6 g cm^{-3} (Figure 11D).

Sublett et al. [2021] also performed analyses on pure CO_2 in a high-pressure optical apparatus at 22 °C, 50 °C, 100 °C, 200 °C and 350 °C, for pressures spanning 10–500 bars. Their parameterization (Figure 11B) is broadly similar to that of Wang et al. [2019], except it shows a larger sensitivity to temperature at the lowest densities, and the change in calculated density vs. true density is more linear with increasing density (Figure 11C). Finally, Hagiwara et al. [2021a] (hereafter H2021) investigate the sensitivity of higher density fluids, using pressures from 7-250 MPa for 23-200 °C, yielding densities of ~0.65-1.2 g cm^{-3} (Figure 11C). This confirms the downward trend in temperature sensitivity towards higher densities indicated by Wang et al. [2019].

To assess these model predictions (Figure 11), we perform heating tests on natural olivine-hosted FIs with densities spanning 0.4-0.9 g cm^{-3} (Figure 11D). Double polished wafers were placed within a LinkamTHMS600 stage and spectra were acquired with the TruePower meter set to 6 mW for temperatures between 37 °C and 70 °C. Over this temperature window, thermal expansion of the olivine induces negligible volume and thus density changes [<0.1 % following the approach of Hagiwara et al. 2018; see also Yamamoto and Kagi 2008]. To account for the possibility of instrument drift and non-reversable changes to the inclusion by heating, analyses were run with increasing temperature and then decreasing temperature (2 repeats made at each temperature). We fit a linear regression to temperature vs. apparent density calculated using the UCB densimeter, and extrapolate this regression to calculate the density at 37 °C (Figure 12). For each acquisition, we then subtract this calculated 37 °C density from the density calculated from the measured splitting using the 37 °C densimeter, and compare this to the prediction of the three models (Figure 11D, Figure 12).

As predicted by all three models, heating causes the 37 °C densimeter to underestimate the true density, and the density change is largest for FI with densities of ~0.4-0.7 g cm^{-3}, and smallest for ones at 0.9 g cm^{-3}. The Wang et al. [2019] model does a very good job of predicting the density change for 6 fluid inclusions (Figure 11D, Figure 12).

These temperature-density gradients can be used to assess the uncertainty introduced by measuring the sample temperature. The ±2 °C error introduced from the thermocouple plus offsets of 1-2 °C we observed based on the type of tape used to stick the thermocouple down and the exact positioning of the thermocouple contributes an additional error of ~±0.0016-0.0048 g cm^{-3}, (depending on the density of the fluid being analysed). For FG04_A1_4, the average analytical error calculated in `DiadFit` is ~0.004 g cm^{-3}, while the standard deviation of hundreds of acquisitions over many years is 0.0085 g cm^{-3}. Assuming the worst-case scenario of temperature contributing the full 0.0048 g cm^{-3}, this leaves a residual error of 0.0057 g cm^{-3} unaccounted for. The remaining variability may reflect subtle variations in laser wavelength with time, laser positioning within the sample, and temperature drift which is not fully accounted for by Ne line corrections (e.g. Figure 6). However, to a first order estimate, peak fitting errors are a good estimate of the uncertainty of Raman analyses within a factor of two, and are certainly not several orders of magnitude too small [c.f. Bakker 2021].

3.7 Laser power

These heating tests give us a framework to assess the effect of laser power on CO_2 density. This has been a topic of extensive debate in the literature. Fall et al. [2011] suggest that the laser doesn't heat the inclusion more than a few 10ths of a degree, based on observations that many inclusions analysed in their laboratory which homogenized during microthermometry at temperatures very close to laboratory temperature still showed distinctive liquid and gaseous phases when analysed

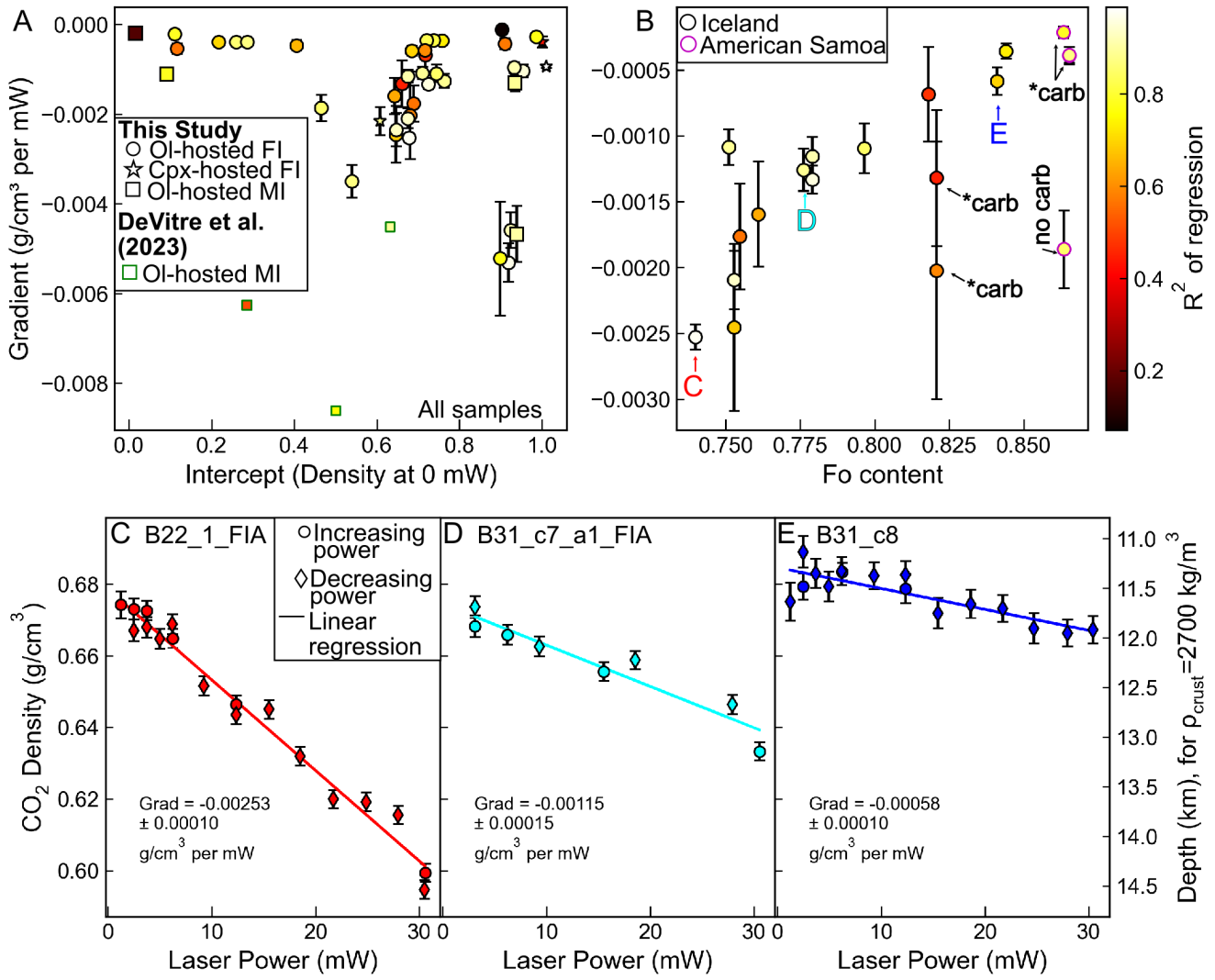


Figure 13: Relationship between CO_2 density and laser power at the sample. [A] Gradient of laser power vs. density for all the heating tests conducted here (individual regressions shown in Supplementary Material 1 Figures). Symbol shapes show the host mineral and whether the sample is a fluid inclusion or melt inclusion. Squares with green edges show the three melt inclusions from DeVitre et al. [2023a]. [B] Gradient of each power series vs. the forsterite content of the host crystal. FI with a carbonate peak are indicated with arrows. [C-E] Laser power series on three Icelandic FI indicated by arrows in [B]. The circle symbols show data collected at increasing laser power, and the diamonds at decreasing laser power. The gradients for the best-fit linear regression are shown.

