UC Berkeley
UC Berkeley Previously Published Works

Title
PyIRoGlass: An open-source, Bayesian MCMC algorithm for fitting baselines to FTIR spectra of basaltic-andesitic glasses

Permalink
https://escholarship.org/uc/item/9151x8zk

Journal
Volcanica, 7(2)

ISSN
2610-3540

Authors
Shi, Sarah
Towbin, William Henry
Plank, Terry
et al.

Publication Date
2024

DOI
10.30909/vol.07.02.471501

Copyright Information
This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

RESEARCH METHODS VOLCANICA

PyIRoGlass: An open-source, Bayesian MCMC algorithm for fitting baselines to FTIR spectra of basaltic-andesitic glasses

Sarah C. Shi*[α], W. Henry Towbin[α], Terry Plank[α], Anna Barth[α,β], Daniel Rasmussen[α,γ], Yves Moussallam[α], Hyun Joo Lee[α], and William Menke[α]

[α]Lamont-Doherty Earth Observatory, Columbia University. New York, NY USA.
[β]University of California, Berkeley. Berkeley, CA USA.
[γ]National Museum of Natural History, Smithsonian Institution. Washington, DC USA.

Submitted: 2023-10-31 | Accepted: 2024-06-12 | Published: 2024-07-26

Editor: K. Iacovino | Typesetter: J. Farquharson

Abstract

Quantifying volatile concentrations in magmas is critical for understanding magma storage, phase equilibria, and eruption processes. We present PyIRoGlass, an open-source Python package for quantifying concentrations of H_2O and CO_2 species in the transmission FTIR spectra of basaltic to andesitic glasses. We leverage a dataset of natural melt inclusions and back-arc basin basalts with volatiles below detection to delineate the fundamental shape and variability of the baseline underlying the CO_3^{2-} and $H_2O_{m,1635}$ peaks, in the mid-infrared region. All Beer-Lambert Law parameters are examined to quantify associated uncertainties. PyIRoGlass employs Bayesian inference and Markov Chain Monte Carlo sampling to fit all probable baselines and peaks, solving for best-fit parameters and capturing covariance to offer robust uncertainty estimates. Results from PyIRoGlass agree with independent analyses of experimental devolatilized glasses (within 6 %) and interlaboratory standards (10 % for H_2O, 6 % for CO_2). We determine new molar absorptivities for basalts, $\varepsilon_{H_2O_{t,3550}} = 63.03 \pm 4.47$ L/mol · cm and $\varepsilon^{2-}_{CO_{3,1515,1430}} = 303.44 \pm 9.20$ L/mol · cm; we additionally update the composition-dependent parameterizations of molar absorptivities, with their uncertainties, for all H_2O and CO_2 species peaks. The open-source nature of PyIRoGlass ensures its adaptability and evolution as more data become available.

Keywords: Volatiles; FTIR; Open-Source; Python; Bayesian; Markov Chain Monte Carlo.

1 Introduction

The loss of volatiles from magma during exsolution and degassing modulates phase equilibria [Sisson and Grove 1993], crystallinity [Blundy and Cashman 2001], density [Ochs III and Lange 1999], and viscosity [Hess and Dingwell 1996] from magma genesis to eruption. Major volatile components are H_2O and CO_2, both of which have strong pressure-dependent solubilities. Analyses of H_2O and CO_2 in mineral-hosted melt inclusions are often utilized to estimate the conditions of magma storage and evolution, whereas analyses of synthetic glasses inform understandings of solubility and phase equilibria. H_2O and CO_2 in volcanic glasses and synthetic products are commonly measured with two analytical techniques: Fourier Transform Infrared Spectroscopy (FTIR) and Secondary Ion Mass Spectrometry (SIMS). Here we focus on the FTIR spectroscopic technique, with the goal of improving the fitting of baselines and peaks in spectra and quantifying uncertainties in the resulting concentrations.

FTIR spectroscopy utilizes infrared light to excite materials, where infrared absorbance relates to the rotational, vibrational, and electronic energies of molecular bonds. Absorbance is proportional to concentration, allowing for the quantification of species with the Beer-Lambert Law:

$$c = \frac{AM}{\varepsilon l \rho}, \tag{1}$$

where c is concentration, A is absorbance, M is the molar mass of the absorbing volatile species (g·mol^{-1}), ε is the absorptivity of the species (L·mol^{-1}·m), l is the optical path length or thickness (m), and ρ is density (kg·m^{-3}). H_2O and CO_2 species in silicate glasses absorb light in the near-infrared (12,800–4000 cm^{-1}) and mid-infrared (4000–200 cm^{-1}) wavenumber regions (Figure 1). Volatile species vibration is observed through infrared absorbance, with the vibrations occurring as bending or stretching. Dissolution of H_2O in silicate liquids forms both molecular water (H_2O_m) and hydroxyl (OH^-) species with concentration dependence. The vibration of H_2O species combines bending and stretching energies in two dimensions [von Aulock et al. 2014]. Two H_2O_m peaks are found at ~5200 and ~1635 cm^{-1}, and one OH^- peak is found at ~4500 cm^{-1} (Figure 1). Total H_2O (H_2O_t = H_2O_m+OH^-) is represented by the ~3550 cm^{-1} peak. Linear relationships between absorbance and concentration dominate, but non-linear relationships can be introduced when increased sample thicknesses result in insufficient light transmission to the detector, causing the $H_2O_{t,3550}$ peak to saturate with ragged peak tops and absorbances exceeding two [von Aulock et al. 2014; McIntosh et al. 2017]. When the $H_2O_{t,3550}$ peak is saturated, the combination of $H_2O_{m,5200}$ or $H_2O_{m,1635}$ with OH^-_{4500} species must be used instead. Dissolution of CO_2 forms molecular carbon dioxide (CO_2) and carbonate (CO_3^{2-}). The vibration of CO_2 species combines bending and stretching in three dimensions [von Aulock et al. 2014]. The CO_2 peak occurs at ~2350 cm^{-1} and the CO_3^{2-} doublet occurs at ~1515 and ~1430 cm^{-1} (Figure 1). We restrict our treatment to basaltic to andesitic glasses, where the CO_3^{2-} species dominate [Dixon and Pan 1995; Brooker et al. 2001a; b; Lowenstern 2001].

*✉ sarah.shi@columbia.edu

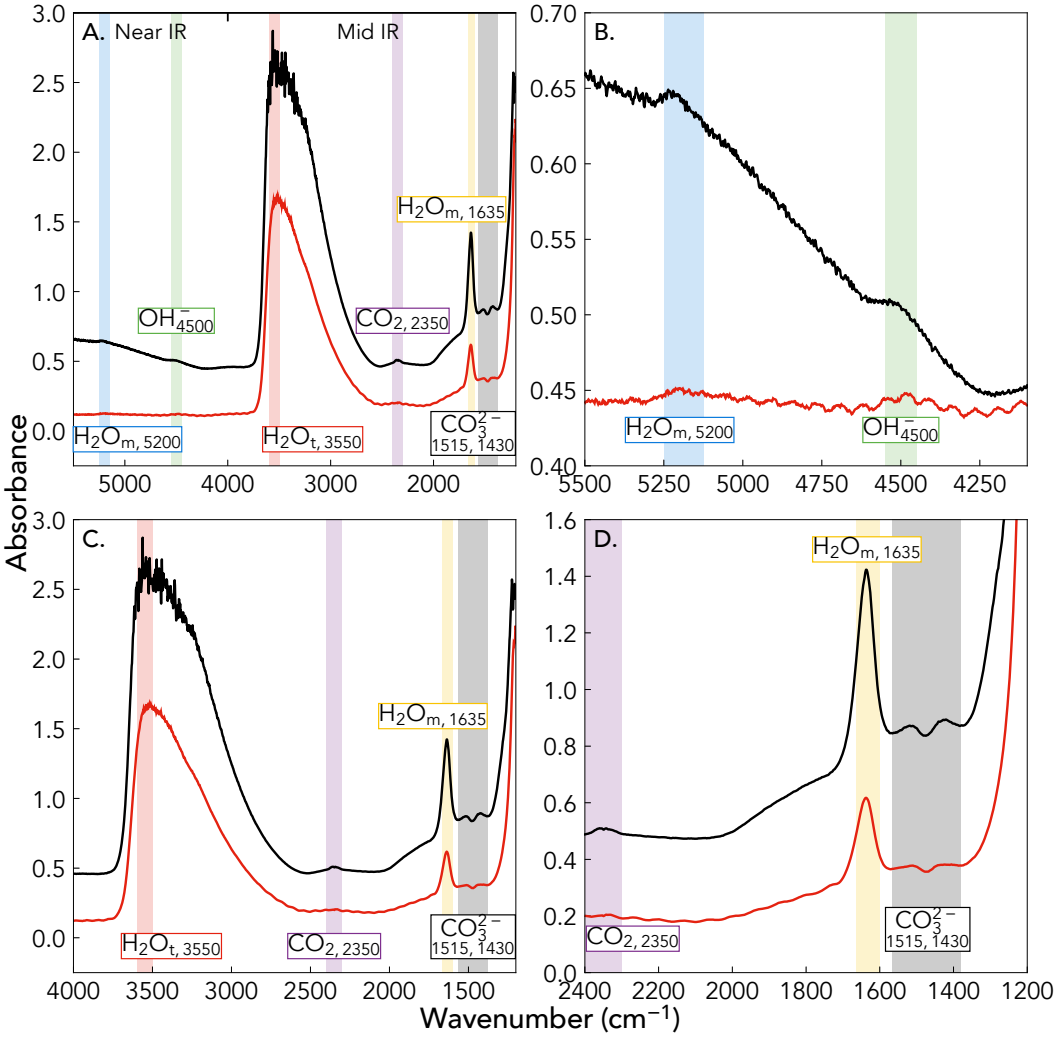

Figure 1: Transmission FTIR spectra of olivine-hosted melt inclusions VF18-AC4-OL49 (black) and VF18-AC4-OL53 (red), from Volcán de Fuego, Guatemala (see Section 3.2). OL49 is 91 ± 3 µm thick and OL53 is 39 ± 3 µm thick. Greater thicknesses increase absorbance and contribute to saturation, seen when the absorbance of $H_2O_{t,3550}$ exceeds two and the peak is ragged. [A] Transmission FTIR spectra through the near-IR and mid-IR regions. [B] $H_2O_{m,5200}$ and OH^-_{4500} peaks in the near-IR region. OL53's absorbance spectrum is shifted upwards for visualization. OL49 has strong signal-to-noise ratios with well-defined peaks, but OL53 has low signal-to-noise ratios with interference fringes. Saturated samples have high signal-to-noise ratio peaks in the near-IR region. OL49 contains 1.82 ± 0.35 wt.% $H_2O_{m,5200}$ and 1.25 ± 0.26 wt.% OH^-_{4500}. OL53's near-IR peaks are not considered given lack of saturation in $H_2O_{t,3550}$. [C] $H_2O_{t,3550}$ peak. OL49 is saturated in $H_2O_{t,3550}$, introducing non-linearity into the Beer-Lambert Law and preventing usage. OL53 contains 4.04 ± 0.43 wt.% $H_2O_{t,3550}$. [D] $H_2O_{m,1635}$ peak and CO_3^{2-} doublet. Rapid absorbance upturn following the $CO^{2-}_{3,1430}$ peak prevents a linear baseline from being fitted, motivating the development of PyIRoGlass. OL49 contains 1.30 ± 0.19 wt.% $H_2O_{m,1635}$ and 753 ± 36 ppm CO_2 and OL53 contains 1.49 ± 0.26 wt.% $H_2O_{m,1635}$ and 735 ± 62 ppm CO_2, with concentrations from PyIRoGlass.

The determination of volatile concentrations in silicate glasses requires the measurement of peak amplitudes or areas, defined as the difference in absorbance or integrated absorbance between peak and baseline. Variability and subjectivity in baseline definition constitutes a significant, and typically unreported, uncertainty in most studies determining volatile concentrations. The $H_2O_{m,5200}$, OH^-_{4500}, and $H_2O_{t,3550}$ peaks in the near-infrared to mid-infrared region can be well quantified with linear to near-linear baselines between the two absorbance minima on either side of the peak. The $H_2O_{m,1635}$ peak and CO_3^{2-} doublet peaks pose greater challenges due to the steep and sharp increase of baseline absorbance at wavenumbers lower than 1430 cm^{-1}. The convolution of the $H_2O_{m,1635}$ peak and the $CO^{2-}_{3,1515}$ doublet peak in glass spectra is a further complication. Devolatilized glass baselines are the gold standard but are not possible to apply in melt inclusion studies, where large pieces of glass are not readily available. These devolatilized glasses are either synthesized from chemistry-matched glass powders or allowed to experimentally lose H_2O and CO_2. Previous studies have approximated baselines beneath the $H_2O_{m,1635}$ peak and the CO_3^{2-} doublet peaks using the spectra of chemistry- and matrix-

matched devolatilized glasses [e.g. Dixon et al. 1988; Newman et al. 2000], splines and flexicurves [e.g. Dixon and Stolper 1995], linear functions [e.g. Behrens et al. 2009; Shishkina et al. 2010; 2014], and other curve functions [e.g. Dixon and Clague 2001]. Such approaches are highly subjective, difficult to reproduce, and seldom reported with uncertainties. The main motivation in our development of `PyIRoGlass` is to provide an open-source code that reproducibly fits transmission FTIR baselines and peaks with quantified uncertainties.

2 Components of `PyIRoGlass`

We examine each free parameter within the Beer-Lambert Law (A, ε, l, ρ) to develop `PyIRoGlass`, a `Python` implementation of a Bayesian algorithm with Markov Chain Monte Carlo (MCMC) sampling for fitting baselines and peaks for all H_2O and CO_2 species in basaltic to andesitic glasses. There are several components to `PyIRoGlass`:

I. Section 2.1 Absorbance (A): We assess the fundamental shape and variability of the infrared baseline using principal component analysis (PCA) applied to a dataset of absorbance spectra for melt inclusions from the Aleutian arc [Rasmussen 2019] and back-arc basin glasses [Pearce et al. 1994] with volatiles below detection, relevant for basaltic to andesitic silicate melts. Application of PCA to this dataset of spectra with volatiles below detection allows for the assessment of the fundamental shape of the baseline and the spectral features contributing to its variability [Carvajal et al. 2016].

II. Section 2.2 Molar Absorptivity (ε): We refit the relationships between molar absorptivity and composition with implicit iterative inversions, updating the parameterizations of Mandeville et al. [2002] and Dixon and Pan [1995] with new data, to derive molar absorptivity with uncertainty. We also calibrate new linear and integrated molar absorptivities for a series of synthetic basalts [Moussallam et al. 2024].

III. Section 2.3 Thickness (l), Density (ρ): To calculate thickness, we develop an automated implementation of the reflectance FTIR interference fringe method [Nishikida et al. 1996; Tamic et al. 2001; Wysoczanski and Tani 2006; Sun et al. 2007]. Because glass density is significantly affected by H_2O concentration, we outline an iterative approach to calculating density and H_2O concentration together.

IV. Section 2.4 `PyIRoGlass` integrates all the above steps by employing a Bayesian algorithm with Markov Chain Monte Carlo sampling to find best-fit spectral baselines and peaks, given inputs of the infrared spectrum, thickness, and composition (to calculate molar absorptivity). The output of `PyIRoGlass` consists of peak amplitudes and concentrations calculated from the Beer-Lambert Law (Equation 1) for each CO_2 and H_2O species peak, with their uncertainties.

2.1 Assessing variability in baselines and peaks

We consider two different approaches to estimating baselines and extracting peak amplitudes. Baselines for the $H_2O_{m,5200}$, OH^-_{4500}, and $H_2O_{t,3550}$ peaks are nominally linear (Figure 1), so we use an asymmetric least squares method to determine the baseline, further discussed in Section *2.1.1*. In the wavenumber region between 2400 and 1250 cm^{-1}, the baseline has a complex shape beneath the $H_2O_{m,1635}$ peak and the $CO^{2-}_{3,1515}$ and CO^{2-}_{1430} doublet peaks. Moreover, the tail on the $H_2O_{m,1635}$ peak can convolute with the $CO^{2-}_{3,1515}$ peak. We thus develop an empirical model for the baseline and peaks in this region, based on a dataset of natural glasses with volatiles below detection, described in Section *2.1.2*.

2.1.1 *Asymmetric least squares baselines for* $H_2O_{m,5200}$, OH^-_{4500}, *and* $H_2O_{t,3550}$ *peaks*

The $H_2O_{m,5200}$, OH^-_{4500}, and $H_2O_{t,3550}$ baselines are fit with the asymmetric least squares (ALS) method [Eilers 2004]. ALS is commonly used as an iterative solver for processing Raman and infrared spectra with interpolated baseline fittings, balancing smoothness and asymmetry while remaining fast [Eilers 2004; Peng et al. 2010; Lee et al. 2017]. Data are preprocessed by first applying a median filter to remove single spike noise. We vary the peak fitting region surrounding each peak and fit three repeat baselines with ALS to assess uncertainty within the peak shape and location. Nominal peak fitting regions are varied by ±200 cm^{-1} for the $H_2O_{t,3550}$ peak and ±25 cm^{-1} for the $H_2O_{m,5200}$ and OH^-_{4500} peaks to fit the range of potential peaks and to assess the uncertainty in peak fitting region. The baselines are subtracted from the peak. The three baseline-subtracted peaks are processed with kriging interpolation to reduce noise and to obtain maximum peak amplitude [Krige 1951]. Kriging interpolation utilizes sample data of limited resolution to predict and interpolate data of higher resolution with estimates of uncertainty in each interpolated value. The mean peak absorbance and standard deviation in absorbance are determined from the three baseline-subtracted peaks. $H_2O_{t,3550}$ saturation, observed when the maximum absorbance of $H_2O_{t,3550}$ exceeds two, prompts the return of an error message and suggests the use of the $H_2O_{m,1635}$ and OH^-_{4500} peaks for determining H_2O_t. The use of the $H_2O_{m,1635}$ over the $H_2O_{m,5200}$ peak is further discussed in Section 3.4.