by Raman spectroscopy. However, Hagiwara et al. [2021b] simulate laser heating with a numerical model, concluding that laser powers of just 10 mW can cause the temperature to rise tens or hundreds of degrees. Specifically, they show that the absorption coefficient of the host mineral is the most important parameter determining the amount of heating, because the CO_2 inclusion itself has very little absorption at the wavelengths of the laser. They perform Raman analyses at varying laser powers, and use a thermometer based on hot bands to determine the laser heating coefficient B (°C per mW of laser power) for FIs in olivine, orthopyroxene, clinopyroxene, spinel and quartz. It is notable how much lower their B values are for quartz (0.010–0.014 °C mW^{-1}) vs. mafic minerals such as olivine (0.02-1 °C mW^{-1}). In their tests, as laser power is changed from 4-14 mW, the olivine FI shows a ~0.02 g cm^{-3} decrease in density while the quartz inclusion shows no change outside uncertainty. Thus, while Fall et al. [2011] do not note the host mineralogy of the inclusions they draw their conclusions based on, the inclusion shown in their figure that they reference as evidence for a lack of laser heating is hosted in quartz. If the Fall et al. [2011] observations are from quartz inclusions, they are not in conflict with the results of Hagiwara et al. [2021b] who predict minimal heating for quartz. DeVitre et al. [2023a] also demonstrate the effect of laser heating on calculated CO_2 densities through the analyses of olivine-hosted melt inclusion vapour bubbles with densities of ~0.3, 0.5 and 0.6 g cm^{-3}. They show a clear drop in split-

ting (and thus density) with increasing power (~0.1 g cm^{-3} from 0–20 mW using TruePower readings).

It is vital to determine the effect of laser heating in different samples and bring some consensus to the literature, because it is very tempting to use high laser powers for analysis to obtain strong signals and small peak fitting errors. For example, Supplementary Material 1 Figure S20 shows the peak fitting error associated with the analysis of FIs and vapour bubbles using laser powers ranging from ~5-35 mW. A laser power of ~5 mW yields a peak fitting error 2-3× higher than a laser power of 35 mW. This reduction in error is highly beneficial for low density samples like melt inclusion vapour bubbles (often 0.02-0.04 g cm^{-3}), where peak fitting errors can introduce an uncertainty of 20–50 % in density, but this is only true if the higher laser powers do not change calculated density.

We perform analyses using TruePower values of ~5 to 50 mW (scaling to 2.2-30 mW at sample depending on the objective) for 4 olivine-hosted melt inclusion vapour bubbles from Mauna Ulu, Kīlauea [Wieser et al. 2019] and the Cascades [Wieser et al. 2023], and 29 olivine-hosted FIs from Iceland [Kahl et al. 2025], Fogo volcano [DeVitre et al. 2023b], and American Samoa (kindly supplied by Drew Downs). We also analyse 3 pyroxene-hosted FI from the Galápagos [Gleeson et al. 2025]. These MI and FI span a wide range of CO_2 densities (Figure 13). microthermometry measurements and broad wavelength scans performed by the studies cited above demonstrate that the vast majority of these inclusions are pure CO_2 [e.g. Fogo and Iceland; DeVitre and Wieser 2024; Kahl et al. 2025]. Some of the Galápagos fluid inclusions show anomalously low freezing temperatures but no other Raman-active peaks, so are inferred to contain small amounts of Raman-inactive noble gases [Gleeson et al. 2025]. We also perform power series on the FI and melt inclusion vapour bubbles that were used for heating series (Figure 14). As for temperature, we analyse each FI with increasing then decreasing laser power.

We also conduct a series of tests to determine if a steady state is reached, given that Hagiwara et al. [2021b] show that the time it takes for an inclusion to reach 90 % of its max temperature upon laser heating is 18.5 s, and that this time is heavily dependent on the boundary condition. Using the WITec slow series function where the laser remains on during the entire interval, we collect 15 spectra spaced between 5 s and 30 s, and fit each of these spectra individually. There are no significant and consistent decreases in CO_2 density or changes with increasing # of acquisitions (Supplementary Material 1 Figure S21), indicating the majority of heating occurs in the first 5-10 s after the laser is turned on. The power series show that the area ratio of the sum of the two hotbands relative to the sum of the two diads increases with increasing laser power, providing evidence that the laser is increasing the internal temperature of the inclusion (Supplementary Material 1 Figure S22). There is no notable change in the hotband to diad ratio during the slow series acquisitions, also indicating that a steady state is reached quickly (within 2-4 % of the total acquisition time for our typical 45 s × 5 accumulations, Supplementary Material 1 Figure S21).

The vast majority of fluid inclusions show a notable drop in CO_2 density as the laser power is increased (Figure 13, Supplementary Material 1 Figure S23-26). We calculate the gradient of a linear regression between the laser power at the sample and CO_2 density, and find highly variable gradients for different samples (~0 to -0.006 g cm^{-3} per mW, Figure 13A). In general, FIs show weaker negative gradients than melt inclusions from the same samples (see also Supplementary Material 1 Figure S27). However, FIs show great diversity in gradients. One fluid inclusion (FI 101C from Fogo volcano) shows a very large change in density (~0.95 to 0.8 g cm^{-3}, Figure 13D). This corresponds to a change in calculate depth of 8 km in the crust. The three Icelandic FI shown in Figure 12C-E show density change varies from −0.08 g cm^{-3} to −0.018 g cm^{-3} for a 30 mW change in laser power. The drop in density for B22_1_FIA corresponds to a change in calculated depth of ~ 3 km (Figure 12C, RH-axis). These changes in density are far larger than the uncertainty on each Raman measurement (see error bars, Figure 12C-E).

Hagiwara et al. [2021b] suggest that the absorption coefficient of the host phase is the most important parameter affecting laser heating. A wide variety of absorption coefficients for olivine have been reported in the literature, making it hard to quantitatively model the heating process. For their laser heating simulations, Hagiwara et al. [2021b] use an olivine absorption coefficient of 1.14 cm^{-1} based on measurements of a hydrothermal olivine from the Cola Peninsula, Russia [Fo_{95}; Taran and Matsyuk 2013]. In contrast, Chassé et al. [2015] report an olivine absorption coefficient at wavenumbers corresponding to a 532 nm green laser (~18800 cm^{-1}) of 11 cm^{-1} (the Fo content is not given). Other papers have shown that the absorption coefficient of olivine is sensitive to the Fo content, and the crystallographic orientation of the host [Hanzen et al. 1978; Ullrich et al. 2002; Taran and Matsyuk 2013]. For example, at 532 nm, a $Fo_{0.90}$ crystal has an absorption coefficient of ~6 cm^{-1} parallel to a vs. 3 cm^{-1} parallel to b [Ullrich et al. 2002]. For $Fo_{0.10}$, the coefficient parallel to the b-axis is ~97 cm^{-1} and 76 cm^{-1} parallel to c. For comparison, Chassé et al. [2015] show absorption coefficients for basaltic glass and melt inclusions of ~100 cm^{-1} at 532 nm. The order of magnitude higher absorption coefficient of basaltic glass vs. typical igneous olivine forsterite contents (0.70–0.90) account for the higher rates of heating seen in melt inclusions vs. fluid inclusions.

We investigate the relationship between heating rates and host forsterite contents for a subset of Icelandic and American Samoa crystals with analyses of the host mineral chemistry, given the expected increase in absorption coefficient as the olivine becomes more Fayalitic. Our data shows a strong correlation between the measured gradient and the Fo content (Figure 13B), but not between measured gradient and CO_2 density, indicating that the absorption coefficient of the host is likely controlling the amount of heating [consistent with the theoretical models of Hagiwara et al. 2021b]. However, olivine chemistry cannot explain all the variability in laser heating rates. Crystal 101 from Fogo volcano contains three FI in the same secondary trail within ~100 µm of each other. FIC experiences a very dramatic drop in density,

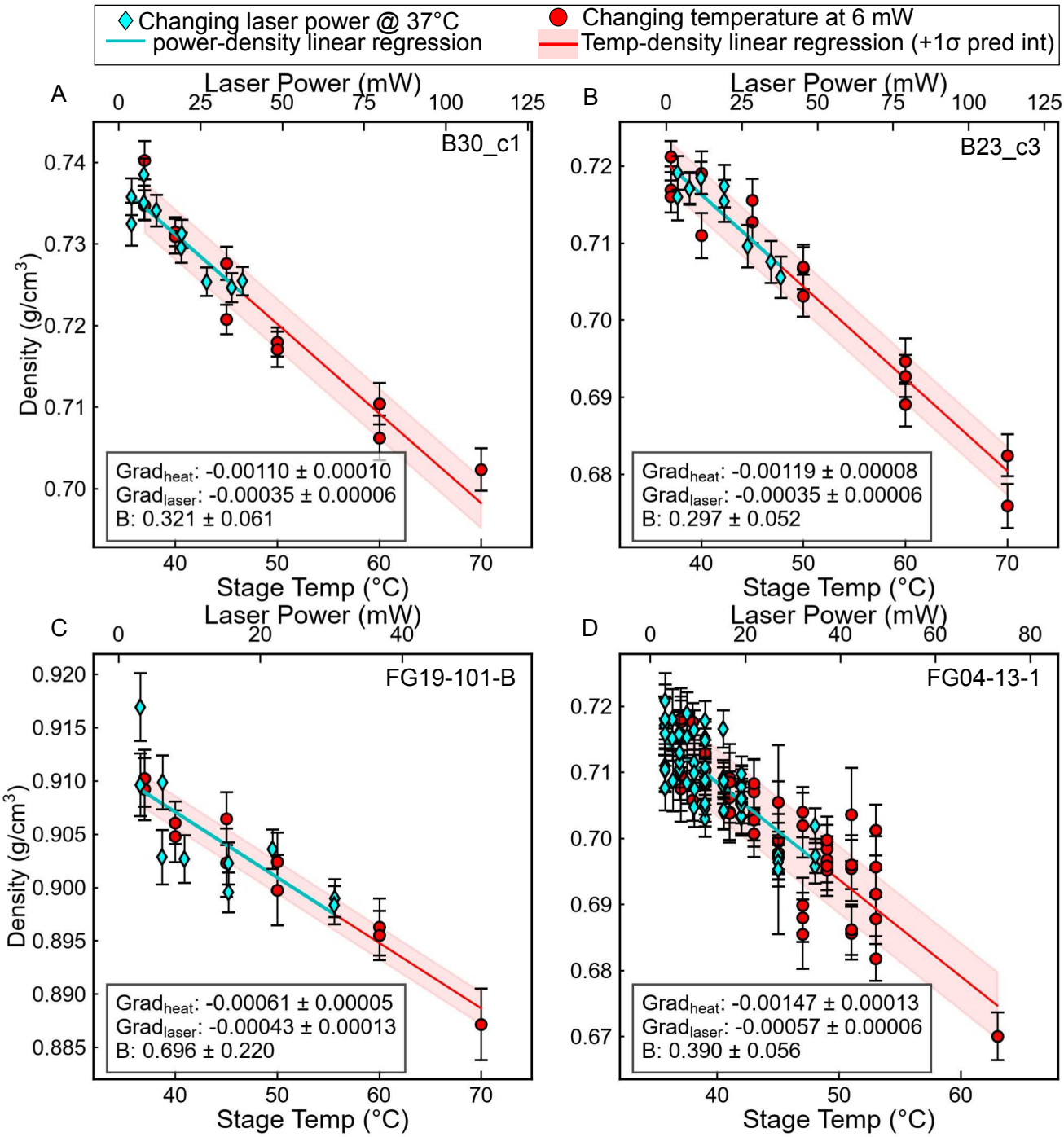


Figure 14: Comparison of heating and laser power series. For each FI, we perform acquisitions at temperatures from 37–70 °C (red circles) and 6 mW of TruePower, yielding 3.7 mW of power at sample for [A]–[C] using the 50× objective and 4.6 mW for [D] using the 40× objective. Results from changing the laser power (6 to 50 mW of TruePower) are also shown with the sample temperature controlled at 37 °C (cyan diamonds). We adjust the limits of the two x-axes such that the gradients of these two data series are identical. The gradient of each series is shown, along with the *B* value.

from ~0.93 g cm^{-3} to ~0.75 g cm^{-3} over 30 mW of heating (gradient of -0.0053 g cm^{-3} per mW). In contrast, FIB experiences no change in density outside of uncertainty (gradient of -0.0001 g cm^{-3} per mW), and FIA experiences a minor change, with a gradient of -0.001 g cm^{-3} per mW (Supplementary Material 1 Figure S26). Hagiwara et al. [2021b] show that in addition to absorption coefficient, the host crystal thickness and radius can also affect heating rates, with higher heating coefficients for larger crystals, and lower heating rates for larger host radius. However, these parameters are also the same for these three fluid inclusions. We initially speculated that the extreme heating experienced by FIC may result from its

proximity to the surface (~6 µm depth). However, when the crystal was flipped over and a power series was performed from the other side (~68 µm depth), the gradient was almost identical (Supplementary Material 1 Figure S30). 3D mapping reveals that 101C contains carbonate (Supplementary Material 1 Figure S31). It has been suggested that microparticles which are not visible optically can enhance the laser heating effect [e.g. Burke 2001], so it is possible that carbonate is playing this role in this inclusion. The two Icelandic samples that show higher heating than expected based on their Fo contents (Figure 12B) also contain notable carbonate in 3D maps. However, of the three FI from American Samoa (pink edges, Figure 12B), the two which follow the gradient-Fo trend defined by the Icelandic samples contain abundant carbonate, while the one with more negative gradients contains no carbonate. The three Mg# 79 pyroxenes also show very variable heating amounts (Supplementary Material 1 Figure S32), despite the fact these samples are very homogenous and there are no apparent color differences between crystals [Gleeson et al. 2025].