2.1.2 *Baselines for the* $H_2O_{m,1635}$, $CO^{2-}_{3,1515}$, *and* $CO^{2-}_{3,1430}$ *peaks*

The compositional variability and small size of melt inclusions (often <50 µm in diameter) present challenges for determining FTIR baselines for the $H_2O_{m,1635}$ and CO^{2-}_3 doublet. While the FTIR spectra of experimentally devolatilized glasses are one of the best ways to determine baselines for the complex spectral region underlying the peaks, these glasses unfortunately do not exist for all melt inclusion compositions. The approach we have taken is to use natural melt inclusions and back-arc basin (BAB) glasses, with volatiles below detection, to determine a family of baseline shapes in this spectral region. Melt inclusions with volatiles below detection are typically considered failures in the attempt to characterize the volatile concentrations of magmas, reflecting potential decrepitation [Maclennan 2017], high detection limits (Section 3.6), or simply volatile-poor systems. We make use of a dataset of glasses with volatiles below detection here to define baseline shapes for natural basalts and andesites, with raw spectra shown in Figure 2. The dataset consists of 57 trans-

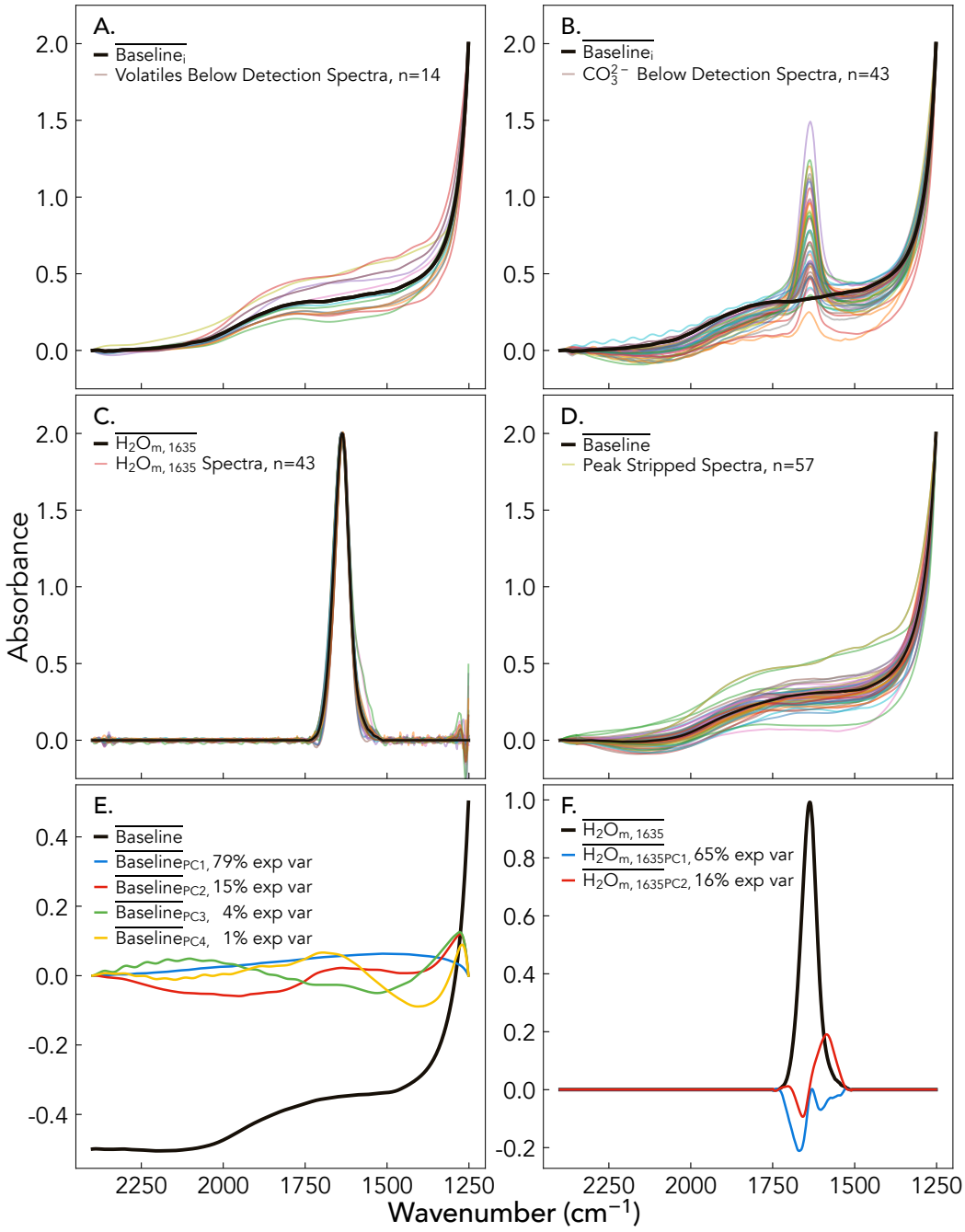

Figure 2: Transmission FTIR spectra for the dataset of basaltic to andesitic melt inclusions and glasses with volatiles below detection used to identify the range of possible baselines for fitting $H_2O_{m,1635}$ and CO_3^{2-} doublet peaks. Spectra range from $16 - 136$ µm and have 1 µm normalized peak amplitudes of less than 10^{-4}. [A] Absorbance spectra in the natural dataset used to generate the initial mean baseline ($\overline{Baseline_i}$). Spectra are linearly rescaled to span an absorbance range of $0 - 2$ in the wavenumber fitting range of 2400 to 1250 cm^{-1}. [B] Absorbance spectra with CO_3^{2-} below detection are fit with $\overline{Baseline_i}$ in a least squares optimization to identify the $H_2O_{m,1635}$ peak shape, allowing for peak-stripping to generate synthetic spectra with volatiles below detection. These spectra are similarly linearly rescaled. [C] $H_2O_{m,1635}$ peaks identified by least squares baseline subtraction are used to generate the mean $H_2O_{m,1635}$ peak shape ($\overline{H_2O_{m,1635}}$) and two associated principal components of $\overline{H_2O_{m,1635}}_{PC1-2}$. [D] The full dataset of spectra with volatiles below detection following peak-stripping of absorbance spectra, using least squares fitting of $\overline{H_2O_{m,1635}}$ and $\overline{H_2O_{m,1635}}_{PC1-2}$. These spectra generate the final mean baseline ($\overline{Baseline}$) and four associated principal components of $\overline{Baseline}_{PC1-4}$ shown in panel [E]. [E] $\overline{Baseline}$ and $\overline{Baseline}_{PC1-4}$ describing variability in the baseline. Panel text shows the percentage of explained variance by that principal component (exp var). [F] $\overline{H_2O_{m,1635}}$ and $\overline{H_2O_{m,1635}}_{PC1-2}$ describing variability in the peak.

mission FTIR spectra of Aleutian melt inclusions [Rasmussen 2019] and back-arc basin glasses [Newman et al. 2000] that span a compositional range from basalt to andesite (SiO_2 = 43–60 wt.%, TiO_2 = 0.7–5.4 wt.%, Al_2O_3 = 10.9–20.8 wt.%, FeO = 6.3–22.2 wt.%, MgO = 2.1–7.0 wt.%, CaO = 4.3–15.6 wt.%, Na_2O = 2.2–4.4 wt.%, K_2O = 0.2–1.9 wt.%). Some spectra lack both detectable $H_2O_{m,1635}$ and CO_3^{2-} peaks (n = 14), while others have $H_2O_{m,1635}$ peaks with no detectable CO_3^{2-} peaks (n = 43).

Definition of baseline and peak shapes is iterative. First, we use the absorbance spectra, with both $H_2O_{m,1635}$ and CO_3^{2-} below detection, to generate an initial mean baseline or $\overline{\text{Baseline}_i}$ (Figure 2A). Spectra are shifted and scaled to ensure equal absorbance in the wavenumber fitting range from 2400 to 1250 cm^{-1}. We describe variability within the baseline with principal component analysis (PCA). Four principal components (PC) describe more than 95 % of the variance within the initial baseline. These spectra are confirmed to have $H_2O_{m,1635}$ or CO_3^{2-} below detection, as the 1 µm normalized peak amplitudes are less than 10^{-4} across thicknesses of 16–136 µm when fit with an initial least squares baseline fitting with $\overline{\text{Baseline}_i}$. The 1 µm normalized peak amplitudes of 10^{-4} establishes an initial operational limit for recognizable peaks, but the detection limits will be quantified later in Section 3.6.

Next, we use the spectra containing $H_2O_{m,1635}$ but with CO_3^{2-} below detection to define peak shapes and to create a greater number of baselines. Spectra with $H_2O_{m,1635}$ but with CO_3^{2-} below detection are first fit for the $H_2O_{m,1635}$ peak, then stripped of it (Figure 2B). We fit the spectra with $\overline{\text{Baseline}_i}$ and four principal components to isolate the shape of the $H_2O_{m,1635}$ peak, which is neither Gaussian nor Lorentzian, due to the rapid downturn of the $H_2O_{m,1635}$ peak. We thus identify the mean $H_2O_{m,1635}$ peak ($\overline{H_2O_{m,1635}}$) and describe variability within the peak with PCA (Figure 2C, 2F). The first two principal components of $\overline{H_2O_{m,1635}}_{PC1-2}$ accommodate variations in $H_2O_{m,1635}$ peak shape and location from its nominal peak location of ~1635 cm^{-1}. $\overline{H_2O_{m,1635}}_{PC1}$ is similar to the negative second derivative of the $H_2O_{m,1635}$ peak; $\overline{H_2O_{m,1635}}_{PC2}$ is similar to the first derivative of the $H_2O_{m,1635}$ peak. We restrict our analysis to the first two principal components, which account for over 80 % of the variance within the peak shape. Subsequent principal components introduce considerable noise and were therefore excluded. Identification of the $H_2O_{m,1635}$ peak shape further allows for $H_2O_{m,1635}$ peak subtraction from the spectra with CO_3^{2-} below detection with least squares, generating additional spectra with $H_2O_{m,1635}$ and CO_3^{2-} below detection. Spectra with CO_3^{2-} below detection are thus converted to spectra with both $H_2O_{m,1635}$ and CO_3^{2-} below detection (Figure 2D).

We iteratively identify and define the mean baseline ($\overline{\text{Baseline}}$) with associated variability described by the addition or subtraction of four principal components of $\overline{\text{Baseline}}_{PC1-4}$ (Figure 2E, Figure 3). For samples that contain both $H_2O_{m,1635}$ and CO_3^{2-}, we disentangle the convoluted tails of the overlapping $H_2O_{m,1635}$ and the $CO_{3,1515}^{2-}$ peaks by identifying $\overline{H_2O_{m,1635}}$ with $\overline{H_2O_{m,1635}}_{PC1-2}$ and the Gaussian CO_3^{2-} peaks (Figure 2F, Figure 4). Given the sharp upturn in absorbance while approaching 1250 cm^{-1}, we provide additional accommodation for shifting and tilting the $H_2O_{m,1635}$ peak and CO_3^{2-} doublet with a linear shift. The code associated with baseline generation is available in the `PyIRoGlass` GitHub repository*.

We can therefore describe the range of possible baselines while simultaneously fitting peaks appropriate for basaltic to andesitic glasses with the net equation:

$$\widehat{A} = \widehat{B} + \widehat{H_2O_{m,1635}} + \widehat{CO_3^{2-}} + \widehat{L}, \tag{2}$$

with $\widehat{A}$ as the best-fit absorbance spectrum.

$\widehat{B}$, the best-fit baseline, is fit by:

$$\widehat{B} = x_0\overline{B} + x_1\overline{B}_{PC1} + x_2\overline{B}_{PC2} + x_3\overline{B}_{PC3} + x_4\overline{B}_{PC4}, \tag{3}$$

where the x_n terms are best-fit scaling coefficients, $\overline{B}$ is the mean baseline, and $\overline{B}_{PCn}$ is the n-th principal component describing variability in the baseline.

$\widehat{H_2O_{m,1635}}$, the best-fit $H_2O_{m,1635}$ peak, is fit by:

$$\widehat{H_2O_{m,1635}} = y_0\overline{H_2O_{m,1635}} + y_1\overline{H_2O_{m,1635}}_{PC1} + y_2\overline{H_2O_{m,1635}}_{PC2}, \tag{4}$$

where the y_n terms are best-fit scaling coefficients, $\overline{H_2O_{m,1635}}$ is the mean $H_2O_{m,1635}$ peak, and $\overline{H_2O_{m,1635}}_{PCn}$ is the n-th principal component describing variability in the $H_2O_{m,1635}$ peak. The $\overline{H_2O_{m,1635}}$ term fits peak amplitude and the $\overline{H_2O_{m,1635}}_{PC1-2}$ terms accommodate lateral variations in peak location.

$\widehat{CO_3^{2-}}$, the best-fit CO_3^{2-} peak, is fit by:

$$\widehat{CO_3^{2-}} = \frac{a}{2\sigma^2} \cdot e^{-(\upsilon^- - \mu)^2}, \tag{5}$$

where a is the peak amplitude, σ is the peak half-width, υ^- is the wavenumber fitting range, and µ is the central wavenumber of the peak. The µ parameter allows for slight deviations (±15 cm^{-1}) in peak location.

$\widehat{L}$, the linear function allowing for tilt and vertical offset of the spectrum, is described by:

$$\widehat{L} = m\upsilon^- + b, \tag{6}$$

where m is the slope, υ^- is the wavenumber fitting range, and b is the intercept.

Peak and baseline coefficients and parameters are sampled and fit concurrently within an MCMC framework with Bayesian inference to accounts for uncertainties and covariance. The best-fit peak amplitudes and uncertainties for the $H_2O_{m,1635}$ and CO_3^{2-} doublet peaks are exported from the algorithm and used to determine volatile concentrations with uncertainties.

*https://github.com/sarahshi/PyIRoGlass/tree/main/Baseline_PCA

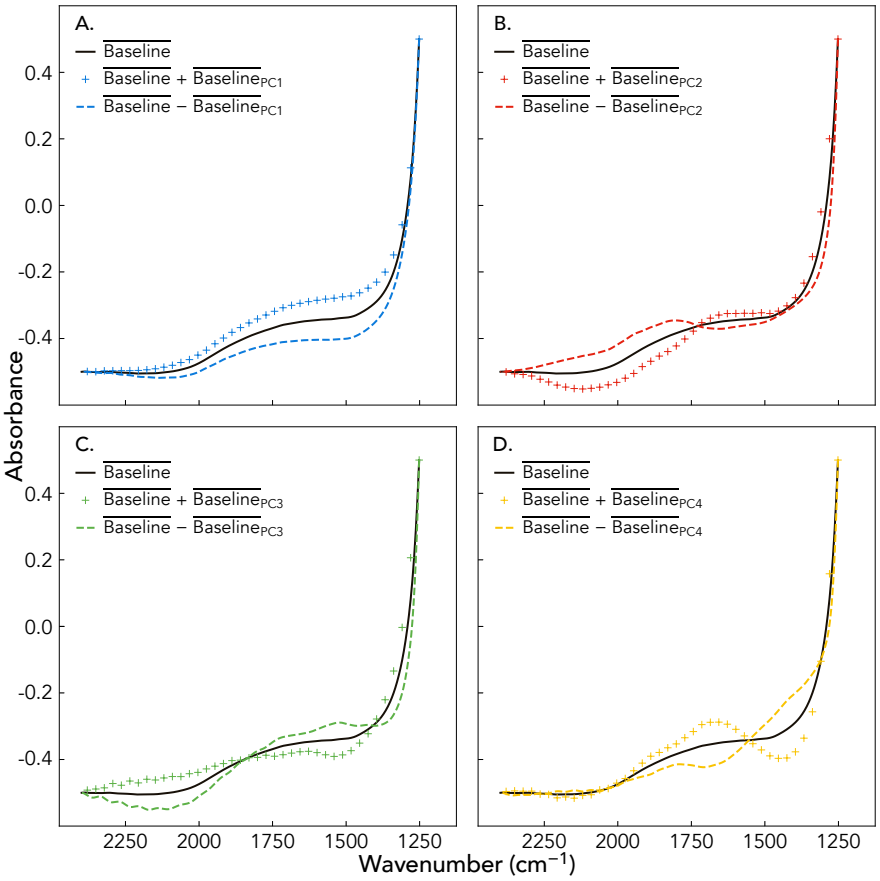

Figure 3: (Left) Mean baseline with volatiles below detection ($\overline{\text{Baseline}}$) with $\overline{\text{Baseline}}_{\text{PC1}-4}$ describing variability within the baseline.
[A] $\overline{\text{Baseline}}\pm\overline{\text{Baseline}}_{\text{PC1}}$ most closely resembles the range of baselines in Figure 2D.
[B] $\overline{\text{Baseline}} \pm \overline{\text{Baseline}}_{\text{PC2}}$.
[C] $\overline{\text{Baseline}} \pm \overline{\text{Baseline}}_{\text{PC3}}$.
[D] $\overline{\text{Baseline}} \pm \overline{\text{Baseline}}_{\text{PC4}}$.

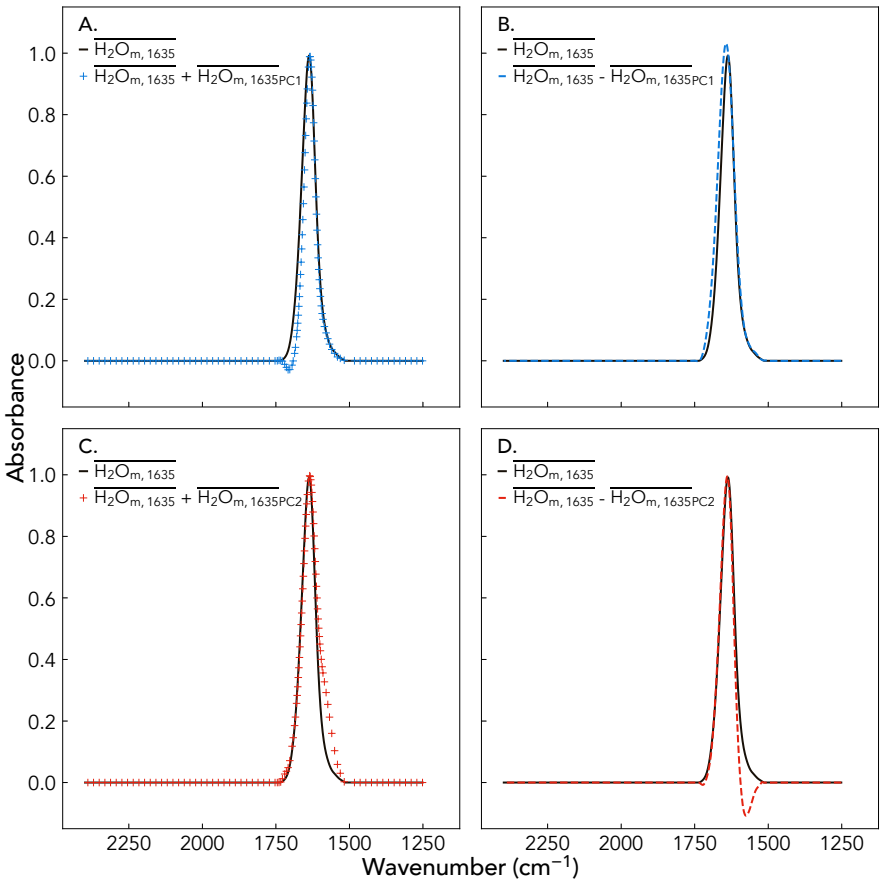

Figure 4: (Right) Mean $H_2O_{m,1635}$ peak ($\overline{H_2O_{m,1635}}$) with $\overline{H_2O_{m,1635}}_{PC1-2}$ describing variability in the nominal peak location.
[A] $\overline{H_2O_{m,1635}}$ + $\overline{H_2O_{m,1635}}_{PC1}$.
[B] $\overline{H_2O_{m,1635}}$ - $\overline{H_2O_{m,1635}}_{PC1}$.
[C] $\overline{H_2O_{m,1635}}$ + $\overline{H_2O_{m,1635}}_{PC2}$.
[D] $\overline{H_2O_{m,1635}}$ - $\overline{H_2O_{m,1635}}_{PC2}$.