As predicted from the modelling of Hagiwara et al. [2021b], we also note that the gradient depends on the heating apparatus used, which controls the boundary condition by which the crystal loses heat. When FI 101C is heated when embedded in crystalbond on a glass slide, it experiences a larger amount of heating (gradient of ~0.008 g cm^{-3} mW^{-1}) than when placed on the Linkam stage (Supplementary Material 1 Figure S33). Overall, these power series demonstrate that while laser heating can be partially predicted based on the absorption coefficient of the host, we still observe extreme large changes in density for specific FI, replicated across different analytical sessions. Given that many melt and FIs are analysed prior to knowing the Fo content (as that is often the final step of measurement) and that we cannot diagnose why a few FI show such extreme heating, this work demonstrates that it is absolutely vital to use low laser powers, as the heating response is currently not entirely predictable. Thus, we recommend each laboratory investigate laser heating using similar tests to those performed here for new sample suites to determine the correct laser power to use, as the heating response is currently not predictable.

Performing laser heating and laser power tests on the same FI within the Linkam stage allows us to estimate the heating coefficient, B (°C mW^{-1} of laser power) for this specific geometry (B values will vary depending on the exact mounting mechanism used and the resulting boundary conditions). For each FI we subsample the power series to be within the range of densities seen in the heating series, and calculate the gradient of density vs. temperature, and density vs. laser power. On Figure 14 (and Supplementary Material 1 Figure S34), we adjust the power and temperature scale such that the gradients of the two datasets align. We calculate B as:

$$B = \frac{\text{slope of density vs. laser power (g cm}^{-3}\text{ mW}^{-1})}{\text{slope of density vs. temperature (g cm}^{-3}\text{ °C}^{-1})} \quad (7)$$

This is valid for heating vs. power series collected on a single fluid inclusion where the sample and FI geometry, density and fluid composition are fixed, assuming that the two series cover a similar temperature range. This method also assumes a steady state is reached rapidly during exposure to the laser, and that the temperature and power results are linear over the regression ranges used (see Figure 14).

B-value errors were derived from the gradient uncertainties of each ordinary least squares (OLS) linear regression, computed via `scipy.stats.linregress`, and propagated in quadrature. B values mostly span ~0.3-0.7 °C mW^{-1}, although a few exceed 1 (Supplementary Material 1 Figure S34). This compares to 0.02-1°C mW^{-1} from Hagiwara et al. [2021b], and ~0.36 °C mW^{-1} from Wang et al. [2024].

Overall, we suggest that users should conduct analyses of a subset of any new samples at different laser powers, to determine the change in density as a function of power for their choice of analytical conditions (the heating rate will vary as a function of the host phase, host geometry, and possibly the method used to mount the samples during analysis if heat sinks are present). By determining the gradient of laser power vs. density (Figure 13), the user can determine the maximum acceptable power that keeps density changes within a range so as not to compromise the objectives of their study. For example, for our Icelandic inclusions which exhibit density-power gradients of 0.0004-0.003 g cm^{-3} per mW (Figure 13B), laser powers of <3-4 mW limit the effects of laser heating to ~0.01 g cm^{-3}, which is comparable to one standard deviation of repeated standard measurements in our laboratory (0.009 g cm^{-3}; Figure 5), and offsets between microthermometry and Raman Spectroscopy [DeVitre and Wieser 2024].

3.8 Other strategies to optimize signal

Having to use low laser powers to minimise heating is a real issue in terms of getting enough signal to minimise peak fitting errors. In addition to optimizing the acquisition times (Figure 10), it is worth thinking about other ways to optimize signals. Our repeated measurements of FI standards (Figure 5) not only allow us to track the stability of the densimeter, but also the intensity of signals with time. If the FI standard returns lower peak heights than normal, we analyse a Si chip which yields ~2000 cts for a 10 s acquisition when the system is well aligned. We use the oscilloscope feature ('live spectra' on HORIBA) to perform micro adjustments on the pinhole position to enhance laser transmission through the fibre to achieve these counts. We find that these adjustments can easily increase signals by a factor of 2-3. We typically adjust the system every few weeks, although sometimes it goes months with no decline in signal, and other times we have to readjust after a few days. Typically, adjustments are needed after HVAC failures where the temperature of the laboratory has fluctuated, suggesting this adjustment is required based on thermal expansion/contraction of components. On a mirror-based instrument, mirror alignments will have a similar effect.

Once the system's optics are optimized, it is worth considering the analytical geometry. The depth of the FI within the sample has been discussed in a variety of papers. Frezzotti et al. [2012] note that the depth resolution of Raman spectroscopy degrades with depth in the sample, so it is best to analyse FI within 30 µm of the surface. Burke [2001] states that '*usually*

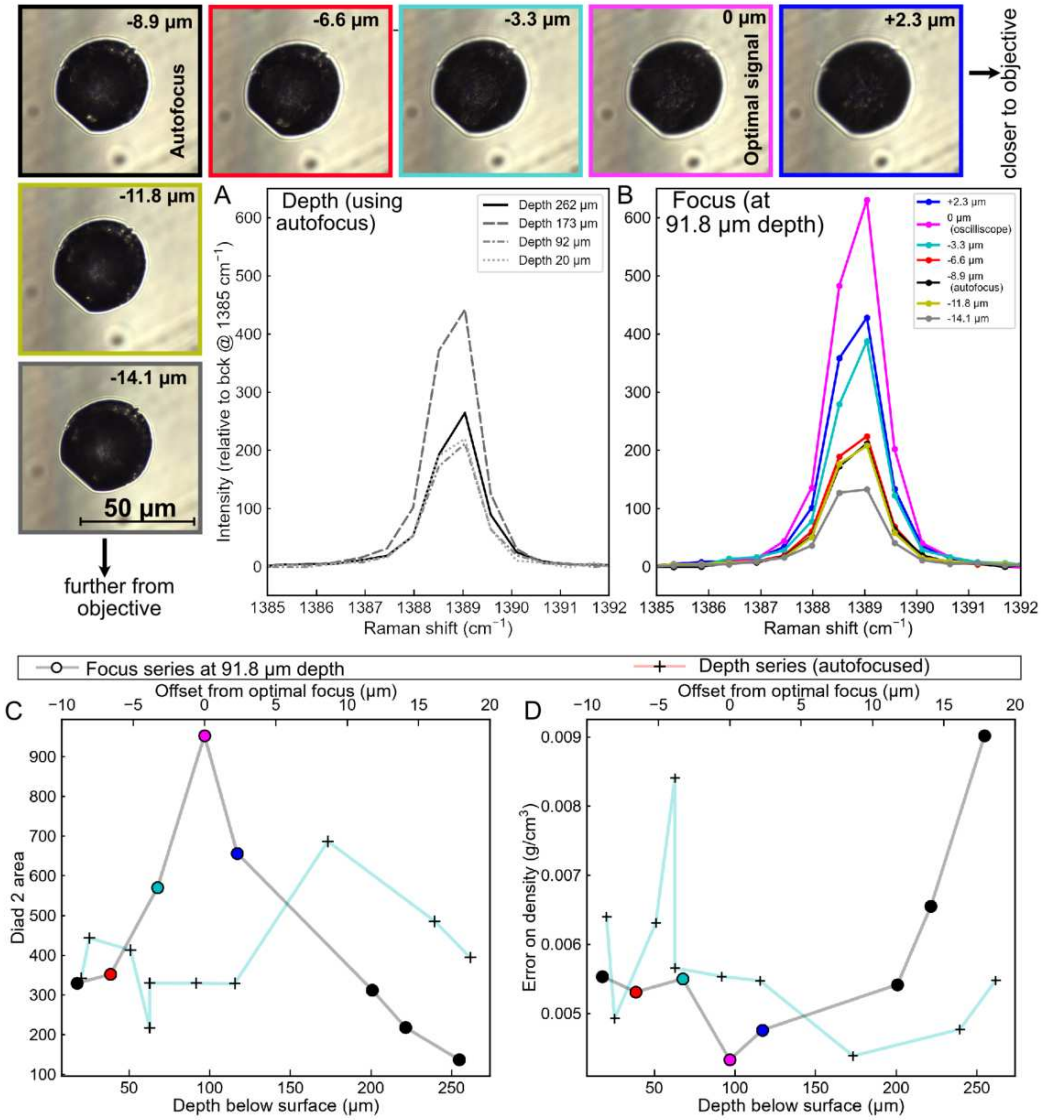


Figure 15: Investigation of how signal strengths are affected by focus and depth in the sample. [A] Region of the spectra surrounding diad2, showing the signal strength for different depths of the FI in the sample during serial sectioning. After each polishing round, the autofocus function was used to find the optimum depth. [B] When the FI was at 91.8 µm depth in the sample, we acquired spectra at 7 different focus positions. The black spectra and dot represents the focus position found by the autofocus, and the pink spectra and dot the best focus position using the oscilloscope feature. The images on the left-hand side show what the FI looked like at each focus. [C] Change in diad2 intensity and [D] density error as a function of depth in the sample (circles) and focus (crosses). It is apparent that small changes in focus (top x-axis) overwhelm the changes from inclusion depth (bottom x-axis).

one strives to analyse inclusions as near the surface as possible inclusions deeper than 50 µm start to give problems, especially if their density is not too high'. To assess the influence of FI depth, we individually ground 3 olivine-hosted FI down, performing analyses after each round of grinding and polishing. We measure the depth by focusing on the surface of the sample in reflected light, and then focusing down to the melt inclusion. We correct the movement of the Z-stage to true depth accounting for the refractive index of olivine (N_{olivine} ~ 1.7, 1.64–1.67 for Fo, 1.84–1.89 for Fa):

$$\text{True depth} = N_{\text{olivine}} \times Z_{\text{stage movement}}. \quad (8)$$

We use the WITec autofocus function at each depth. Surprisingly, we find no coherent trends in signal intensity (quantified here using the area of diad2) or the error on calculated density with increasing depth in the sample (Figure 15A, Supplementary Material 1 Figure S35–37). However, we do find big changes with minor changes in focus at a single sample depth. To demonstrate this, we perform acquisitions at different focuses when the FI was 91.8 µm deep in the sample.

We first find the optical focus yielding a sharp inclusion edge (yellow box and spectra, Figure 15). We then use the WITec autofocus feature at our maximum laser power. This collects spectra for ~3 s at various depths to find the maximum intensity in user-selected spectral window. The Z-position was ~ 2 µm close to the surface than the best optical focus and yielded very similar signal strengths (black spectra and box). We then used the oscilloscope function to manually scroll through Z-positions to find the best signal (pink box and spectra). The 6 focus positions between −14.1 µm and 0 µm all yield optically reasonable images, and yet the signal strength varies by a factor of 3. Interestingly, the blue focus position (+2.3 µm) yields an optically poor image, but still a better focus than the oscilloscope feature found. This comparison clearly demonstrates that focus has far more effect on spectral strength than changing the depth in the sample by ~250 µm (Figure 15C,D).

This result was replicated for several inclusions (Supplementary Material 1 Figure S35–S37). We also find that the signal can vary quite considerably spatially, and small shifts in x-y position can also result in stronger signals. This becomes very apparent when taking 2D transects through and across FI (Supplementary Material 1 Figure S38–39). We suggest that when signals are weak, the autofocus or oscilloscope feature should be used to investigate different x-y-z positions around the optimal focus. Variability in signal strength can also be seen by performing 2D and 3D maps of signal intensity. We perform these tests at high laser power (~30 mW at sample) for speed, and the drop the power to perform the analyses at the optimal Z-position. Our samples do not show any evidence for damage when exposed to high laser powers, but care should be taken with new samples which may experience Fe-nanolite formation [e.g. if glass is exposed to high laser powers; Di Genova et al. 2018], or possibly even decrepitation (we have never personally observed this occurring).

4 Conclusions

In summary, we investigate the effect of several instrument parameters and data reduction methods on the accuracy and precision of Raman analyses. Our findings can be summarized as follows:

1. Spectral non-linearity means that the spectral window/center and drift correction method (e.g. choice of Ne and Ar lines) used for densimeter calibration must be replicated when analysing unknowns in future sessions. Spectral resolution alone cannot account for the diversity in published densimeters, shown by the fact even a single instrument can generate a wide variety of densimeter relationships as the drift correction strategy or spectral center is adjusted. We instead speculate that densimeter diversity in the literature is likely related to non-linearity, causing the splitting to vary depending on which CCD pixels the CO_2 diad is projected onto.

2. Estimates of peak fitting error from Pseudovoigt curves, combined with uncertainty in the Ne line correction model, account for approximately half the scatter seen in repeated Raman analyses of the same inclusion over ~ 2.5 years (~0.004 g cm^{-3} for FI with densities of >0.6 g cm^{-3}). These low uncertainties make suggestions of very large peak fitting errors untenable. The additional 0.004 g cm^{-3} of scatter seen in repeated standard acquisitions likely results from uncertainty in the temperature of the sample and room, combined with variations in laser wavelength and other parameters with time.

3. We recommend tracking the density of a single standard during every run to ensure there has been no drift from the laboratory densimeter relationship. These standards also help to track the efficiency of the instrument (e.g. checking the peak heights are as high as previous days), so the user knows when re-alignment is required.

4. In a temperature-controlled laboratory, drift during and between analytical sessions was minimal using our WITec alpha300R Raman. However, when the laboratory temperature is perturbed, the Ne line correction helps to reduce instrument drift by a factor of ~3. Thus, while it is better to have a Raman in a temperature-controlled laboratory, it is not essential. Large temperature swings of +5 °C result in an additional uncertainty of <0.01 g cm^{-3} if Ne lines are used for drift correction.