2.2 Calibrating molar absorptivity or extinction coefficient

Molar absorptivities or extinction coefficients (ε) determine the intensity at which light attenuates when passing through a material, relating absorbance to concentration. Calibration of ε involves pairing measurements of transmission FTIR absorbance spectra with H_2O or CO_2 concentrations determined by an independent volatile analytical technique. ε is dependent on glass composition. The ε related to H_2O species is dependent on the tetrahedral cation fraction of $\tau = (Si^{4+} + Al^{3+})/\text{total cations}$, while the ε related to CO_2 species is dependent on the cation fraction of $Na^+/(Na^+ + Ca^{2+})$, which we will term η. ε data are drawn from studies with independent methods of volatile measurement, such as Elemental Analyzer (EA), Elastic Recoil Detection Analysis (ERDA), experimental mass balance, Karl Fischer Titration (KFT), manometry, Nuclear Magnetic Resonance (NMR), Nuclear Reaction Analysis (NRA), Secondary Ion Mass Spectrometry (SIMS), and vacuum heating.

We calibrate linear molar absorptivities (peak amplitude-based, ε) and integrated molar absorptivities (peak area-based, ε_i) for the $H_2O_{t,3550}$ and CO_3^{2-} peaks of an array of synthetic basalts (Table 3) analyzed by ERDA and SIMS for H_2O and NRA for CO_2 by Moussallam et al. [2024] (Supplementary Material Section S2.1). We further compile a new dataset of published ε from studies performed between 1982 to 2024 (Table 1). In the full dataset, τ spans the range of 0.50–0.90 and η spans 0.23–0.84. We examine whether these compositional parameters remain the best descriptors of ε. We refit the linear relationships between $\varepsilon H_2O_{m,5200}$, εOH^-_{4500}, $\varepsilon H_2O_{t,3550}$, $\varepsilon H_2O_{m,1635}$, and εCO_3^{2-} and their corresponding compositional parameters using a Newtonian inversion technique. This inclusion of more recent data and broader compositional ranges was done following the last parameterizations of ε with τ by Mandeville et al. [2002] and with η by Dixon and Pan [1995].

Molar absorptivities ε related to H_2O species increases with τ, reflecting increased polymerization of the glass [Stolper 1982]. Tetrahedral cations have also been proposed to compete with non-tetrahedral cations (M = Mg^{2+}, Ca^{2+}, Na^+) to bond with free hydroxyl groups in melt [Pandya et al. 1992; Mercier et al. 2010]. More depolymerized melts will have higher proportions of free hydroxyl, which translates to more M-$(OH)_x$ bonding and hydrogen bonding [Xue 2009; Mercier et al. 2010]. Dixon et al. [1995] and Mandeville et al. [2002] demonstrate the dependence of the $\varepsilon H_2O_{m,5200}$, εOH^-_{4500}, and $\varepsilon H_2O_{m,1635}$ on glass τ. $\varepsilon H_2O_{t,3550}$ is also positively correlated with τ, but this relationship is less straightforward than originally thought [Mercier et al. 2010]. Quench temperatures and rates drive the formation of variable proportions of H_2O_m and OH^- [Silver and Stolper 1989; Stolper 1989]. $\varepsilon H_2O_{t,3550}$ is thus considered as a combination of two endmember molar absorptivities, $\varepsilon H_2O_{m,3550}$ and εOH^-_{3550}, multiplied by the relative proportion of each respective species [Newman et al. 1986; Okumura et al. 2003; McIntosh et al. 2017]. Studies quantifying $\varepsilon H_2O_{m,3550}$ and εOH^-_{3550} are the gold standard, but are currently limited in compositional range. McIntosh et al. [2017] considered the range of τ = 0.746–0.800, spanning silicic compositions with rhyolitic to albitic glasses. More experiments are needed to properly quantify the species-dependent nature of the $H_2O_{t,3550}$ peak. We find a positive relationship between $\varepsilon H_2O_{t,3550}$ and τ, despite complications presented by speciation. While endmember ε data for the $H_2O_{t,3550}$ peak remain limited, $\varepsilon H_2O_{t,3550}$ is best described by τ (Figure 5). We do not observe a strong relationship in the full dataset between ε and NBO/T, calculated by the method outlined in Mysen and Richet [2018].

Dixon and Pan [1995] first demonstrated the dependence of ε for CO_3^{2-} on η. The proportion of Na^+ to Ca^{2+} modulates the dissolution of C as molecular CO_2 and/or CO_3^{2-}. As the proportion of Na^+ increases, εCO_3^{2-} decreases (Figure 5). Carbon dissolves as both molecular CO_2 and CO_3^{2-} in NaAl-rich silicate glasses (albitic, jadeitic, nephelinitic compositions) [Mysen and Virgo 1980a; b; Fine and Stolper 1985; Dixon and Pan 1995], whereas carbon dissolves solely as CO_3^{2-} in Ca and CaMg-rich silicate glasses (diopsidic, sodamelilitic, and akermanitic compositions) [Fine and Stolper 1986; Dixon and Pan 1995]. εCO_3^{2-} data are often limited to either the $\varepsilon CO^{2-}_{3,1515}$ or $\varepsilon CO^{2-}_{3,1430}$ peak, which are also assumed to be approximately equal [Dixon and Pan 1995]. We parameterize a combined εCO_3^{2-} with data from $\varepsilon CO^{2-}_{3,1515}$ and $\varepsilon CO^{2-}_{3,1430}$, where available.

Uncertainties exist within both the compositional parameter (τ, η) and molar absorptivity (ε). Higher uncertainties in compositional parameters may exist in older experimental studies where electron microprobe measurements quantified a limited number of oxide components. Variability in ε can be attributed to differences in the baseline fitting method. $\varepsilon H_2O_{m,5200}$ and εOH^-_{4500} have been quantified with both linear and Gaussian baselines, resulting in differences in the determined ε. In order to account for uncertainties in our regression of ε versus compositional parameter, we apply a variant of the inversion described in Chapter 9 of Menke [2018], with the implicit model:

$$f(x) = -y + m_0 + m_1 p = 0, \tag{7}$$

where m is the vector of coefficients to be solved, p is predictor variable or the compositional parameter, y is the dependent variable or the ε value, and x is a vector containing ε, compositional parameters, and coefficients. The inversion applies Newton's method for determining the minimum of a function in an inverse problem. The implicit inversion derives strength from the ability to incorporate experimental and compositional uncertainties into the calibration and to utilize information on the shape of the error for a trial solution to successively derive improved fits for model parameters, which minimize error. The full development and application of the implicit inversion technique is given in the Supplementary Material and the inversion code is available in the `PyIRoGlass` GitHub repository*.

We derive best-fit parameters and covariance matrices, evolving from the initial least squares fit to reach a minimum on the error surface, and compare these parameters against

*https://github.com/sarahshi/PyIRoGlass/tree/main/Epsilon_Inversion

those from previous work (Figure 5). The problem is sufficiently overdetermined such that the solution does not depend significantly on the starting model and converges within ten iterations. Best-fit compositional parameter multipliers (m_1) decrease and intercepts (m_0) increase for $\varepsilon H_2O_{m,5200}$, εOH^-_{4500}, and $\varepsilon H_2O_{m,1635}$ from the previous parameterizations of Mandeville et al. [2002]; compositional parameter multipliers increase and intercepts decrease for εCO_3^{2-} from Dixon and Pan [1995]. No previous parameterizations of $\varepsilon H_2O_{t,3550}$ exist. Variance within calibrated ε differs significantly from the low variance $\varepsilon H_2O_{m,5200}$ and $\varepsilon H_2O_{m,1635}$ to high variance $\varepsilon H_2O_{t,3550}$. The $\varepsilon H_2O_{m,5200}$ covariance matrix is small given low variance within the data. The $\varepsilon H_2O_{m,1635}$ covariance matrix is similarly small given few calibration data. The εOH^-_{4500} covariance is slightly larger given higher variance within the data. $\varepsilon H_2O_{m,3550}$ variance is greater due to the increased variability in speciation due to quench temperatures and quench rates. εCO_3^{2-} variance is similarly greater, particularly within the $\varepsilon CO^{2-}_{3,1515}$ calibrations, likely given convolution of the $H_2O_{m,1635}$ and the $CO^{2-}_{3,1515}$ tails.

Two forms of error—the error of calibration and the error of a single application of the inversion to molar absorptivity—can be quantified. We define the error of calibration as being equal to the posterior covariance in ε, between the predicted and experimental ε. The 68 % confidence interval of a single application of the model incorporates the calibration uncertainty as well as the analytical uncertainty of the calculated compositional parameter, applied with:

$$c_T = Zc_mZ^T + mc_zm^T, \tag{8}$$

where Z is a vertical matrix of the measured compositional parameter, c_m is the diagonal posterior covariance on the model parameter coefficients, m is the vertical matrix of the posterior model parameter coefficients, and c_z is a diagonal matrix of the uncertainty of measured compositional parameter. We are thus able to derive ε with a 68 % confidence interval to account for both model and analytical uncertainty in a marked improvement to previous calibrations, where there are no estimates of the uncertainty of calibration. Application of the inversion with Equation 8 to the mid-ocean ridge basalt (MORB) composition of Dixon et al. [1995] with τ = 0.627 predicts $\varepsilon H_2O_{m,5200}$ = 0.642±0.284, εOH^-_{4500} = 0.583±0.264, $\varepsilon H_2O_{m,1635}$ = 27.545±8.161, and $\varepsilon H_2O_{t,3550}$ = 60.524±7.088 L/mol·cm. Application of the inversion to the Etna alkali basalt of Shishkina et al. [2014] with η = 0.367 predicts εCO_3^{2-} = 300.433±15.842 L/mol·cm. Uncertainties associated with individual applications of the inversion are larger than those cited within the original studies, but more accurately capture the uncertainty of calibration.

2.3 Calculating thickness and density

We develop an automated implementation for calculating thicknesses from reflectance spectra by the interference fringe method. Interference fringes are produced by the interactions between reflected light and reflections internal to the sample, with fringe wavelengths inversely proportional to thickness and proportional to refractive index of the sample [Nishikida et al. 1996; Tamic et al. 2001; Wysoczanski and Tani 2006; Sun et al. 2007]. Interference fringes from the reflectance spectra of two Volcán de Fuego olivines are shown in Figure 6. Nishikida et al. [1996] define the relationship applicable to glasses and olivine:

$$l = \frac{m}{2n(v_1 - v_2)}, \tag{9}$$

where l is the thickness of analyzed area, m is the number of fringes in the wavenumber range, n is the refractive index of the material, and v_1 and v_2 are the highest and lowest wavenumbers in the interval.

Signal-to-noise ratios associated with interference fringes in reflectance spectra can be low. Reflectance spectra are preprocessed with median filtering to remove potential single spike noise and successively processed with Savitzky-Golay filtering from the `SciPy` package [Virtanen et al. 2020], fitting low-degree polynomials to data with linear least squares, shown in Figure 6 [Savitzky and Golay 1964]. Savitzky-Golay filtering improves signal quality by removing additional noise, but impacts peak amplitudes variably. Sinusoidal peaks with smooth extrema are not strongly impacted by filtering (Figure 6A), but peaks with sharp extrema are reduced in amplitude (Figure 6B). The locations of the peaks in filtered spectra, however, do not differ significantly despite the difference in amplitude. Local extrema associated with each interference fringe wave are located by identifying regions of greatest absorbance change. Maxima are surrounded by points that are lower in absorbance and minima are surrounded by points that are greater in absorbance. Individual local extrema identification by differentiation can fail with noisier spectra, but consideration of all peaks within the wavenumber range allows for uncertainty assessment. Thicknesses are calculated from the wavenumber differences (λ) between all adjacent local extrema in the specified wavenumber range. Uncertainties in thickness are calculated from the standard deviation of all local maxima and minima-derived thicknesses. Individual thicknesses falling outside one standard deviation of the mean are removed.

Thicknesses from interference fringes in the wavenumber range of 2800 to 2000 cm^{-1} were compared against those determined using a Mitutoyo 543-783B Digimatic Indicator for Volcán de Fuego olivines, to assess the uncertainty associated with this automated method. A refractive index of 1.709 was determined using the orientation- and forsterite-dependent parameterization of olivine refractive index from Howie et al. [1992] with a forsterite content of 72 (the mean composition measured next to the studied melt inclusions). Mean thicknesses calculated from interference fringes compare well with those from digital micrometer measurements (Figure 7). Agreement in thickness measurements is demonstrated by the high concordance correlation coefficient (CCC, a measurement of the agreement between two variables [Lin 1989]), low root-mean-square error (RMSE), low relative root-mean-square error (RRMSE), high coefficient of determination (R^2), slope (m) nearing 1, and intercept (b) nearing 0. Mean thickness uncertainties from five repeat measurements are similar

Table 1: ε calibration data for H_2O and CO_2 species, including new linear molar absorptivities (peak amplitude-based) determined in this study for a series of synthetic basalts [Moussallam et al. 2024]. τ values recalculated from reported compositions. Additional information regarding the calculation of calibration uncertainties and the integrated molar absorptivities (peak area-based) are provided in the Supplementary Material and the `PyIRoGlass` GitHub repository. Continued on next page.

Reference	Glass composition	Analytical method	τ	$\varepsilon H_2O_{m,5200}$ (L/mol·cm)	εOH^-_{4500} (L/mol·cm)	$\varepsilon H_2O_{m,3550}$ (L/mol·cm)	$\varepsilon H_2O_{m,1635}$ (L/mol·cm)
Shishkina et al. [2014]	Nephelinite, E2624	KFT, EA	0.508			56.7	
Shishkina et al. [2014]	Basanite, A2549	KFT, EA	0.532			57.3	
Shishkina et al. [2014]	Alkali basalt, B2518	KFT, EA	0.585			60.3	
Shishkina et al. [2014]	Alkali basalt, B2507	KFT, EA	0.585			60.5	
Mercier et al. [2010]	Basanite	KFT	0.588			43.96	
Acocella et al. [1984]	Na_2O*3SiO_2	Experimental loading	0.600	0.634	0.267		
Lesne et al. [2011]	Basalt (Vesuvius)	KFT	0.607	0.57	0.70		
Behrens et al. [2009]	Ultrapotassic melt	KFT, EA	0.622	1.02	0.62	63.9	
Dixon et al. [1995]	MORB	Manometry	0.627	0.62	0.67	63	25
Ohlhorst et al. [2001]	Basalt	KFT	0.629	0.66	0.57		
Lesne et al. [2011]	Basalt (Stromboli)	KFT	0.630	0.62	0.97		
Lesne et al. [2011]	Basalt (Etna)	KFT	0.631	0.56	0.75		
Stolper [1982]	Basalt SLNT	Manometry	0.636			64	
Stolper [1982]	Basalt Bouvet CHN	Manometry	0.639			72	
Pandya et al. [1992]	Basalt	Mass spectrometry	0.643			61	
Jendrzejewski et al. [1997]	MORB	Manometry	0.645			78	
Shi et al. [2024], this study	Basalt	ERDA, SIMS	0.646			63.03	
Vetere et al. [2011]	Shoshonite	KFT	0.659	1.03	0.80		
Di Matteo et al. [2006]	Shoshonite	KFT	0.663			60	
Shishkina et al. [2010]	Basalt	KFT, EA	0.664	0.65	0.69	65	
Shishkina et al. [2010]	Basalt	KFT, EA	0.664	0.65	0.69	65	
Yamashita et al. [1997]	Tholeiite	Manometry	0.665	0.84	0.85	64	
Mercier et al. [2010]	Basalt	KFT	0.666			62.8	
Stolper [1982]	Basalt Marianas 46D-I-1	Manometry	0.668			67	
Yamashita et al. [1997]	High Al Basalt	Manometry	0.671	0.84	0.85	64	
Fiege et al. [2015]	Mg-poor basalt	KFT	0.689	0.848	1.161		
Stolper [1982]	Basalt Tasaday	Manometry	0.691			75	
Jakobsson [1997]	Icelandite	EA	0.708	1.25	0.95		42
Di Matteo et al. [2006]	Latite	KFT	0.712			60	
Carroll and Blank [1997]	Phonolite	Manometry	0.720	1.10	1.25		
Ohlhorst et al. [2001]	Andesite	KFT	0.722	0.90	0.77		
King et al. [2002]	Andesite	Manometry, SIMS	0.743	1.08	1.15	70.3	40.8
Mandeville et al. [2002]	Fe-Bearing Andesite	Manometry	0.746	1.07	0.79	62.32	42.34
Di Matteo et al. [2004]	Trachyte	KFT	0.755	1.36	1.58		
Vetere et al. [2006]	Fe-Free Andesite	KFT	0.756	1.04	0.92		
Silver et al. [1990]	Jadeite Glass	Manometry	0.756	1.13	1.12		
Silver et al. [1990]	CAS-E2	Manometry	0.774	1.07	0.85		
Yamashita et al. [1997]	Dacite	Manometry	0.782	1.60	0.94	68	
Ohlhorst et al. [2001]	Dacite	KFT	0.783	1.17	1.02		
Allabar and Nowak [2020]	Peralkaline rhyolite	KFT	0.791	1.20	1.50		
Mandeville et al. [2002]	Fe-Free Andesite	Manometry	0.795	1.46	0.89	69.21	52.05
Withers and Behrens [1999]	Albite, 300 K	KFT	0.797	1.47	1.27		
Silver et al. [1990]	Orthoclase, KAS	Manometry	0.799	1.87	1.43		
Behrens [1995]	Albite	KFT	0.799	1.49	1.28		
Silver and Stolper [1989]	Albite	Manometry, NMR	0.799	1.67	1.13	70	49
Behrens et al. [1996]	Orthoclase, K100	KFT	0.806	1.65	1.58		
Stolper [1982]	Hydrated obsidian OBS G	Manometry	0.811			55	
Stolper [1982]	Hydrated obsidian OBS I	Manometry	0.813			66	
Stolper [1982]	Hydrated obsidian OBS E	Manometry	0.813			73	
Nowak and Behrens [1995]	Haplogranite	KFT	0.850	1.79	1.56		
Stabile et al. [2020]	Pantellerite	KFT	0.853	1.97	0.99		
Leschik et al. [2004]	Rhyolite	KFT	0.856			80	
Withers and Behrens [1999]	Rhyolite	KFT	0.856	1.69	1.47		
Newman et al. [1986]	Rhyolite	Manometry	0.859	1.61	1.73	100	55
Hauri et al. [2002]	Rhyolite	SIMS	0.859			90	56
Okumura and Nakashima [2005]	Rhyolite	KFT	0.860	1.75	1.42		
Ihinger et al. [1994]	Rhyolite	Manometry	0.860	1.86	1.50	80	55
Aubaud et al. [2009]	Rhyolite	ERDA	0.860			95	
Okumura et al. [2003]	Rhyolite	KFT	0.864			75	
Silver et al. [1990]	Rhyolite	Manometry	0.864	1.61	1.73		
Dobson et al. [1989]	Rhyolite	Manometry	0.867			88	
Stolper [1982]	Synthetic albite A78-56	Manometry	0.902			88	

Table 1 [cont.]: ε calibration data for H_2O and CO_2 species, including new linear molar absorptivities (peak amplitude-based) determined in this study for a series of synthetic basalts [Moussallam et al. 2024]. η values recalculated from reported compositions. Additional information regarding the calculation of calibration uncertainties and the integrated molar absorptivities (peak area-based) are provided in the Supplementary Material and the PyIRoGlass GitHub repository.

Reference	Glass composition	Analytical method	η	$\varepsilon CO_{3,1515}^{2-}$ (L/mol·cm)	$\sigma\varepsilon CO_{3,1515}^{2-}$ (L/mol·cm)	$\varepsilon CO_{3,1430}^{2-}$ (L/mol·cm)	$\sigma\varepsilon CO_{3,1430}^{2-}$ (L/mol·cm)
Behrens et al. [2009]	Ultrapotassic	KFT, EA	0.232	308	110		
Shi et al. [2024], this study	Basalt	NRA	0.236	303.44	9.20	303.44	9.20
Vetere et al. [2014]	Phonotephrite AH1	KFT, EA	0.237			294	16
Jendrzejewski et al. [1997]	MORB	Manometry	0.245	398	10	398	10
Brounce et al. [2021]	Boninite	EA	0.263			265	30
Dixon and Pan [1995]	Basanite	EA	0.268	284	5	281	6
Fine and Stolper [1986]	Basalt	NRA	0.268	375	20	375	20
Shishkina et al. [2014]	MORB, 169oxi	KFT, EA	0.276	359	25	306	11
Shishkina et al. [2010]	Basalt	KFT, EA	0.277	361	43	317	23
Thibault and Holloway [1994]	Leucitite	SIMS	0.288	355		355	
Shishkina et al. [2014]	Ferro-basalt, SC1	KFT, EA	0.303	372	14	329	15
Shishkina et al. [2014]	Basanite, A2549	KFT, EA	0.361	349	25	306	32
Shishkina et al. [2014]	Alkali basalt, OB93	KFT, EA	0.367	385	18	341	19
Shishkina et al. [2014]	Alkali basalt, Etna	KFT, EA	0.367	394	27	360	24
Shishkina et al. [2014]	Alkali basalt, B2507	KFT, EA	0.391	380	30	311	41
Vetere et al. [2014]	Phonotephrite AH2	KFT, EA	0.424			244	15
Vetere et al. [2011]	Shoshonite	KFT, EA	0.438			356	18
Jakobsson [1997]	Icelandite	EA	0.507	180	15	190	15
Vetere et al. [2014]	Phonotephrite AH3	KFT, EA	0.537			264	15
Vetere et al. [2014]	Latite	KFT, EA	0.590			215	8
Duncan and Dasgupta [2015]	Rhyolite	SIMS	0.841	207	22	143	10

to previously reported values of less than 3 µm [Nichols and Wysoczanski 2007; von Aulock et al. 2014; Allison et al. 2019].

Melt density is a function of glass composition and can be calculated from the gram formula weights and partial molar volumes, from Lesher and Spera [2015] or Iacovino and Till [2019], at ambient temperature and pressure of the analytical conditions at 25°C and 1 bar. The two models yield similar densities, differing by a maximum of 2 % for the range of Volcán de Fuego melt inclusions. The large molar volume of H_2O significantly impacts density and requires the implementation of an iterative solver. Density is initially assumed to have no contribution from H_2O but is iteratively updated with the calculated amount of $H_2O_{t,3550}$ or $H_2O_{m,1635} + OH^-_{4500}$ if the sample is saturated (see Section 3.4). The density calculation is repeated to ensure stability and generally converges within ten iterations.

2.4 PyIRoGlass: an MCMC algorithm for calculating volatile concentrations with uncertainties

The net equation describing the absorbance spectrum of a glass or melt inclusion, with υ^- defined as the wavenumber fitting range from 2400 to 1250 cm^{-1}, is expressed with the following form:

$$\begin{aligned}\widehat{A} = {} & x_0\overline{B} + x_1\overline{B}_{PC1} + x_2\overline{B}_{PC2} + x_3\overline{B}_{PC3} + x_4\overline{B}_{PC4} \\ & + y_0\overline{H_2O_{m,1635}} + y_1\overline{H_2O_{m,1635}}_{PC1} + y_2\overline{H_2O_{m,1635}}_{PC2} \\ & + \frac{a_{1515}}{2\sigma_{1515}^2}\cdot e^{-(\upsilon^- - \mu_{1515})^2} + \frac{a_{1430}}{2\sigma_{1430}^2}\cdot e^{-(\upsilon^- - \mu_{1430})^2} \\ & + m\upsilon + b \end{aligned} \quad (10)$$

where x_0, x_1, x_2, x_3, x_4, y_0, y_1, y_2, a_{1515}, a_{1430}, σ_{1515}, σ_{1430}, μ_{1515}, μ_{1430}, m, and b are the best-fit coefficients or parameters (Table 2). x_n and y_n describe best-fit coefficients for the mean baseline or peak and the principal components associated with the baseline or peak. The remaining variables describe best-fit parameters.

Table 2: PyIRoGlass coefficients and parameters in Equation 10.

Abbreviation	Coefficient or Parameter
$\overline{B}$	Mean volatiles below detection baseline
$\overline{B}_{PCn}$	n-th principal component of baseline
x_n	Best-fit coefficient for baseline-related terms
$\overline{H_2O_{m,1635}}$	Mean $H_2O_{m,1635}$ peak
$\overline{H_2O_{m,1635}}_{PCn}$	n-th principal component of $H_2O_{m,1635}$ peak
y_n	Best-fit coefficient for $H_2O_{m,1635}$-related terms
$a_{1515/1430}$	CO_3^{2-} peak amplitude
$\mu_{1515/1430}$	CO_3^{2-} peak center
$\sigma_{1515/1430}$	CO_3^{2-} peak half-width
υ^-	Wavenumber fitting range
m	Slope of linear tilt
b	Intercept of linear tilt

These best-fit coefficients or parameters are first initialized with ordinary least squares. The parameter space of the modeled absorbance of the transmission FTIR spectrum is sampled and explored by leveraging a Markov Chain Monte

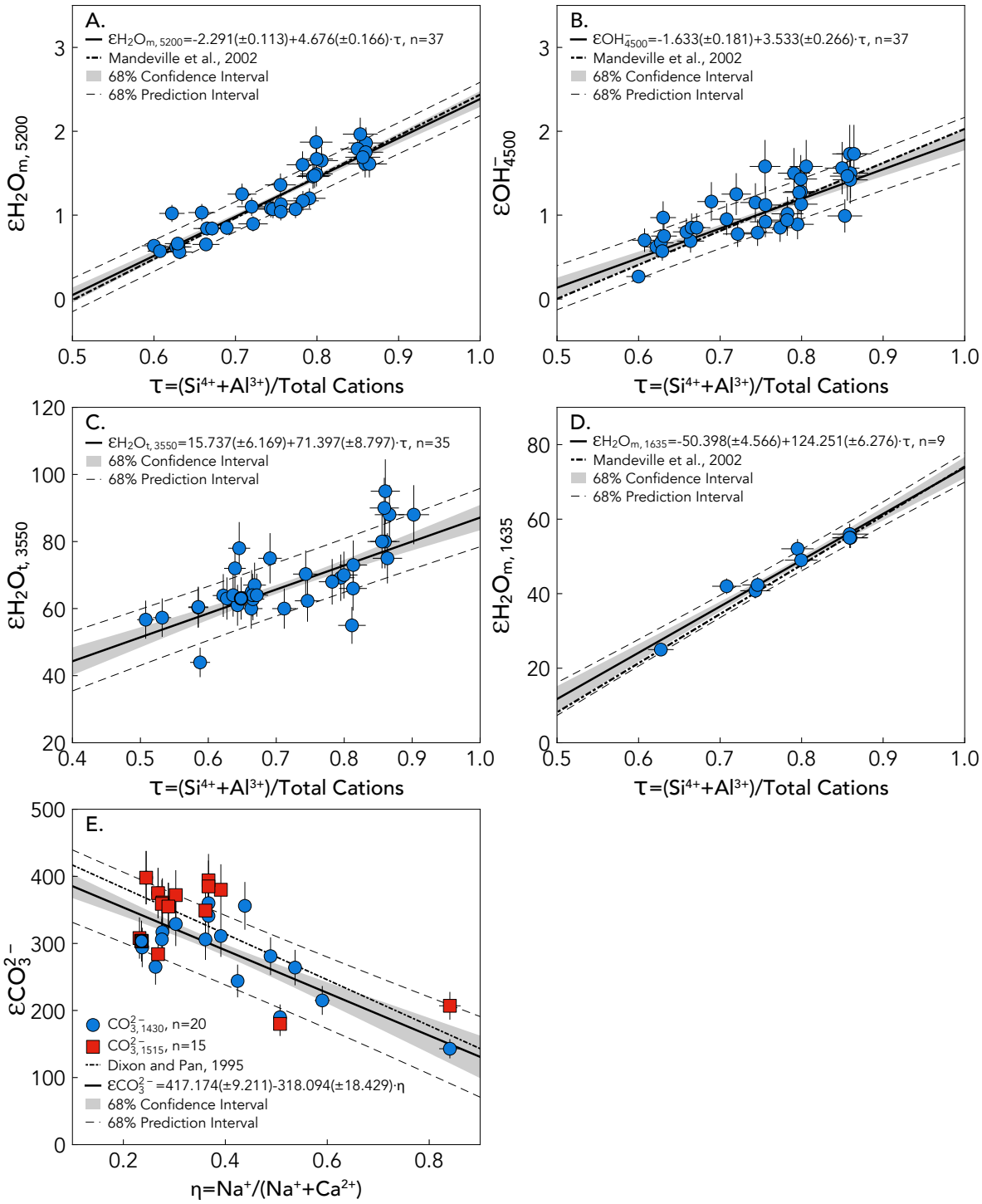

Figure 5: Tetrahedral cation fraction $\tau = (Si^{4+} + Al^{3+})/\text{total cations}$ against: [A] $\varepsilon H_2O_{m,5200}$. [B] εOH^-_{4500}. [C] $\varepsilon H_2O_{t,3550}$, with the new calibrated $\varepsilon_{H_2O_{t,3550}} = 63.03 \pm 4.47$ L/mol · cm demarcated by the bold scatter symbol. Section S2.1 (Table S1 and Figure S1) of the Supplementary Material for additional information on calibration. [D] $\varepsilon H_2O_{m,1635}$. E. Cation fraction $\eta = Na^+/(Na^+ + Ca^{2+})$ against εCO_3^{2-}, with the new calibrated $\varepsilon_{CO^{2-}_{3,1515,1430}} = 303.44 \pm 9.20$ L/mol · cm demarcated by the bold scatter symbol. See Section S2.1 (Table S1 and Figure S1) of the Supplementary Material for additional information on calibration.

Carlo sampler in a Bayesian parametric framework, to account for uncertainty. In this framework of Bayesian inference, the posterior joint probability distribution of model parameters is quantified as a function of the prior probability of model parameters and a likelihood function. Bayes' Theorem dictates that the posterior joint probability distribution is updated as more information becomes available, following Cubillos et al. [2017]:

$$P(\theta|y, M) \propto P(\theta|M)P(y|\theta, M), \tag{11}$$

where y denotes data, θ is the set of parameters, M is the credible region for the set of model parameters, $P(\theta|M)$ is the prior probability distribution (not incorporating new information), $P(y|\theta, M)$ is the likelihood, and $P(\theta|y, M)$ is the posterior probability distribution of parameters. The prior probability distribution of parameters is defined by the expected range and distributions of these parameters. The likelihood serves as the probability density function of the modeled data, given an array of parameters. A Markov Chain Monte Carlo (MCMC) algorithm was utilized to generate random samples from the parameter space that hold a probability density distribution that is proportional to the posterior probability distribution. The MCMC approach utilizes a series of jumps between samples, to effectively traverse and explore parameter space. The open-source `Python` package, `Multi-Core Markov Chain Monte Carlo` (`MC`^{3}), developed by Cubillos et al. [2017], integrates these statistical methods and was implemented for this study.

Fixed absorbance spectra provide the information required to refine posterior probability distributions of model parameters. Prior probability distributions describing the baseline, peaks, and lines can be characterized using either uniform or Gaussian distributions. Uniform priors assign equal probabilities across their defined range, making all values equally probable. In contrast, Gaussian priors assign probabilities based on a Gaussian probability distribution, characterized by its mean and standard deviation. Priors for the parameters relating to the baseline and H_2O_m peak are bounded by the range of variability in principal component space, when the principal component vectors are fit back to PCA space; this is akin to a linear transform from wavenumber to principal component space. These initial priors were limited after testing with the test dataset of Section 3.2 to avoid overfit-

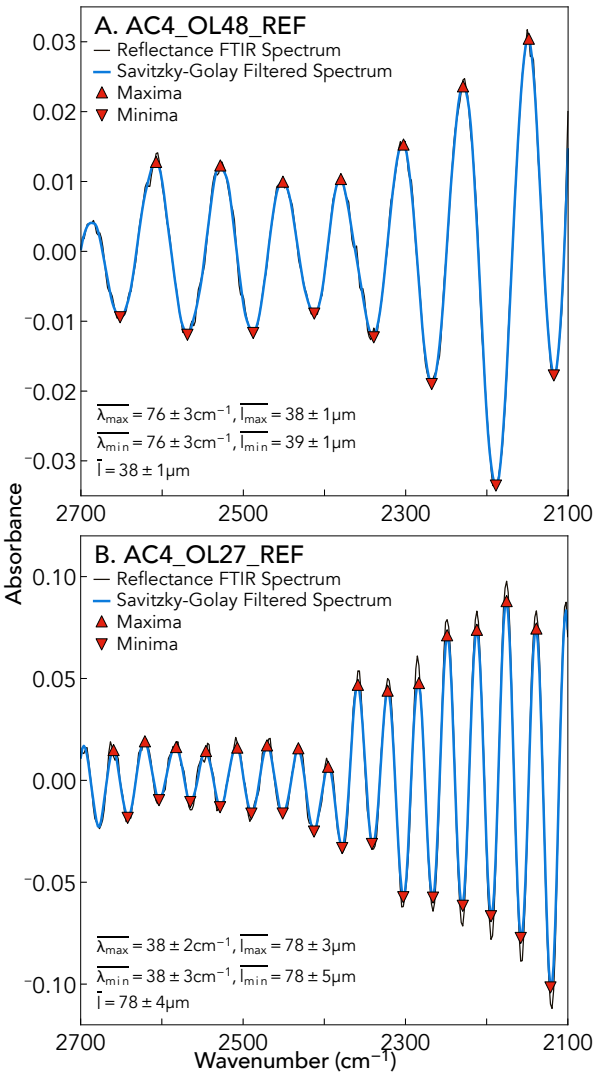

Figure 6: Reflectance FTIR spectra with interference fringes for olivines from Volcán de Fuego, Guatemala. The relationship between wavenumber differences (λ) and wafer thickness is negative, as larger differences are associated with lower thicknesses. [A] Olivine wafer VF18-AC4-OL48 (black) with Savitzky-Golay smoothed spectrum (blue) and identified extrema (red triangles). OL48 spectrum and filtered spectrum are sinusoidal and smooth, with strong agreement between the raw and smoothed spectra. [B] Olivine wafer VF18-AC4-OL27 with initially sharp, singular peaks; Savitzky-Golay smoothed spectrum has reduced amplitudes. Local extrema wavenumbers locations remain well reproduced. Local extrema are more closely spaced than those from OL48, thus OL27 is thicker.

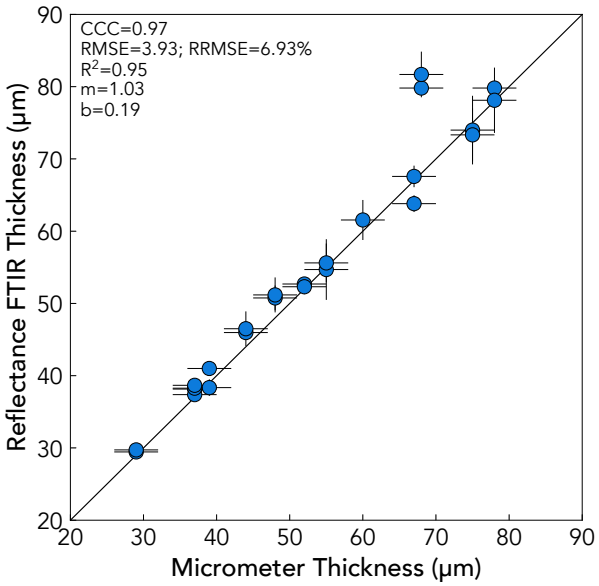

Figure 7: Thicknesses from interference fringes of FTIR spectra compared to those from digital micrometer. Uncertainties on interference fringe thicknesses are calculated from the standard deviation of all maxima- and minima-derived thicknesses within the wavenumber range of 2800 to 2000 cm^{-1}. Uncertainties on thicknesses from the digital micrometer are set at the stated uncertainty of 3 μm. Agreement between calculated and measured thicknesses is indicated by high CCC, low RMSE (μm unit) and RRMSE (unitless), high R^2, m nearing 1, and b nearing 0.

ting. Parameters related to the baseline, H_2O_m peak, CO_3^{2-} doublet peak amplitudes, and linear adjustment are sampled uniformly due to the larger variability in these parameters in the absence of *a priori* knowledge. Priors for the CO_3^{2-} doublet peak locations and half-widths are constrained by prior knowledge about this peak. The CO_3^{2-} doublet peak locations and half-widths are defined with Gaussian priors, reflecting the well-established nature of these parameters in the absorbance spectra of basaltic to andesitic glasses. The initial parameters for these priors are optimized with the least squares Trust Region Reflective algorithm, which is particularly suited for well-constrained, large problems [Branch et al. 1999].