5. Laser heating is a significant issue, particularly for materials with higher optical absorption coefficients (e.g. silicate glass, lower Fo olivine). For all but the most strongly laser-heated FI investigated here, ~3-4 mW of laser power at the sample changes the calculated density by less than the uncertainty on repeated FI measurements (<0.009 g cm^{-3}, Figure 5) while still providing sufficient signal to obtain good peak fits. However, the amount of laser heating should be investigated when working on new sample suites given variable heating rates observed in this study, even within a single crystal.

6. To help counteract the relatively weak signals achieved using low laser powers, the analyst should investigate the play off between uncertainty and acquisition time for their specific instrument. They should experiment with different focuses using an autofocus/live spectra function of the oscilloscope at high power to find the optimum signal within the fluid. Focus has a far larger effect than depth to the fluid within the sample (even at depths of ~250 µm). Care should be taken with this approach, as some samples may be sensitive to decrepitation and damage (e.g. Fe nanolite formation in glass) when exposed to high laser powers.

We suggest that prior to performing Raman density measurements, users conduct the various tests here to optimize analytical conditions:

1. Choose a spectral window over which to collect data, and if for any reason you want to change this after calibration, investigate the impact of this change on the density of a natural or synthetic FI.

2. Use a temperature logger to explore variations in laboratory temperature. If the temperature fluctuates by several degrees, perform repeated analysis of a natural or synthetic fluid inclusion, and test whether the selected drift correction method (Ne lines or repeated standard measurements) minimises this drift or not.

3. Investigate the relationship between laser power and signal using the available power options, and between analytical errors and acquisition time and accumulations.

4. For the samples of interest (with their different absorption coefficients), investigate the effects of laser heating to choose a laser power that gives sufficient signal without inducing density changes.

Author Contributions

PW and CLD conceived the project. PW, CLD and IS collected the Raman data and performed peak fitting. PW wrote the manuscript. All authors contributed to the discussion of ideas and the editing of the manuscript.

Acknowledgements

PW and CD acknowledge funding from NSF EAR 2217371, EAR 2342156, the Sloan foundation, the Berkeley Rose Hills Innovator Program and the Berkeley Heising Simons Faculty Fellowship. IS was funded through a SURF and Ramsden summer fellowship at UCB. The authors are very grateful to the producers of `lmfit`, without which this study would not have been possible. We thank Maren Kahl for supplying the Icelandic samples, Drew Downs for the American Samoa samples, and Matthew Gleeson for the Galápagos samples. We are grateful to Alex Bearden for his early contributions to investigating laser heating in some Mauna Loa vapour bubbles. We thank Charles le Losq and two anonymous reviewers for extremely detailed comments and suggestions that greatly improved the manuscript.

Data Availability

All data and Jupyter Notebooks needed to recreate the figures in this study are provided in the Github repository (https://github.com/PennyWieser/Raman_Best_Practices). We also supply a supporting spreadsheet (Supplementary Material 2) containing processed data for each figure (e.g. peak fits etc).

Copyright notice

© The Author(s) 2026. This article is distributed under the terms of the Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

Allison, C. M., K. Roggensack, and A. B. Clarke (2021). "Highly explosive basaltic eruptions driven by CO_2 exsolution". *Nature Communications* 12(1), page 217. DOI: 10.1038/s41467-020-20354-2.

Anderson, O. E., M. G. Jackson, A. S. Pamukçu, E. F. Rose-Koga, V. Le Roux, F. Klein, K. T. Koga, G. A. Gaetani, and A. A. Price (2024). "Extensive H_2O degassing in deeply erupted submarine glasses inferred from Samoan melt inclusions: The EM2 mantle source is damp, not dry". *Chemical Geology* 651, page 121979. DOI: 10.1016/j.chemgeo.2024.121979.

Bakker, R. J. (2021). "The perfection of Raman spectroscopic gas densimeters". *Journal of Raman Spectroscopy* 52(11), pages 1923–1948. DOI: 10.1002/jrs.6245.

Bearden, A., P. Wieser, C. DeVitre, C. Gordon, K. Lynn, and B. Monteleone (2026). "Petrological, Textural, and Volatile Constraints on the Architecture of Mauna Loa's Magmatic System over 170 years (1852–2022)". *Journal of Petrology*, agag086. DOI: 10.1093/petrology/egag086.

Burke, E. A. (2001). "Raman microspectrometry of fluid inclusions". *Lithos* 55(1-4), pages 139–158. DOI: 10.1016/s0024-4937(00)00043-8.

Chassé, M., G. Lelong, P. van Nijnatten, I. Schoofs, J. de Wolf, L. Galoisy, and G. Calas (2015). "Optical Absorption Microspectroscopy (μ-OAS) Based on Schwarzschild-Type Cassegrain Optics". *Applied Spectroscopy* 69(4), pages 457–463. DOI: 10.1366/14-07628.

Dayton, K., E. Gazel, P. Wieser, V. R. Troll, J. C. Carracedo, H. La Madrid, D. C. Roman, J. Ward, M. Aulinas, H. Geiger, F. M. Deegan, G. Gisbert, and F. J. Perez-Torrado (2023). "Deep magma storage during the 2021 La Palma eruption". *Science Advances* 9(6), eade7641. DOI: 10.1126/sciadv.ade7641.

DeVitre, C. and P. Wieser (2024). "Reliability of Raman analyses of CO_2-rich fluid inclusions as a geobarometer at Kīlauea". *Geochemical Perspectives Letters* 29, pages 1–8. DOI: 10.7185/geochemlet.2404.

DeVitre, C. L., C. M. Allison, and E. Gazel (2021). "A high-precision CO_2 densimeter for Raman spectroscopy using a Fluid Density Calibration Apparatus". *Chemical Geology* 584, page 120522. DOI: 10.1016/j.chemgeo.2021.120522.

DeVitre, C. L., K. Dayton, E. Gazel, A. Pamukçu, G. Gaetani, and P. E. Wieser (2023a). "Laser heating effect on Raman analysis of CO_2 co-existing as liquid and vapor in olivine-hosted melt inclusion bubbles". *Volcanica* 6(2), pages 201–219. DOI: 10.30909/vol.06.02.201219.

DeVitre, C. L., E. Gazel, R. S. Ramalho, S. Venugopal, M. Steele-MacInnis, J. Hua, C. M. Allison, L. R. Moore, J. C. Carracedo, and B. Monteleone (2023b). "Oceanic intraplate explosive eruptions fed directly from the mantle". *Proceedings of the National Academy of Sciences* 120(33), e2302093120. DOI: 10.1073/pnas.2302093120.

Di Genova, D., A. Caracciolo, and S. Kolzenburg (2018). "Measuring the degree of "nanotilization" of volcanic glasses: Understanding syn-eruptive processes recorded in melt inclusions". *Lithos* 318-319, pages 209–218. DOI: 10.1016/j.lithos.2018.08.011.

Fall, A., B. Tattitch, and R. J. Bodnar (2011). "Combined microthermometric and Raman spectroscopic technique to determine the salinity of H_2O–CO_2–NaCl fluid inclusions based on clathrate melting". *Geochimica et Cosmochimica Acta* 75(4), pages 951–964. DOI: 10.1016/j.gca.2010.11.021.

Fermi, E. (1931). "Über den Ramaneffekt des Kohlendioxyds". *Zeitschrift für Physik* 71(3-4), pages 250–259. DOI: 10.1007/bf01341712.