We implement `PyIRoGlass` using 10^6 samples, 9 chains, and $2 \cdot 10^4$ burn-in samples. We employ the Snooker Updater Differential Evolution Markov Chain method (DEMC) for random MCMC sampling [ter Braak 2006; ter Braak and Vrugt 2008]. In the Snooker Updater DEMC algorithm, N chains are computed in parallel. The algorithm computes the differences between the chain from which the jump is being proposed and the remaining chains ($N - 1$), utilizing the differences to determine the scale and orientation of each proposed jump [ter Braak and Vrugt 2008]. The chains converge towards a posterior distribution, ensuring that evolution is oriented and scaled appropriately. We apply a Gelman-Rubin convergence criterion with a threshold of 1.01 to assess convergence. The Gelman-Rubin convergence criterion determines differences both within and across chains. When the variance between chains is minimal, the Gelman-Rubin convergence criterion should approach 1, indicating that the chains have effectively converged to the posterior distribution [Gelman and Rubin 1992]. The MCMC will thus terminate early. The Snooker Updater DEMC optimizes acceptance rates and efficiency. Processing a single transmission spectrum and generating all output files takes approximately 20 seconds on a M2 Max Macbook Pro with 12 CPU cores. The same task takes about two minutes on the cloud-based `Python` development environment of Google `Colab`.

`PyIRoGlass` calculates volatile concentrations with their associated uncertainties using a final MCMC scheme with all parameters in the Beer-Lambert Law. ALS (Section *2.1.1*) and `MC`^{3} baseline-derived peak amplitudes (Section *2.1.2*) are used to determine concentrations of H_2O and CO_2 species, with glass composition used to determine the composition-dependent ε (Section 2.2) and to iteratively determine density (Section 2.3), all as a function of sample thickness (Section 2.3). The best-fit peak amplitudes, molar absorptivities, thicknesses, and densities are applied in the Beer-Lambert Law to determine the concentration of each peak. The uncertainty associated with each concentration is taken by running a simple Monte Carlo error assessment in which all parameters, except molar mass, are allowed to vary with Gaussian distributions for $5 \cdot 10^5$ samples.

3 Application and validation of `PyIRoGlass`

3.1 `PyIRoGlass` structure and packaging

3.1.1 *Installation*

`PyIRoGlass` can be locally installed with with `Python` versions 3.8–3.12 through the command line:

```
pip install PyIRoGlass
```

We further provide a cloud-based `Python` development environment in Google `Colab`, with links provided on the `PyIRoGlass` GitHub repository*. These implementations create interactive computing environments without necessitating local installations of `Python`. In these cloud-based `Python` development environments, installation instead occurs within a code cell:

```
!pip install PyIRoGlass
```

Following installation, the `PyIRoGlass` package must then be imported and initialized with:

```
import PyIRoGlass as pig
```

Functions from `PyIRoGlass` are then called with the abbreviated package name, followed by the function name. The primary functions required for processing transmission FTIR spectra with `PyIRoGlass` include the classes and method functions for loading spectra and/or chemistry and thickness data, and functions calculating baselines and calculating concentrations. The `load_spectrum_directory` method in the `SampleDataLoader` class identifies and loads all FTIR spectra within a directory into a dictionary and is called as follows:

```
loader = pig.SampleDataLoader(args)
loader.load_spectrum_directory()
```

The `load_chemistry_thickness` method in the `SampleDataLoader` class loads chemistry and thickness data and is called as follows:

```
loader = pig.SampleDataLoader(args)
loader.load_chemistry_thickness()
```

Alternatively, both the directory of FTIR spectra and file with chemistry and thickness data can be loaded simultaneously with the `load_all_data` method in the `SampleDataLoader` class as follows:

```
loader = pig.SampleDataLoader(args)
loader.load_all_data()
```

The function then calculating baselines for each FTIR spectrum in the dictionary is called as follows:

```
pig.calculate_baselines(args)
```

and the function calculating concentrations from the peak amplitudes returned from the `calculate_baselines` function is called as follows:

```
pig.calculate_concentrations(args)
```

*https://github.com/sarahshi/PyIRoGlass

All functions require the input of specific variables within the parentheses, called arguments (abbreviated to `args`).

The additional function for processing reflectance FTIR spectra identifies peak and trough locations and determines thicknesses. The `SampleDataLoader` class and associated methods are reused for loading reflectance FTIR spectra. The function determining all thicknesses is called as follows:

```
pig.calculate_mean_thickness(args)
```

The `PyIRoGlass` code is actively maintained on GitHub, where the most recent version and full development history are readily accessible in the repository [Shi 2024]. The current stable `PyIRoGlass` release (Version 0.6.3) with all code, documentation, and data, has been archived on Zenodo with a static DOI (10.5281/zenodo.12735203). All available functions and arguments can be accessed and explored in the `PyIRoGlass` Read the Docs*. The cloud-based `Python` development environments are recommended for ease of use for `Python` beginners, but significant computational gains are provided by running the code locally.

3.1.2 *Data input*

The two required types of input files include FTIR spectrum files and the chemistry and thickness file. Each FTIR spectrum is input as a CSV file, unmodified from the Thermo Scientific Nicolet iN10 MX Fourier Transform Infrared Spectrometer outputs, with a wavenumber and absorbance column. The chemistry and thickness file is structured as a CSV file with columns with headers: Sample, SiO_2, TiO_2, Al_2O_3, Fe_2O_3, FeO, MnO, MgO, CaO, Na_2O, K_2O, P_2O_5, Thickness, and Sigma_Thickness. Sample input files are provided in the `PyIRoGlass` GitHub repository.

3.1.3 *Code input*

Computational workflow examples for `PyIRoGlass` are provided in the Google `Colab` run file example on the `PyIRoGlass` GitHub repository. The workflow is structured such that the sole modifications required are for the paths to input files, shown as follows:

```
import os
spectrum_path = os.path.join(os.getcwd(),
                'Inputs/TransmissionSpectra')
chemistry_thickness_path = os.path.join(
                           os.getcwd(),
                           'Inputs/ChemThick.csv')
export_path = os.path.join(os.getcwd(),
                           'Outputs')
```

These lines of code are used to navigate to the paths of directories and files. In this case, we have a directory called `PyIRoGlass` with the primary code run file (where this code would be found) and a directory called Inputs. The Inputs directory contains a sub-directory of transmission FTIR spectra and a CSV of chemistry and thickness data. The sub-directory is called TransmissionSpectra and contains the transmission FTIR spectra of interest, which can be batched and run at the same time. One additional file containing chemistry and thickness data for all the spectra of interest is called 'ChemThick.csv'. The following functions within `PyIRoGlass` are used to load and read these files as `Python` dictionaries or DataFrames.

```
loader = pig.SampleDataLoader(
        spectrum_path = spectrum_path,
        chemistry_thickness_path =
                chemistry_thickness_path
                        )
dfs_dict, chemistry,
        thickness = loader.load_all_data()
```

The spectra can then be batch processed through `PyIRoGlass` with the saved outputs:

```
peak_amplitude_df,
          failures = pig.calculate_baselines(
                 dfs_dict,
                 export_path)
concentrations = pig.calculate_concentrations(
                 peak_amplitude_df,
                 chemistry,
                 thickness,
                 export_path)
```

The `calculate_baselines` function, fitting baselines to all spectra, takes the arguments of a dictionary of transmission FTIR spectra and a string with the export path and returns the peak_amplitude_df DataFrame. All parameters—most importantly, peak amplitudes with uncertainties—sampled from the `MC`^{3} and ALS fitting of all species peaks are returned within the peak_amplitude_df DataFrame. The `calculate_concentrations` function, calculating concentrations from peak amplitudes, takes the arguments of a DataFrame with peak amplitudes and uncertainties, a DataFrame with sample chemistry, a DataFrame with sample thickness, and a string with the export path. `calculate_concentrations` returns the concentrations DataFrame which contains volatile concentrations with uncertainties, peak saturation information, signal-to-noise-ratios, density, and molar absorptivities. Both DataFrames are saved for future reference.

3.1.4 *Data and file output*

`PyIRoGlass` returns six forms of outputs: DataFrames (.csv) including volatile concentrations with uncertainties, peak amplitudes, densities, and molar absorptivities; figures (.pdf) of peak fits, histograms of explored parameter space, pairwise plots of explored parameter space, and trace plots of explored parameter space; log (.log), text (.txt), zipped (.npz) files recording explored parameter space and statistics; pickle (.pkl) files containing best-fit baselines. We will walk through one instance of `PyIRoGlass` and show the output files in Section 3.3. Additional examples of output files are provided in the `PyIRoGlass` GitHub repository.

3.1.5 *Unit testing*

To ensure the reliability and robustness of the `PyIRoGlass` package for FTIR spectra processing, we implemented a comprehensive suite of unit tests. Unit tests are a foundational component of software development. The primary objective

*pyiroglass.readthedocs.io

of unit tests is to validate that each function in the package works as expected and to detect potential regressions during development as new features are added, modifications are made, or package versions are updated. We develop unit tests for most functions within the code:

1. **Data Loading**: Ingest data to ensure consistent format for package integration.

2. **Fitting Functions**: Validate Markov Chain Monte Carlo outputs of VF18-AC4-OL49 and VF18-AC4-OL53 for peak amplitude consistency up to the third decimal, permitting some variability within the sampling regime. Confirm ALS-derived peak amplitudes for consistency up to the fifth decimal. Ensure creation of output directories and files.

3. **Concentration, Density, Epsilon Calculations**: Cross-check algorithms against select samples with variable glass compositions, for both saturated and unsaturated transmission FTIR spectra. Ensure that volatile concentrations agree up to the fifth decimal place and that Monte Carlo errors agree within 0.01 wt.% for H_2O and 1 ppm for CO_2.

4. **Thickness Calculations**: Assess the agreement in thicknesses derived from interference fringes of reflectance FTIR up to the fifth decimal.

5. **Plotting**: Implement tests to ensure plots are generated accurately and consistently.

6. **Inversions**: Examine Newtonian inversions for molar absorptivity to ensure correct determinations of fitting parameters and uncertainties to the fifth decimal.

We utilize the `pytest` package and framework for executing tests [Krekel et al. 2004]. The choice of `pytest` was motivated by its widespread adoption in the `Python` community, its rich ecosystem, and its flexibility in handling test scenarios. At each instance of modification to the package code, the unit tests are automatically executed within GitHub Workflows. This ensures that the performance and integrity of the software remains intact with every change. Our unit testing currently covers 94 % of the `PyIRoGlass` code. This provides a quantifiable metric of how much of the software's functionality is tested automatically for reproducibility. Through this extensive and tailored test suite, we aim to capture potential anomalies, inaccuracies, or failures in the spectrum processing pipeline, ensuring the reliability of `PyIRoGlass` across its breadth of functionalities.

3.2 Volcán de Fuego melt inclusions: a validation dataset

We demonstrate the performance of `PyIRoGlass` with a validation dataset of melt inclusions from the 2018 eruption of Volcán de Fuego, Guatemala. This exercise underscores the variable quality of output peak information. Samples consist of Aguas Calientes (AC4, IGSN: TAP000081) tephra fall deposits from the largest ash plume of the Volcán de Fuego eruption of June 3, 2018. These samples are the focus of a study currently underway [Shi et al. 2021]. Natural glass standards and olivine-hosted melt inclusions (MIs) were analyzed with the Thermo Scientific Nicolet iN10 MX Fourier Transform Infrared Spectrometer at the Lamont-Doherty Earth Observatory. Dry and CO_2-scrubbed air purged the machine and measurements were made on a liquid nitrogen-cooled MCT-A detector. Our methodology for preparing samples and analyzing melt inclusions by FTIR is further described in Supplementary Material Section S3. Periodic internal check standard analyses of back-arc basin glass D1010 and melt inclusions CN-C-OL1', CN92C-OL2, and ETF46 were performed through each analytical session, as in-house check standards. We found it to be good practice to measure these check standards regularly to ensure data quality.

3.3 Volcán de Fuego melt inclusions: parameter posteriors and spectrum fits

Posterior distributions from one application of `PyIRoGlass` inform our understanding of the relationships between parameters and their underlying uncertainties. Figure 8 shows the parameter posterior probability density distributions, when applied to melt inclusion VF18-AC4-OL49. Strong correlations exist between the posterior probabilities of the baseline-generating coefficients of x_0 scaling $\overline{B}$, x_1 scaling $\overline{B}_{PC1}$, x_2 scaling $\overline{B}_{PC2}$, x_3 scaling $\overline{B}_{PC3}$, x_4 scaling $\overline{B}_{PC4}$, m, and b. This suggests that there is non-uniqueness in the sampled best-fit solutions, resulting in a strong positive covariance structure between x_0, x_1, and x_3, as well as a negative covariance structure between x_0, x_2, and x_4. The posterior probabilities of m and b are strongly negatively correlated, which further highlights the non-unique balance between parameters.

The posterior probabilities of the parameters fitting the CO_3^{2-} doublet reveal positive correlations in the peak-amplitude related, a_{1515} and a_{1430}. This indicates that the CO_3^{2-} doublet peaks should possess similar amplitudes, consistent with the assumptions of Dixon and Pan [1995]. The coefficients, y_n, which fit the $H_2O_{m,1635}$ peak are correlated with the parameters fitting the $CO_{3,1515}^{2-}$ peak, which is expected given the proximity of the peaks and potential convolution of the tails. The peak amplitude modulating coefficient, y_0, for $\overline{H_2O_{m,1635}}$ shows no correlation with the parameters fitting the $CO_{3,1515}^{2-}$ peak. The peak position modulating coefficients, y_1 and y_2, respectively show positive and negative correlations with σ_{1515}, the parameter controlling the half-width of the $CO_{3,1515}^{2-}$ peak. These findings underscore the importance of accounting for covariance between coefficients and parameters in ensemble Bayesian frameworks when fitting baselines and peaks to FTIR spectra, especially for accurately assessing volatile concentration uncertainties.

Figure 9 shows the resulting baseline and peak fits from `PyIRoGlass` for melt inclusion VF18-AC4-OL49, which is also the subject of Figure 1 and of the unit testing described in Section *3.1.5*. The melt inclusion is saturated in $H_2O_{t,3550}$, given the ragged peak and absorbance exceeding two, resulting in the need to combine $H_2O_{m,1635}$ and OH^-_{4500} to quantify H_2O_t (Figure 9A). We sample the uncertainty in near-infrared peak amplitude by fitting three ALS baselines with variable wavenumber fitting ranges to the $H_2O_{m,5200}$ and OH^-_{4500} peaks (Figure 9B). The baseline-subtracted $H_2O_{m,5200}$ and OH^-_{4500} peaks are then interpolated with kriging to reduce

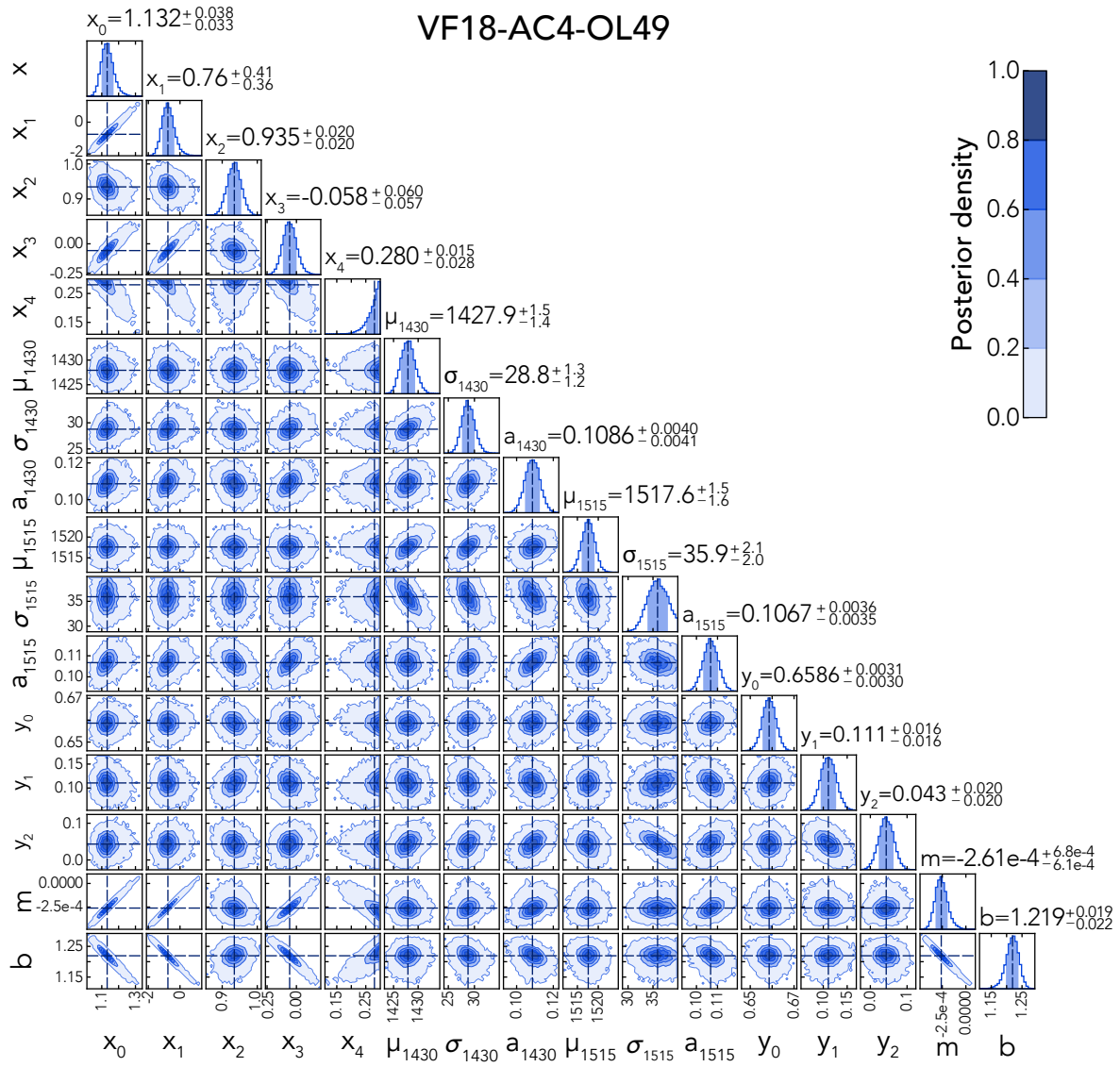

Figure 8: Posterior probability density distribution for the 16 coefficients or parameters of Equation 10 for olivine-hosted melt inclusion VF18-AC4-OL49, fit with PyIRoGlass over 10^6 samples. See Table 2 for additional information about the coefficients and parameters. x_n best-fit coefficients generating the baseline scale the $\overline{\text{Baseline}}$ and $\overline{\text{Baseline}_{\text{PC1-4}}}$ vectors. y_n best-fit coefficients for fitting the $H_2O_{m,1635}$ peak scale the $\overline{H_2O_{m,1635}}$ and $\overline{H_2O_{m,1635\text{PC1-2}}}$ vectors. Parameters for fitting the CO_3^{2-} doublet include the subscripted parameters a, σ, and μ; parameters for fitting the linear shift include m and b. Several parameters covary, demonstrating the importance of assessing underlying covariance structures to properly account for uncertainty within this ensemble solution.

noise within the peaks (Figure 9C), and quantify the maximum peak amplitude [Krige 1951; Cressie 1990]. Variability within the smoothed peak amplitudes is observed, given uncertainty in the fitting ranges for these peaks. Both near-infrared peaks have signal-to-noise ratios exceeding four, representing meaningful signals. The $H_2O_{t,3550}$ peak is similarly fit with three ALS baselines and median filtered to reduce single spike noise (Figure 9D). The peak is not used in this sample given saturation and the introduction of non-linearity to the Beer-Lambert Law. Baseline and peaks, fit simultaneously, for the $H_2O_{m,1635}$ peak and CO_3^{2-} doublet are generated and sampled through 10^6 iterations (Figure 9E). The peak amplitudes and concentrations for both peaks in the CO_3^{2-} doublet are near identical. We report CO_2 as the mean of the CO_3^{2-} doublet peaks. The PyIRoGlass fit to this transmission FTIR spectrum sees residuals in absorbance of less than 0.01 in the region of interest beneath the $H_2O_{m,1635}$ peak and CO_3^{2-} doublet. The combination of the Bayesian framework and fitting technique allows for robust determinations of both volatile concentrations and their uncertainties.

3.4 Observations on H_2O species peak quality

The derivation of concentrations from the $H_2O_{m,1635}$ and $H_2O_{m,5200}$ peaks prompts the question of their relative quality

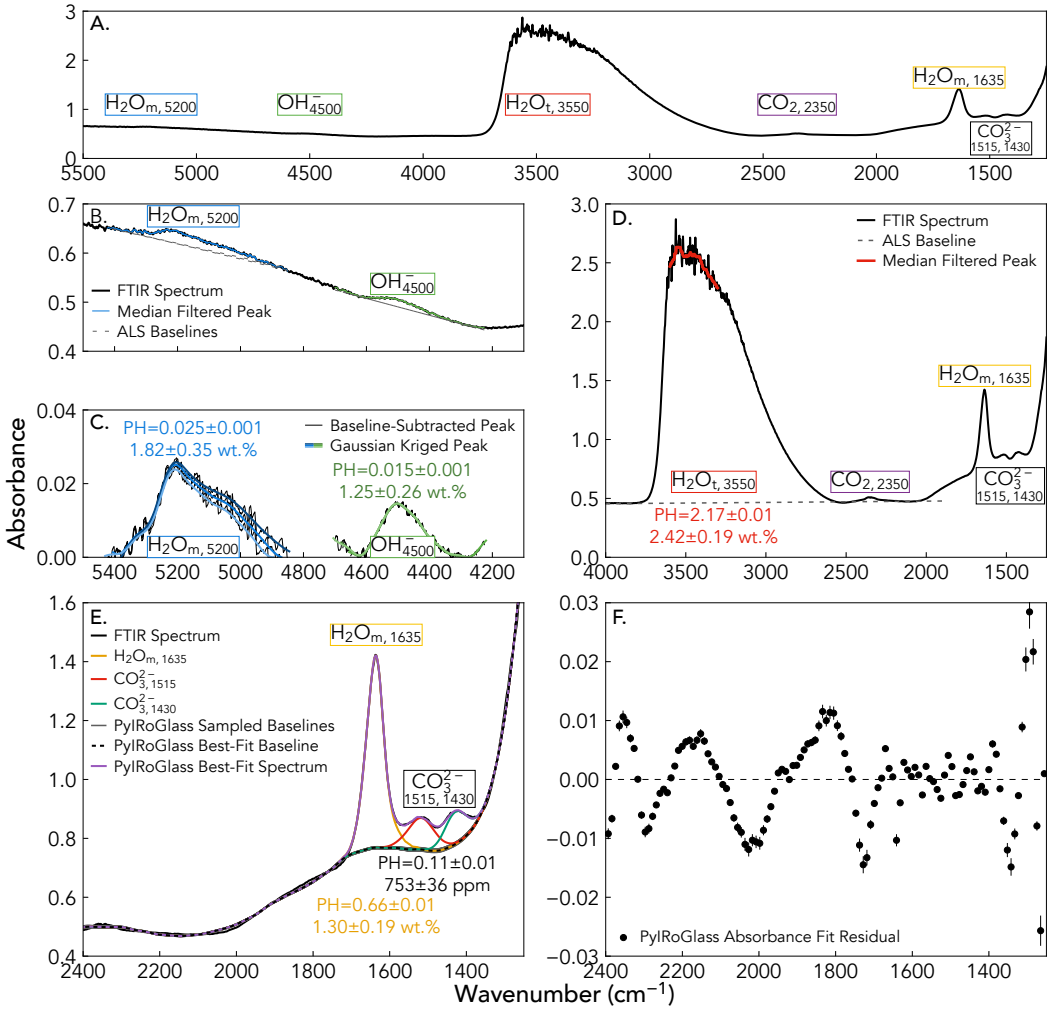

Figure 9: PyIRoGlass baseline and peak fits for the transmission FTIR spectrum of olivine-hosted melt inclusion VF18-AC4-OL49. OL49 has a thickness of 91 ± 3 µm. Uncertainties in peak amplitudes are small within the ensemble solution of PyIRoGlass's fitting framework; most uncertainties in volatile concentrations are introduced by uncertainties in molar absorptivity and thickness. [A] Transmission FTIR spectrum through the near-IR and mid-IR region. OL49's ragged $H_2O_{t,3550}$ peak indicates saturation, prompting the determination of H_2O_t from $H_2O_{m,1635}$ + OH^-_{4500} (see Section 3.4). [B] $H_2O_{m,5200}$ and OH^-_{4500} baselines in the near-IR region. Three ALS-derived baselines are generated by sampling slightly different regions around each peak, to assess uncertainty in peak location. [C] $H_2O_{m,5200}$ and OH^-_{4500} baseline-subtracted peaks, with kriging interpolation. Mean peak amplitudes with uncertainties are determined from the three baseline-subtracted peaks. [D] Saturated $H_2O_{t,3550}$ peak. Three ALS-derived baselines are generated by sampling slightly different regions around each peak, although this saturated peak is not used. [E] Best-fit baselines and peaks for the $H_2O_{m,1635}$ peak and CO_3^{2-} doublet, with baseline fits over 10^6 samples. H_2O_t equals 2.55 ± 0.16 wt.%, combining $H_2O_{m,1635}$ + OH^-_{4500}. [F] Absorbance residual between the best-fit PyIRoGlass spectrum and the measured spectrum. Residuals lie between ±0.01 in the wavenumber region beneath the $H_2O_{m,1635}$ peak and CO_3^{2-} doublet.

and performance. We determine H_2O species concentrations with PyIRoGlass to compare the H_2O_m peaks in the validation dataset of Volcán de Fuego melt inclusions and the internal check standards noted above (Figure 10, Section 3.2). The Volcán de Fuego melt inclusions form the primary basis of this analysis, given the composition- and matrix-matching of the samples. Repeated measurements of the internal check standards allows for assessment of variability in the $H_2O_{m,1635}$ and $H_2O_{m,5200}$ peaks through time. Triplicate analyses of internal check standards were made through each of the 12 analytical sessions, periodically spaced through the session. Both datasets are filtered to only contain analyses where near-infrared peaks have signal-to-noise ratios exceeding four.

$H_2O_{m,5200}$ primarily predicts concentrations greater than those from $H_2O_{m,1635}$ in the Volcán de Fuego melt inclusions (Figure 10A) and internal check standards (Figure 10B, see Section 3.5 and Table 4). Most analyses fall above the 1–1 line and the slopes fit to the datasets exceed 1. Two clusters of internal check standards—back-arc basin glass D1010 ($H_2O_{m,1635}$~0.15 wt.%) and Cerro Negro melt inclusion CN92C-OL2 ($H_2O_{m,1635}$~2.60 wt.%)—highlight the variability observed in the $H_2O_{m,5200}$ peak. Concentrations from

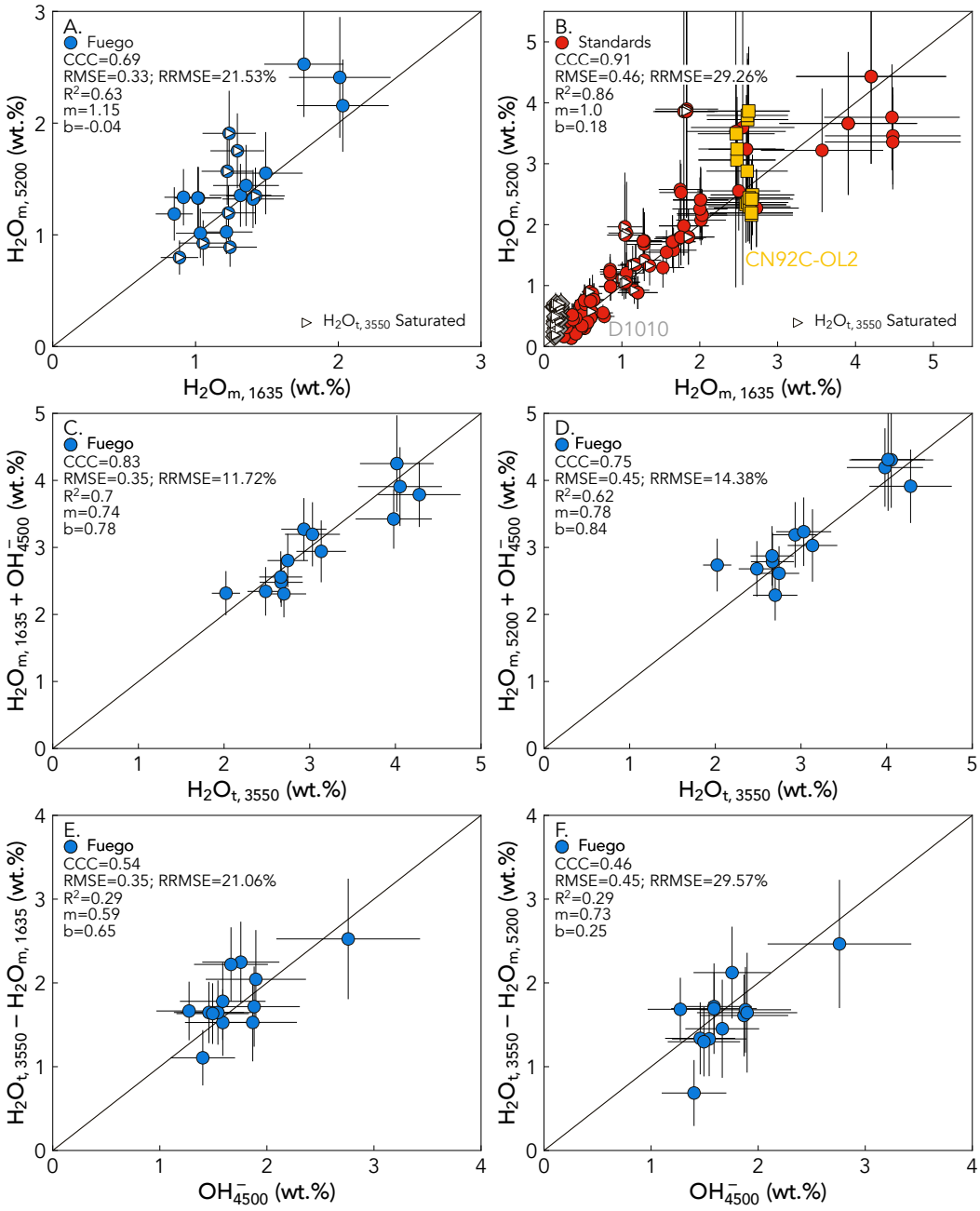

Figure 10: Comparison of the $H_2O_{m,1635}$ and $H_2O_{m,5200}$ peaks. $H_2O_{t,3550}$-saturated glasses are marked with a white triangle. [A] $H_2O_{m,1635}$ and $H_2O_{m,5200}$ concentrations in Volcán de Fuego melt inclusions. Most melt inclusion analyses have higher concentrations of $H_2O_{m,5200}$ than $H_2O_{m,1635}$. [B] $H_2O_{m,1635}$ and $H_2O_{m,5200}$ concentrations in the internal check standards given in Section 3.2. Most analyses have higher concentrations of $H_2O_{m,5200}$ than $H_2O_{m,1635}$. Significant variability is observed for replicate analyses of internal standards D1010 (grey diamonds) and CN92C-OL2 (yellow squares), forming expansive vertical bands. Note that sample D1010 is saturated in $H_2O_{t,3550}$ despite its low $H_2O_{m,1635}$, due to its thickness of 231 µm and concentration of 0.93 ± 0.10 wt.% H_2O_t. [C] H_2O_t from $H_2O_{m,1635} + OH^-_{4500}$. [D] H_2O_t from $H_2O_{m,5200} + OH^-_{4500}$. [E] OH^-_{4500} from $H_2O_{t,3550} - H_2O_{m,1635}$. [F] OH^-_{4500} from $H_2O_{t,3550} - H_2O_{m,5200}$.

the $H_2O_{m,1635}$ peak are tightly constrained, while concentrations from the $H_2O_{m,5200}$ peak vary significantly. Variation within the $H_2O_{m,5200}$ peak concentrations approach 1.00 wt.% for D1010 and span over 2.00 wt.% for CN92C-OL2. Some of the variability is captured within the uncertainty of the peak fits, but the reproducibility of $H_2O_{m,1635}$ concentrations allows for greater confidence to be placed in the peak fits. The variability in the data collected from the near-infrared region can be attributed to the small amplitude of the peaks. The stability of the mid-infrared $H_2O_{m,1635}$ peak is due to its greater amplitude with respect to the baseline, as a more significant peak.

Unsaturated glass analyses further allow for analysis of the quality of near-infrared peaks. Reconstructing H_2O_t from

the combination of either H_2O_m and the OH^-_{4500} concentrations allows for comparison with concentrations determined from the $H_2O_{t,3550}$ peak. This comparison reveals that using the $H_2O_{m,1635}$ peak results in higher concordance correlation coefficients (CCC), a measure of the agreement between two variables [Lin 1989], as well as lower root-mean-square errors (unit-dependent), lower relative root-mean-square errors (unitless), and higher coefficients of determination (Figure 10C, 10D). The concordance correlation coefficient serves as a more comprehensive measure by capturing both the correlation and agreement between the reconstructed concentrations and the reference. While the coefficient of determination provides information on the proportion of variance in the data explained by the prediction model and the root-mean-square error provides insight into the precision of predictions, neither gives a complete picture of how well predictions agree with actual values. The concordance correlation coefficient integrates both these aspects, ensuring that not only are our predictions closely related to actual values (high correlation) but also that they closely match the actual values (high agreement). Hence, in assessing the accuracy of reconstruction, a higher concordance correlation coefficient serves as a more robust indicator of efficacy [Lin 1989]. Reconstructing OH^- from the subtraction of either H_2O_m peak from the $H_2O_{t,3550}$ peak and comparing those concentrations against the OH^-_{4500} peak similarly shows that the $H_2O_{m,1635}$ peak yields higher concordance correlation coefficients, lower root-mean-square errors and relative root-mean-square errors, and higher coefficients of determination (Figure 10E, 10F). We recommend use of the $H_2O_{m,1635}$ peak over the $H_2O_{m,5200}$ peak when calculating H_2O_t for saturated samples, due to the better accuracy, precision, and reproducibility of the $H_2O_{m,1635}$ peak across sessions.

3.5 Synthetic and natural glasses: a test dataset

Two questions remain regarding the performance of `PyIRoGlass`, as we test the performance of the MCMC technique with Bayesian inference. First, are we reproducing the 'true' devolatilized baseline? We answer this question by comparing `PyIRoGlass` baselines to a set of experimentally devolatilized ('true') baselines (Table 3, Section *3.5.1*). Second, are we reproducing absolute concentrations? We answer this question by comparing `PyIRoGlass` results with published values for an array of interlaboratory standards (Table 4, Section *3.5.2*).

3.5.1 *`PyIRoGlass` performance on synthetic glasses with devolatilized baselines*

We assess the performance of `PyIRoGlass` baselines against the 'true' devolatilized baseline by turning to a test dataset consisting of 28 synthetic glasses of basaltic, andesitic, and basanitic compositions from Lee et al. [2024] and Moussallam et al. [2024]. These glasses range in H_2O from 0.2 to 5.85 wt.% and in CO_2 from 0.09 to 1.60 wt.%. For each of these glasses, a corresponding chemistry- and matrix-matched devolatilized glass was available (Table 3). The volatilized and corresponding devolatilized glasses are synthesized from the same starting materials and are compositionally near-identical (agreement within 10 %, with a mean deviation of 4 %). All experimental details can be found in Lee et al. [2024] and Moussallam et al. [2024].

For each of the 28 volatilized glasses, we apply `PyIRoGlass` to the transmission FTIR spectrum to determine the best-fit baseline and peak amplitudes. We then compare these `PyIRoGlass` peak amplitudes with those determined from a simple least squares baseline subtraction of the devolatilized glass spectrum from the volatilized glass spectrum. Replicate analyses of these 28 synthetic volatilized glasses, which have variable thicknesses, form a dataset of 148 spectra. All spectra of volatilized glasses, along with their devolatilized counterparts, are available on GitHub and Zenodo [Shi 2024, DOI: 10.5281/zenodo.12735203].