Frezzotti, M. L., F. Tecce, and A. Casagli (2012). "Raman spectroscopy for fluid inclusion analysis". *Journal of Geochemical Exploration* 112, pages 1–20. DOI: 10.1016/j.gexplo.2011.09.009.

Fukura, S., T. Mizukami, S. Odake, and H. Kagi (2006). "Factors Determining the Stability, Resolution, and Precision of a Conventional Raman Spectrometer". *Applied Spectroscopy* 60(8), pages 946–950. DOI: 10.1366/000370206778062165.

Gleeson, M., P. E. Wieser, C. L. DeVitre, S. C. Shi, M.-A. Millet, D. D. Muir, M. J. Stock, and J. Lissenberg (2025). "Persistent High-Pressure Magma Storage beneath a Near-Ridge Ocean Island Volcano (Isla Floreana, Galápagos)". *Journal of Petrology* 66(5), egaf031. DOI: 10.1093/petrology/egaf031.

Guo, X., A. Mughal, D. Dunphy, G. Stone, D. Miller, and S. Huh (2020). "Ambient Temperature–Corrected Mechanical Stress Mapping of Gallium Nitride and Aluminum Indium Gallium Phosphide Films by Raman Scattering Spectroscopy for Characterization of Light-Emitting Diodes". *Physica Status Solidi A* 217(7), page 1900776. DOI: 10.1002/pssa.201900776.

Hagiwara, Y., T. Kawano, K. Takahata, J. Torimoto, and J. Yamamoto (2021a). "Temperature dependence of a Raman CO_2 densimeter from 23°C to 200°C and 7.2 to 248.7 MPa: Evaluation of density underestimation by laser heating". *Journal of Raman Spectroscopy* 52(10), pages 1744–1757. DOI: 10.1002/jrs.6188.

Hagiwara, Y., Y. Sogo, K. Takahata, and J. Yamamoto (2018). "Temperature dependence of CO_2 densimetry using micro-Raman Spectrometry at laboratory conditions". *Geochemical Journal* 52(4), pages 379–383. DOI: 10.2343/geochemj.2.0523.

Hagiwara, Y., J. Torimoto, and J. Yamamoto (2020). "Pressure measurement and detection of small H_2O amounts in high-pressure H_2O–CO_2 fluid up to 141 MPa using Fermi diad splits and bandwidths of CO_2". *Journal of Raman Spectroscopy* 51(6), pages 1003–1018. DOI: 10.1002/jrs.5865.

Hagiwara, Y., L. Yokokura, and J. Yamamoto (2023). "Unlocking ultimate precision of intensity and area ratio measurements in Raman spectroscopy: Insights from simulation, experimentation, and theory and implications for isotope ratio analysis". *Journal of Raman Spectroscopy* 54(12), pages 1440–1464. DOI: 10.1002/jrs.6594.

Hagiwara, Y., K. Yoshida, A. Yoneda, J. Torimoto, and J. Yamamoto (2021b). "Experimental variable effects on laser heating of inclusions during Raman spectroscopic analysis". *Chemical Geology* 559, page 119928. DOI: 10.1016/j.chemgeo.2020.119928.

Hanzen, R., P. Bell, and H. Mao (1978). "Effects of compositional variation on absorption spectra of lunar pyroxenes". *9th Lunar and Planetary Science Conference*, pages 2919–2934.

Howe, T. A., T. E. Christopher, S. Moune, H. Tuffen, and F. Schiavi (2025). "Melt inclusion bubbles provide new insights into crystallisation depths and CO_2 systematics at Soufrière Hills Volcano, Montserrat". *Frontiers in Earth Science* 12, page 1509409. DOI: 10.3389/feart.2024.1509409.

Inoue, Y., R. Okiyama, Y. Hagiwara, and J. Yamamoto (2023). "Raman spectroscopic evaluation of precision of oxygen isotope ratio of carbon dioxide". *Geochemical Journal* 57(2), pages 92–99. DOI: 10.2343/geochemj.gj23007.

Jakubek, R. S. and M. D. Fries (2022). "Calibration of the temporal drift in absolute and relative Raman intensities in large Raman images using a mercury–argon lamp". *Journal of Raman Spectroscopy* 53(1), pages 137–147. DOI: 10.1002/jrs.6259.

Kahl, M., D. J. Morgan, P. E. Wieser, E. Bali, G. H. Guðfinnsson, D. A. Neave, and R. Walshaw (2025). "Crystal-mush remobilization timescales and magma storage depth in the Snæfellsnes Volcanic Zone (W-Iceland): insights from olivine Fe-Mg diffusion chronometry and fluid inclusion barometry". *Bulletin of Volcanology* 87(12), page 118. DOI: 10.1007/s00445-025-01892-3.

Kawakami, Y., J. Yamamoto, and H. Kagi (2003). "Micro-Raman Densimeter for CO_2 Inclusions in Mantle-Derived Minerals". *Applied Spectroscopy* 57(11), pages 1333–1339. DOI: 10.1366/000370203322554473.

Kobayashi, T., J. Yamamoto, T. Hirajima, H. Ishibashi, N. Hirano, Y. Lai, V. S. Prikhod'ko, and S. Arai (2012). "Conformity and precision of CO_2 densimetry in CO_2 inclusions: microthermometry versus Raman microspectroscopic densimetry". *Journal of Raman Spectroscopy* 43(8), pages 1126–1133. DOI: 10.1002/jrs.3134.

Kramida, A. and Y. Ralchenko (1999). *NIST Standard Reference Database 78*. Atomic Spectra Database. National Institute of Standards and Technology. DOI: 10.18434/T4W30F.

Lamadrid, H., L. Moore, D. Moncada, J. Rimstidt, R. Burruss, and R. Bodnar (2017). "Reassessment of the Raman CO_2 densimeter". *Chemical Geology* 450, pages 210–222. DOI: 10.1016/j.chemgeo.2016.12.034.

Le, V.-H., M.-C. Caumon, and A. Tarantola (2021). "FRAnCIs calculation program with universal Raman calibration data for the determination of PVX properties of CO_2–CH_4–N_2 and CH_4–H_2O–NaCl systems and their uncertainties". *Computers & Geosciences* 156, page 104896. DOI: 10.1016/j.cageo.2021.104896.

Lin, F., R. Bodnar, and S. Becker (2007). "Experimental determination of the Raman CH_4 symmetric stretching (ν_1) band position from 1–650bar and 0.3–22°C: Application to fluid inclusion studies". *Geochimica et Cosmochimica Acta* 71(15), pages 3746–3756. DOI: 10.1016/j.gca.2007.05.016.

Liu, D. and B. M. Hennelly (2024). "Wavenumber Calibration Protocol for Raman Spectrometers Using Physical Modelling and a Fast Search Algorithm". *Applied Spectroscopy* 78(8), pages 790–805. DOI: 10.1177/00037028241254847.

McLennan, J. C. and H. D. Smith (1932). "Raman effects with liquid and solid carbon dioxide". *Canadian Journal of Research* 7(6), pages 551–555. DOI: 10.1139/cjr32-104.

Mestari, A., R. Gaufrès, and P. Huguet (1997). "Behaviour of the calibration of a Raman spectrometer with temperature changes". *Journal of Raman Spectroscopy* 28(10), pages 785–789. DOI: 10.1002/(sici)1097-4555(199710)28:10<785::aid-jrs148>3.0.co;2-d.