$CO^{2-}_{3,1515}$ and $CO^{2-}_{3,1430}$ doublet peak amplitudes derived from `PyIRoGlass` and from devolatilized baselines show strong agreement through 50 µm normalized absorbances of 2.25, within 10 % uncertainty bounds (Figure 11A, 11B). Agreement between the samples is demonstrated by the high concordance correlation coefficient, low root-mean-square errors and relative root-mean-square errors, high coefficients of determination, slopes exceeding 0.91, and intercepts approximating 0. Stronger agreement is found in the $CO^{2-}_{3,1430}$ peak amplitudes, indicated by the ratio between peak amplitudes from `PyIRoGlass` and the devolatilized baseline or P/D (Figure 11C, 11D). Scatter (and potentially noise) is evident where peak amplitudes are below 0.50, with P/D ranging from 0.86–1.04. Distinct mean peak amplitude ratios ($\overline{P/D}$) of 0.934 for the $CO^{2-}_{3,1515}$ peak and of 0.950 for the $CO^{2-}_{3,1430}$ emerge. The similarity between baseline fits—and consequently peak amplitudes—and strength of statistical reproduction suggest that `PyIRoGlass` properly derives and constructs baselines to fit basaltic to andesitic glasses. There are some differences between peak amplitudes derived from `PyIRoGlass` and devolatilized baselines, notably a 5.0–6.5 % offset in the two methods (Figure 11C, 11D). A closer examination of the peaks and baselines generated for the test dataset can shed light on the causes of this offset (Figure 12).

Baselines from `PyIRoGlass` and devolatilized glasses can be nearly identical, but also reveal some meaningful differences (Figure 12). Spectra CI-Ref-6 and CI-Ref-10 (Figure 12A, 12B) demonstrate the success of `PyIRoGlass` in reconstructing the devolatilized baseline and in fitting the $H_2O_{m,1635}$ peak and CO^{2-}_3 doublet. `PyIRoGlass` considers the convolution of the $H_2O_{m,1635}$ peak with the $CO^{2-}_{3,1515}$ peak, resulting in the plotted Gaussian $CO^{2-}_{3,1515}$ not reaching the top of the spectrum. On the other hand, the absorbances determined by simple least squares subtraction of the devolatilized baselines do not account for peak convolution and the contribution of the $H_2O_{m,1635}$ peak. This effect is apparent in the slightly lower $\overline{P/D}$ for $CO^{2-}_{3,1515}$ compared to $CO^{2-}_{3,1430}$ (Figure 11C, 11D). The pervasive offset of ~6 % (the uniformly lower $\overline{P/D}$ ratio for both CO^{2-}_3 peaks) requires a different explanation that involves a mismatch in the baseline. Spectra CI-Ref-Bas-9 and CI-Ref-27 (Figure 12C, 12D) both show greater scooping in the devolatilized baselines than predicted by `PyIRoGlass`. Such scooping would lead to greater CO^{2-}_3 absorbances for both

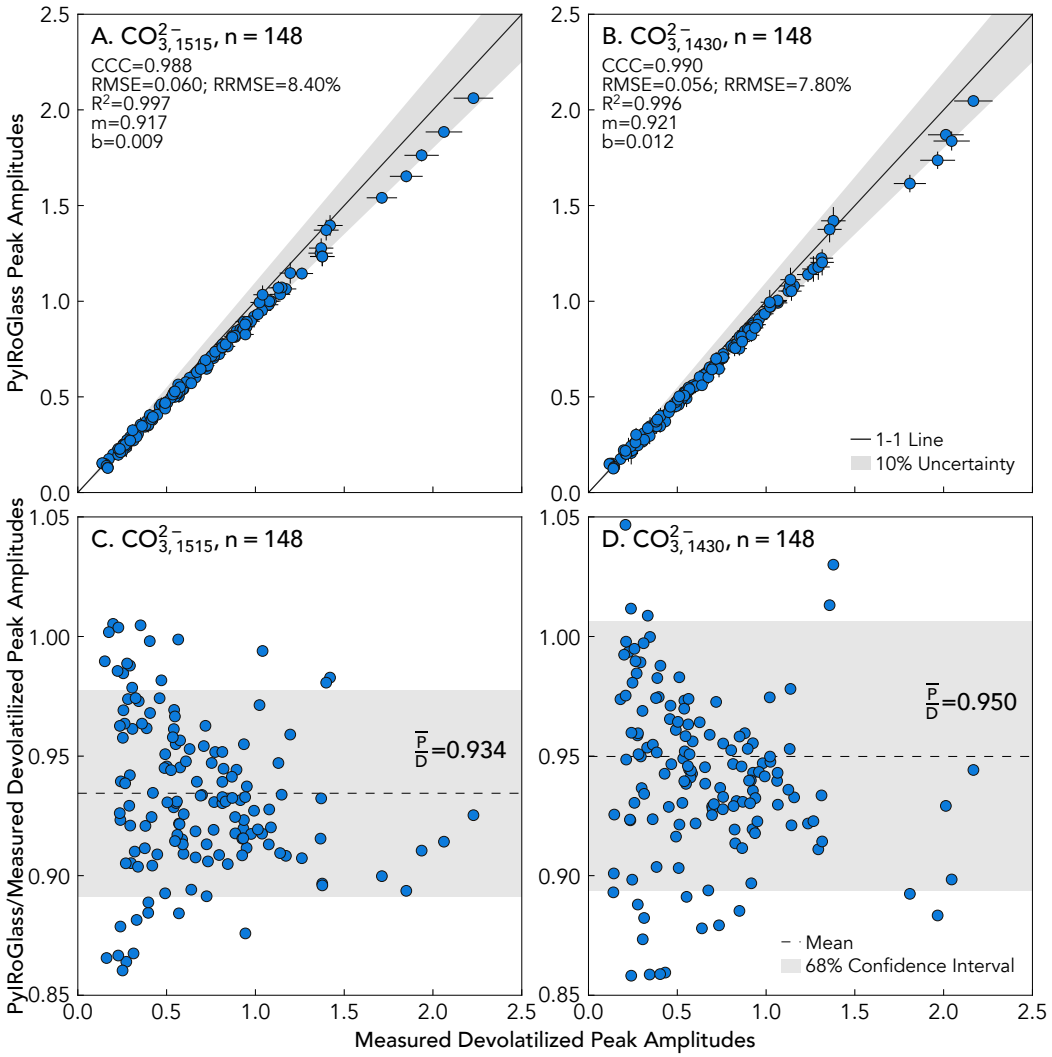

Figure 11: Peak amplitudes between PyIRoGlass and devolatilized baselines for synthesized basalt, andesite, and basanite glasses (Table 3). Peak amplitudes are all normalized to a thickness of 50 μm to demonstrate the significant range of concentrations. 1σ peak amplitude uncertainties from PyIRoGlass are plotted, along with 5 % uncertainty imposed on peak amplitudes from the devolatilized method. [A] Peak amplitudes for $CO_{3,1515}^{2-}$. [B] Peak amplitudes for $CO_{3,1430}^{2-}$. PyIRoGlass peak amplitudes mostly fall within 10 % uncertainty of the devolatilized peak amplitudes, with $CO_{3,1430}^{2-}$ showing stronger agreement. [C] Peak amplitude ratios generated by the two methods (PyIRoGlass/Devolatilized, or P/D) for $CO_{3,1515}^{2-}$, with a mean ratio of 0.934. [D] Peak amplitude ratios for $CO_{3,1430}^{2-}$, with a mean ratio of 0.950. Peak amplitude ratios at lower absorbances incorporate more significant noise.

peaks in the devolatilized baseline approach. While this could be a shortcoming of the baseline dataset input (spectra with volatiles below detection) into PyIRoGlass, it could also be an artifact of the experimental procedure, whereby experimental conditions with no control on oxygen fugacity may have altered the devolatilized glass and therefore the baseline (see Section 3.7). Evidence for this lies in the mismatch between the measured Spectrum CI-Ref-27, and the experimentally devolatilized baseline (Figure 12D), showing a lack of overlap in the region between 1800 and 2300 cm^{-1}. This suggests alterations during the experimental heating that affected the shape of the baseline. The PyIRoGlass ensemble solution, which accounts for a multiplicity of baseline shapes, better matches the measured spectrum for CI-Ref-27. Nonetheless, agreement between baselines from PyIRoGlass and experimental devolatilization within 6 % is still highly satisfactory. And PyIRoGlass can continue to be improved by incorporating baselines from more spectra with volatiles below detection, from a wider array of compositions.

3.5.2 *PyIRoGlass performance on published test dataset*

Next, we compare PyIRoGlass H_2O and CO_2 concentrations calculated from a self-consistent set of FTIR measurements (all at LDEO on the same FTIR, see Table 4 and Supplementary Material for information on analytical conditions). These include previous studies of volatiles in natural back-arc basin glasses measured by FTIR and SIMS [Pearce et al. 1994; Newman et al. 2000; Rasmussen 2019; Barth 2021; Lytle et al.

Table 3: Synthetic volatilized glasses with devolatilized baselines, from Lee et al. [2024] and Moussallam et al. [2024], in the test dataset. Experiments performed by piston cylinder and analyses performed with the Thermo Scientific Nicolet iN10 MX Fourier Transform Infrared Spectrometer, both at the Lamont-Doherty Earth Observatory. 1σ uncertainties propagate analytical, fitting, and replicate analysis uncertainties. Peak amplitudes are normalized to $50\ \mu m$.

Sample	Composition	n	$CO_{2,1515}$ devolatilized peak height	PyIRoGlass $CO_{2,1515}$ peak height	PyIRoGlass $\sigma CO_{2,1515}$ peak height	$CO_{2,1430}$ devolatilized peak height	PyIRoGlass $CO_{2,1430}$ peak height	PyIRoGlass $\sigma CO_{2,1515}$ peak height	PyIRoGlass H_2O (wt.%)	σPyIRoGlass H_2O (wt.%)	PyIRoGlass CO_2 (ppm)	σPyIRoGlass CO_2 (ppm)
CI-LMT-BA3	Synthetic basaltic andesite	1	1.040	1.034	0.012	1.021	0.995	0.016	0.43	0.11	10278	1717
CI-LMT-BA4	Synthetic basaltic andesite	1	1.367	1.251	0.006	1.312	1.225	0.007	0.33	0.04	12553	845
CI-LMT-BA17	Synthetic basaltic andesite	2	2.145	1.973	0.005	2.090	1.958	0.006	0.23	0.05	19023	1778
CI-Ref-4	Synthetic MORB	2	1.007	0.924	0.003	0.992	0.938	0.003	1.11	0.11	8876	542
CI-Ref-6	Synthetic MORB	8	0.553	0.514	0.003	0.544	0.514	0.004	1.02	0.10	4880	294
CI-Ref-8	Synthetic MORB	5	0.930	0.858	0.004	0.913	0.862	0.005	0.88	0.10	8924	640
CI-Ref-9	Synthetic MORB	2	0.909	0.836	0.002	0.883	0.832	0.002	1.61	0.13	8862	422
CI-Ref-10	Synthetic MORB	8	0.626	0.574	0.003	0.635	0.597	0.003	0.80	0.07	5439	866
CI-Ref-11	Synthetic MORB	4	0.891	0.815	0.007	0.872	0.813	0.010	1.27	0.49	7688	2295
CI-Ref-14	Synthetic MORB	2	0.349	0.318	0.007	0.330	0.307	0.009	1.07	0.15	2942	468
CI-Ref-15	Synthetic MORB	3	0.275	0.250	0.013	0.247	0.232	0.016	1.00	1.11	2315	542
CI-Ref-18	Synthetic MORB	5	0.303	0.286	0.006	0.293	0.280	0.008	1.01	0.14	2686	378
CI-Ref-20	Synthetic MORB	5	0.310	0.296	0.005	0.299	0.288	0.007	0.78	0.10	2755	380
CI-Ref-23	Synthetic MORB	4	0.245	0.228	0.003	0.240	0.227	0.004	1.01	0.09	2178	180
CI-Ref-25	Synthetic MORB	10	0.993	0.916	0.004	0.982	0.925	0.005	0.86	0.19	7968	1947
CI-Ref-27	Synthetic MORB	6	0.581	0.546	0.006	0.570	0.543	0.008	1.14	0.19	5163	818
CI-Ref-28	Synthetic MORB	6	0.950	0.883	0.006	0.939	0.884	0.007	1.34	0.16	8433	1084
CI-Ref-bas-3	Synthetic basanite	7	0.682	0.643	0.007	0.675	0.626	0.009	1.51	0.25	6784	952
CI-Ref-bas-5	Synthetic basanite	3	0.222	0.222	0.006	0.225	0.218	0.007	1.62	0.19	2353	639
CI-Ref-bas-6	Synthetic basanite	3	0.937	0.861	0.007	0.930	0.849	0.009	1.69	0.23	9368	1497
CI-Ref-bas-7	Synthetic basanite	9	0.440	0.412	0.009	0.433	0.404	0.011	1.38	0.23	4397	1410
CI-Ref-bas-8	Synthetic basanite	8	0.483	0.456	0.007	0.486	0.450	0.009	1.29	0.21	4719	1235
CI-Ref-bas-9	Synthetic basanite	14	0.651	0.609	0.007	0.655	0.599	0.009	1.30	0.56	7429	4277
ND70-2-01	Synthetic basalt	6	0.155	0.143	0.003	0.133	0.140	0.003	2.12	0.34	1304	122
ND70-3-01	Synthetic basalt	7	0.259	0.248	0.005	0.229	0.237	0.005	3.43	0.97	2263	409
ND70-4-02	Synthetic basalt	8	0.474	0.446	0.003	0.440	0.427	0.004	3.86	0.89	4161	631
ND70-5-02	Synthetic basalt	6	1.356	1.277	0.013	1.292	1.243	0.017	5.16	0.59	12057	1221
ND70-6-02	Synthetic basalt	3	1.833	1.652	0.008	1.940	1.730	0.012	5.85	0.96	16018	1864

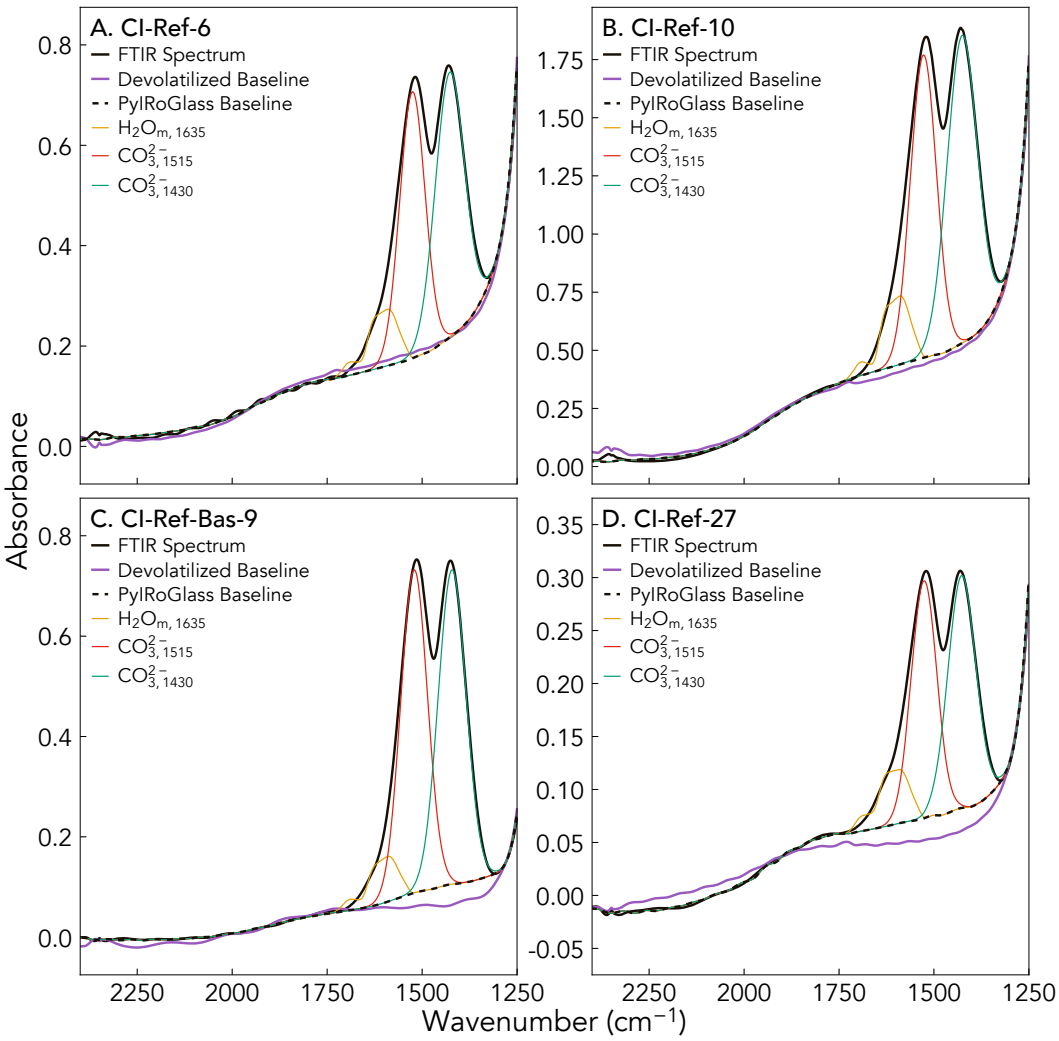

Figure 12: Comparison of PyIRoGlass and devolatilized baselines for four exemplar spectra. Spectra are presented as analyzed, without thickness normalization. [A] CI-Ref-6 with strong baseline agreement, [B] CI-Ref-10 with strong baseline agreement. [C] CI-Ref-Bas-9 with the devolatilized baseline accommodating greater scooping beneath CO_3^{2-} peaks. [D] CI-Ref-27 with the devolatilized baseline accommodating greater scooping beneath CO_3^{2-} peaks, but with divergence of baselines past $1800\ cm^{-1}$.