Moore, L. R., E. Gazel, R. Tuohy, A. S. Lloyd, R. Esposito, M. Steele-MacInnis, E. H. Hauri, P. J. Wallace, T. Plank, and R. J. Bodnar (2015). "Bubbles matter: An assessment of the contribution of vapor bubbles to melt inclusion volatile budgets". *American Mineralogist* 100(4), pages 806–823. DOI: 10.2138/am-2015-5036.

Newville, M., T. Stensitzki, D. B. Allen, M. Rawlik, A. Ingargiola, and A. Nelson (2016). *Lmfit: Non-Linear Least-Square Minimization and Curve-Fitting for Python*. Astrophysics Source Code Library.

Remigi, S., T. Mancini, S. Ferrando, and M. L. Frezzotti (2021). "Interlaboratory Application of Raman CO_2 Densimeter Equations: Experimental Procedure and Statistical Analysis Using Bootstrapped Confidence Intervals". *Applied Spectroscopy* 75(7), pages 867–881. DOI: 10.1177/0003702820987601.

Rosso, K. and R. Bodnar (1995). "Microthermometric and Raman spectroscopic detection limits of CO_2 in fluid inclusions and the Raman spectroscopic characterization of CO_2". *Geochimica et Cosmochimica Acta* 59(19), pages 3961–3975. DOI: 10.1016/0016-7037(95)94441-h.

Song, Y., I. Chou, W. Hu, B. Robert, and W. Lu (2009). "CO_2 Density-Raman Shift Relation Derived from Synthetic Inclusions in Fused Silica Capillaries and Its Application". *Acta Geologica Sinica* 83(5), pages 932–938. DOI: 10.1111/j.1755-6724.2009.00090.x.

Span, R. and W. Wagner (1996). "A New Equation of State for Carbon Dioxide Covering the Fluid Region from the Triple-Point Temperature to 1100 K at Pressures up to 800 MPa". *Journal of Physical and Chemical Reference Data* 25(6), pages 1509–1596. DOI: 10.1063/1.555991.

Sublett, D. M., E. Sendula, H. M. Lamadrid, M. Steele-MacInnis, G. Spiekermann, and R. J. Bodnar (2021). "Raman spectral behavior of N_2, CO_2, and CH_4 in N_2–CO_2–CH_4 gas mixtures from 22°C to 200°C and 10 to 500 bars, with application to other gas mixtures". *Journal of Raman Spectroscopy* 52(3), pages 750–769. DOI: 10.1002/jrs.6033.

Taran, M. N. and S. S. Matsyuk (2013). "Fe^{2+}, Mg-distribution among non-equivalent structural sites M1 and M2 in natural olivines: an optical spectroscopy study". *Physics and Chemistry of Minerals* 40(4), pages 309–318. DOI: 10.1007/s00269-013-0572-x.

Ullrich, K., K. Langer, and K. D. Becker (2002). "Temperature dependence of the polarized electronic absorption spectra of olivines. Part I - fayalite". *Physics and Chemistry of Minerals* 29(6), pages 409–419. DOI: 10.1007/s00269-002-0248-4.

Venugopal, S., F. Schiavi, S. Moune, N. Bolfan-Casanova, T. Druitt, and G. Williams-Jones (2020). "Melt inclusion vapour bubbles: the hidden reservoir for major and volatile elements". *Scientific Reports* 10(1), page 9034. DOI: 10.1038/s41598-020-65226-3.

Wang, H., W. Lu, W. Wang, Q. Liu, and J. Yamamoto (2024). "High-precision analysis of carbon isotopic composition for individual CO_2 inclusions via Raman spectroscopy: Addressing issues arising from the laser-heating effects". *Chemical Geology* 651, page 122014. DOI: 10.1016/j.chemgeo.2024.122014.

Wang, W., M.-C. Caumon, A. Tarantola, J. Pironon, W. Lu, and Y. Huang (2019). "Raman spectroscopic densimeter for pure CO_2 and CO_2-H_2O-NaCl fluid systems over a wide P-T range up to 360 °C and 50 MPa". *Chemical Geology* 528, page 119281. DOI: 10.1016/j.chemgeo.2019.119281.

Wang, X., I.-M. Chou, W. Hu, R. C. Burruss, Q. Sun, and Y. Song (2011). "Raman spectroscopic measurements of CO_2 density: Experimental calibration with high-pressure optical cell (HPOC) and fused silica capillary capsule (FSCC) with application to fluid inclusion observations". *Geochimica et Cosmochimica Acta* 75(14), pages 4080–4093. DOI: 10.1016/j.gca.2011.04.028.

Wieser, P., A. Kent, C. DeVitre, E. Gazel, C. Till, E. Johnson, P. Wallace, and F. Couperthwaite (2023). "Bubble Trouble: Raman measurements of Cascade melt inclusion vapour bubbles indicate substantial underestimation of magma storage depths in legacy data". *AGU Fall Meeting*. V23D-0203.

Wieser, P. E., M. Edmonds, J. Maclennan, F. E. Jenner, and B. E. Kunz (2019). "Crystal scavenging from mush piles recorded by melt inclusions". *Nature Communications* 10(1), page 5797. DOI: 10.1038/s41467-019-13518-2.

Wieser, P. E., S. C. Shi, M. L. M. Gleeson, B. Rangel, C. L. DeVitre, A. T. Bearden, K. J. Lynn, and M. Camille-Caumon (2025). "Fluid inclusion constraints on the geometry of the magmatic plumbing system beneath Mauna Loa — Part 1:lavas and tephras". *Bulletin of Volcanology* 87(10), page 89. DOI: 10.1007/s00445-025-01874-5.

Wojdyr, M. (2010). "*Fityk*: a general-purpose peak fitting program". *Journal of Applied Crystallography* 43(5), pages 1126–1128. DOI: 10.1107/s0021889810030499.

Wright, R. B. and C. H. Wang (1973). "Density effect on the Fermi resonance in gaseous CO_2 by Raman scattering". *The Journal of Chemical Physics* 58(7), pages 2893–2895. DOI: 10.1063/1.1679594.

Yamamoto, J. and Y. Hagiwara (2024). "Precision evaluation of Raman densimetry for carbon dioxide: improvement by correction for the drift effect". *Applied Optics* 63(5), page 1402. DOI: 10.1364/ao.507939.

Yamamoto, J. and H. Kagi (2006). "Extended Micro-Raman Densimeter for CO_2 Applicable to Mantle-originated Fluid Inclusions". *Chemistry Letters* 35(6), pages 610–611. DOI: 10.1246/cl.2006.610.

– (2008). "Application of densimetry using micro-Raman spectroscopy for CO_2 fluid inclusions: a probe for elastic strengths of mantle minerals". *European Journal of Mineralogy* 20(4), pages 529–535. DOI: 10.1127/0935-1221/2008/0020-1825.

Yuan, X. and R. A. Mayanovic (2017). "An Empirical Study on Raman Peak Fitting and Its Application to Raman Quantitative Research". *Applied Spectroscopy* 71(10), pages 2325–2338. DOI: 10.1177/0003702817721527.