2023] or synthetic glasses by FTIR, SIMS, KFT, EA, ERDA, or NRA [Shishkina et al. 2010; Helo et al. 2011; Fiege et al. 2015; Brounce et al. 2021; Moussallam et al. 2024]. Natural glasses include back-arc basin basalts D1010, ALV1833-11, WOK5-4, and ALV1846-9 analyzed with FTIR by Newman et al. [2000]; CD33-12-2-2 and CD33-22-1-1 analyzed with SIMS by Lytle et al. [2023]; ETFSR-OL8 analyzed with SIMS by Barth [2021]. Back-arc basin glasses analyzed with FTIR utilize baselines from a least squares fitting technique, using a single reference baseline and peak components, described by Newman et al. [2000]. Synthetic glasses include the high-Mg basaltic glasses ABWCl-F0x and ABWB-0x analyzed with KFT for H_2O by Fiege et al. [2015]; boninites BF73, BF76, and BF77 analyzed with KFT for H_2O and elemental analyzer (EA) for CO_2 by Brounce et al. [2021]; basalts M35 and M43 analyzed with KFT for H_2O and FTIR for CO_2 by Shishkina et al. [2010]; synthetic basalts of the ND70-series analyzed with SIMS and ERDA for H_2O and NRA for CO_2 by Moussallam et al. [2024]; synthetic basalt NS-1 analyzed with FTIR for H_2O and CO_2 by Helo et al. [2011]; synthetic basaltic andesite INSOL-MX1-BA4 analyzed with FTIR for H_2O and CO_2 by Moussallam et al. [2024]. Utilizing PyIRoGlass to determine volatile concentrations from the re-analyzed transmission spectra of samples, with published concentrations from alternate analytical techniques, allows for the validation of concentration measurements. Glasses analyzed with SIMS, KFT, EA, ERDA, and NRA allow for comparison against more independent analytical techniques, although many calibrations rely on standards analyzed by FTIR.

Application of PyIRoGlass to the test dataset shows that H_2O determined by FTIR primarily agree with published values up to concentrations of approximately 6.0 wt.%, even when the $H_2O_{t,3550}$ peak is saturated and we calculate H_2O_t from the $H_2O_{m,1635}$ and OH^-_{4500} peaks (Figure 13A). Agreement between concentrations is demonstrated by the high concordance correlation coefficient, low root-mean-square error and relative root-mean-square error of ~10 %, high coefficient of determination, slope nearing 1, and intercept nearing

0. The calculated H_2O_t from glasses synthesized by Fiege et al. [2015] are well represented by the combined $H_2O_{m,1635}$ and OH^-_{4500} concentrations despite saturation, discussed in Section 3.4. Offsets observed in back-arc basin glasses D1010, ALV1833-11, WOK5-4, and ALV1846-9 can likely be attributed to differences in fitting methodologies, fitting regions, and potential heterogeneity in analyzed glass wafers. Inversion and published molar absorptivities (63 L·mol^{-1}·cm) agree within 1 L·mol^{-1}·cm. The remaining samples within the test dataset agree within the uncertainties of analysis at concentrations below 6.0 wt.% H_2O, across FTIR, SIMS, KFT, and ERDA.

`PyIRoGlass` demonstrates strong agreement in CO_2 determined by FTIR, SIMS, EA, and NRA through concentrations of 1.6 wt.% for back-arc basin and synthetic glasses (Figure 13B). The low H_2O and low CO_2 back-arc basin basalts, along with the synthetic boninites (BF73, BF76, BF77) and synthetic basalts (NS-1, ND70-series), provide coverage over a significant range of volatile concentrations. For glasses with published concentrations using FTIR, we use the molar absorptivities provided in each respective study to calculate concentrations. This is for the sake of consistency and to highlight the baseline fitting in `PyIRoGlass`, which is our primary objective. Overall, strong agreement in CO_2 concentrations between `PyIRoGlass` and the other published results (Figure 13B) is demonstrated by the high concordance correlation coefficient, low root-mean-square error and relative root-mean-square error of ~6 %, high coefficient of determination, slope nearing 1, and intercept nearing 0. If we used the compositionally-dependent molar absorptivities in Figure 5, results would be ~30 % lower for the Brounce et al. [2021] BF glasses and ~12 % higher for the Helo et al. [2011] NS-1 glass. Clearly, more work resolving the compositional effects on the molar absorptivity of the CO_2 peaks is welcome.

3.6 Detection limit for volatile species

Previous determinations of FTIR detection limits for glasses have primarily relied on either estimations of operational limits (converted to concentration) for recognizable peaks [Fine and Stolper 1985; 1986; Newman et al. 1986; Dixon and Stolper 1995; Newman et al. 2000] or utilized the 'practical' quantification of the lowest concentrations of an analyte in a given sample [Mercier et al. 2010; Lowenstern and Pitcher 2013]. The detection limit, or LOD, is defined as the minimum concentration of an analyte that is distinguishable from a background or blank, and is quantified by calculating three times the standard deviation of the background signal [Analytical Methods Committee 1987]. Quantification of the detection limit occurs within the signal domain (absorbance for FTIR) and is reported within the concentration domain, with the Beer-Lambert Law serving as the calibration function [Delgado 2022]. Concentrations and thickness are correlated by the Beer-Lambert Law. Thicker samples yield stronger signals in absorbance, but risk saturation when this signal is too strong and insufficient light is transmitted to the detector. Achieving the balance between signal intensity and noise is critical to high-quality analyses with low detection limits [Newman et al. 1986].

We assess the detection limits for each volatile peak within the dataset of synthetic and natural glasses (Section 3.5), where concentrations are well determined. The baseline is considered as the background, as it captures the spectrum of a sample without volatiles. To determine the variability in the background signal, we calculate three times the standard deviation of the `PyIRoGlass` sampled baselines. For the 10^6 sampled baselines underlying the $H_2O_{m,1635}$ and CO_3^{2-} peaks, we first remove burn-in samples and then utilize the best-fit peak locations from `PyIRoGlass` to calculate the standard deviation in the sampled baselines (see Figure 9E). This captures the background variability within the baseline. For the three repeat baselines underlying the $H_2O_{m,5200}$, OH^-_{4500}, and $H_2O_{t,3550}$ peaks, we directly calculate the standard deviation in the background with the nominal peak locations. We consider only samples with signal-to-noise ratios exceeding 4 for the $H_2O_{m,5200}$ and OH^-_{4500} peaks, and unsaturated samples for the $H_2O_{t,3550}$ peak. Three times these standard deviations in the baseline absorbances were converted to concentrations with the Beer-Lambert Law, incorporating the ε, l, and ρ values obtained from `PyIRoGlass` as scaling factors.

Detection limits directly correlate with measured sample thickness, and this relationship can be defined with a reciprocal function (Figure 14). The shape of the reciprocal function is rooted in the presence of thickness in the denominator of the Beer-Lambert Law, with the detection limit increasing rapidly as thickness approaches zero. Despite the inextricable nature of thickness within the concentration space, this practice provides a useful quantification of the detection limit and further demonstrates the differential background and detector sensitivity across the analyzed spectral region. The reciprocal function holds the form:

$$\mathrm{LOD} = \frac{a}{l} + b, \tag{12}$$

where l is the wafer thickness, and a and b are fit with least squares optimization from the `scikit-learn` package [Pedregosa et al. 2011].

Detection limits for the $H_2O_{m,1635}$ and CO_3^{2-} peaks, with `PyIRoGlass` MCMC sampled baselines, are well fit by the reciprocal function (Figure 14A, 14B, 14C). The reciprocal functions fitting these peaks are associated with high coefficients of determination (greater than 0.9) and low relative root-mean-square errors (less than 20 %). The strength of the fit can be attributed to the lower noise within this region and the increased constraints associated with the volatiles below detection baselines. At 50 µm, the detection limit for $H_2O_{m,1635}$ is predicted to be 0.02 wt.% and for CO_3^{2-} is predicted to be 47 ppm; at 100 µm, these respectively decrease to 0.01 wt.% and 22 ppm.

The detection limits for the $H_2O_{m,5200}$, OH^-_{4500}, and $H_2O_{t,3550}$ peaks, with `PyIRoGlass` ALS sampled baselines, are associated with significantly more noise within the reciprocal function (Figure 14D, 14E, 14F). The reciprocal functions fitting these peaks have low coefficients of determination (between 0.1 and 0.2) and higher relative root-mean-square errors (greater than 75 %). The weaker nature of this fit can be attributed to increased noise in the spectral region. The lower

Table 4: Back-arc basin glasses (BAB), melt inclusions (MI), and synthetic glasses forming the test dataset, with stated sources—both in terms of publication and analysis location. All glasses were distributed to and re-analyzed by FTIR at Lamont-Doherty Earth Observatory (LDEO), with the exception of the ALV, WOK, and CD series glasses re-analyzed by FTIR at the National Museum of Natural History, Smithsonian Institution. Published analytical technique and locations—FTIR at California Institute of Technology (Caltech) fits baselines with the Newman et al. [2000] least squares fitting technique with a single baseline and $H_2O_{m,1635}$ peak component; FTIR at Institut für Mineralogie, Leibniz Universität Hannover (Hannover) fits tangential baselines; FTIR at American Museum of Natural History (AMNH) and LDEO fits a devolatilized baseline; SIMS analyses performed at Carnegie Institution for Science (CIS) and Centre de Recherches Pétrographiques et Géochimiques, Université de Lorraine (Nancy); ERDA and NRA performed at the Laboratoire d'Etude des Eléments Légers (LEEL) joint CEA-CNRS laboratory at Université Paris-Saclay. Volatile concentrations and uncertainties are presented as provided within the reference within the published concentration column. 1σ uncertainties propagate analytical, fitting, and replicate analysis uncertainties. 10 % uncertainty is imposed when uncertainties are not reported, or where published uncertainties are unrealistic given inherent uncertainties (thickness and molar absorptivity in FTIR). - indicates that there is no published measurement or uncertainty, and bdl indicates that concentrations were below detection limits.

Sample	Reference	Composition	Published analytical technique	n	Published		PyIRoGlass		Published		PyIRoGlass		PyIRoGlass	
					H_2O	σH_2O	H_2O	σH_2O	CO_2	σCO_2	CO_2	σCO_2	CO_2	σCO_2
					(wt.%)		(wt.%)		(ppm)		(ppm)		(ppm) with published ε	
ABWCl-F0x	Fiege et al. [2015]	Synthetic high-Mg basalt	FTIR-Hannover	3	2.97	0.29	3.19	0.48	-	-	435	66	-	-
ABWB-0x	Fiege et al. [2015]	Synthetic high-Mg basalt	KFT-Hannover	4	4.40	0.40	3.88	0.49	-	-	1400	47	-	-
ALV1833-11	Newman et al. [2000]	BAB-Mariana Trough	FTIR-Caltech	3	1.20	0.12	1.02	0.08	102	10	35	9	30	8
21ALV1846-9	Newman et al. [2000]	BAB-Mariana Trough	FTIR-Caltech	3	1.89	0.19	1.42	0.12	bdl	bdl	bdl	bdl	bdl	bdl
23WOK5-4	Newman et al. [2000]	BAB-Mariana Trough	FTIR-Caltech	3	1.60	0.16	1.01	0.08	64	6	21	6	18	5
BF73	Brounce et al. [2021]	Synthetic boninite	KFT/EA/FTIR-Hannover	3	0.72	0.07	0.82	0.06	2995	190	2417	77	3042	97
BF76	Brounce et al. [2021]	Synthetic boninite	KFT/EA/FTIR-Hannover	3	0.67	0.07	0.75	0.06	2336	127	1843	78	2319	98
BF77	Brounce et al. [2021]	Synthetic boninite	KFT/EA/FTIR-Hannover	3	0.70	0.07	0.86	0.08	1030	27	708	57	891	72
CD33-12-2-2	Pearce et al. [1994]; Lytle et al. [2023]	BAB-Lau Basin	SIMS-CIS	3	0.27	0.03	0.27	0.02	170	17	131	14	-	-
CD33-22-1-1	Pearce et al. [1994]; Lytle et al. [2023]	BAB-Lau Basin	SIMS-CIS	3	0.49	0.05	0.53	0.05	109	11	67	16	-	-
D1010	Newman et al. [2000]	BAB-Mariana Trough	FTIR-Caltech	34	1.13	0.07	0.93	0.15	139	14	179	41	149	34
ETFSR-OL8	Barth [2021]	MI-Etna	SIMS-Nancy	3	4.16	0.42	3.85	0.43	-	-	5311	344	-	-
M35	Shishkina et al. [2010]	Synthetic basalt	KFT/FTIR-Hannover	3	4.20	0.12	4.10	0.51	1019	81	975	106	1000	109
M43	Shishkina et al. [2010]	Synthetic basalt	KFT/FTIR-Hannover	3	2.62	0.11	2.52	0.32	3172	265	2735	253	2857	264
ND70-2-01	Moussallam et al. [2024]	Synthetic basalt	ERDA/NRA-CEA-CNRS	6	2.53	0.24	2.12	0.34	1837	35	1304	122	1304	122
ND70-3-01	Moussallam et al. [2024]	Synthetic basalt	ERDA/NRA-CEA-CNRS	7	3.13	0.30	3.43	0.97	2689	54	2263	409	2263	409
ND70-4-02	Moussallam et al. [2024]	Synthetic basalt	ERDA/NRA-CEA-CNRS	8	3.68	0.35	3.86	0.89	4122	65	4161	631	4161	631
ND70-5-02	Moussallam et al. [2024]	Synthetic basalt	ERDA/NRA-CEA-CNRS	6	5.34	0.51	5.16	0.60	12682	105	12057	1221	12057	1221
ND70-6-02	Moussallam et al. [2024]	Synthetic basalt	ERDA/NRA-CEA-CNRS	3	6.26	0.59	5.85	0.96	16847	120	16018	1864	16018	1864
NS-1	Helo et al. [2011]	Synthetic basalt	FTIR-AMNH	3	0.37	0.04	0.35	0.03	3154	158	4273	173	3546	144
INSOL-MX1-BA4	Moussallam et al. [2024]	Synthetic basaltic andesite	FTIR-LDEO	3	0.15	0.01	0.15	0.01	8207	377	8044	367	8044	367

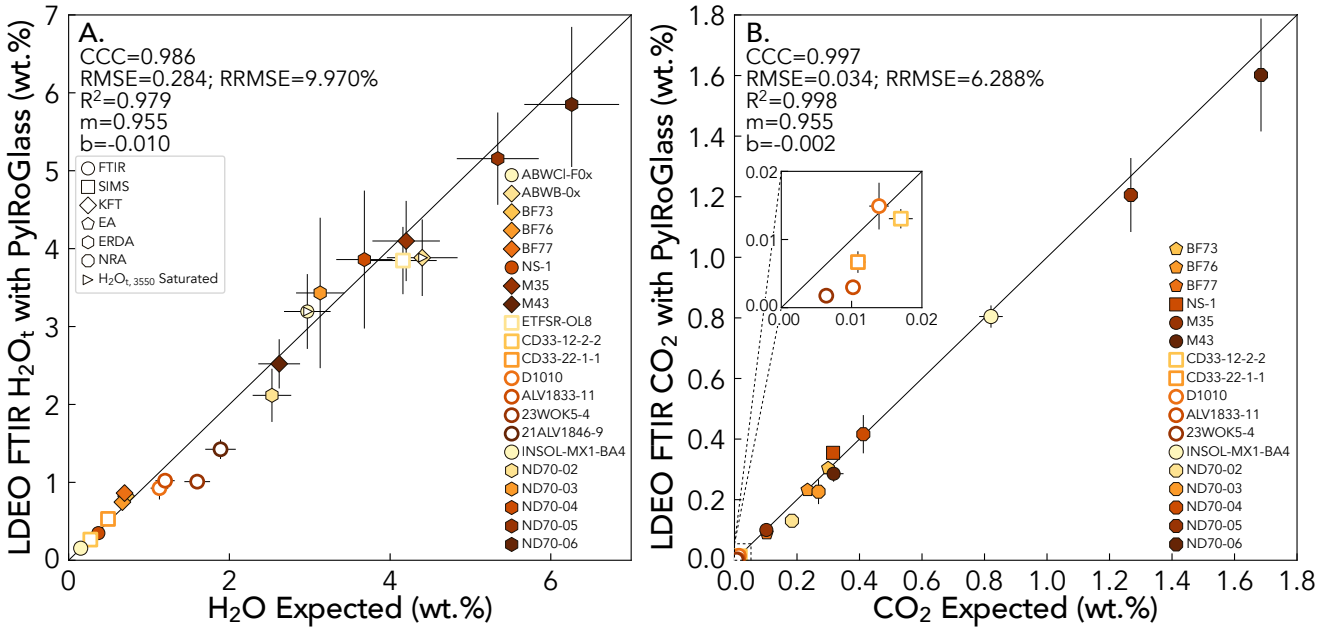

Figure 13: Comparison of volatile concentrations from `PyIRoGlass` and concentrations from published analyses by FTIR, SIMS, KFT, EA, ERDA, or NRA. 1σ uncertainties from `PyIRoGlass` propagate analytical, fitting, and replicate analysis uncertainties. [A] H_2O concentrations from FTIR vs. concentrations from FTIR, SIMS, KFT, or ERDA. Saturated samples H_2O concentrations are represented by combining $H_2O_{m,1635}$ and OH^-_{4500}. Agreement within H_2O concentrations is demonstrated by the high CCC, low RMSE (wt.% unit), low RRMSE of $\sim$ 10 % (unitless), high R^2, m nearing 1, and b nearing 0. [B] CO_2 concentrations by FTIR vs. concentrations from FTIR, SIMS, EA, or NRA. We use the molar absorptivities provided in each respective study to calculate CO_2 concentrations, for samples with published concentrations from FTIR (presented in Table 4). Agreement within CO_2 concentrations is demonstrated by the high CCC, low RMSE (wt.% unit), low RRMSE of $\sim$ 6 % (unitless), high R^2, m nearing 1, and b nearing 0.

amplitude near-infrared $H_2O_{m,5200}$ and OH^-_{4500} peaks are associated with increased variability and noise; this translates to the detection limits associated with these peaks. The detection limits for $H_2O_{m,5200}$ are approximately three times those for the $H_2O_{m,1635}$ peak. This observation further lends credence to the observation of the higher quality of the $H_2O_{m,1635}$ peak over the $H_2O_{m,5200}$ peak (see Section 3.4). The detection limits for the mid-infrared $H_2O_{t,3550}$ peak are low, but are similarly associated with variability and noise. At 50 µm, the detection limit for $H_2O_{m,5200}$ is predicted to be 0.06 wt.%, for OH^-_{4500} is predicted to be 0.05 wt.%, and for $H_2O_{t,3550}$ is predicted to be 0.01 wt.%; at 100 µm, these will respectively decrease to 0.05 wt.%, 0.03 wt.%, and 0.01 wt.%.

We thus quantify the detection limit, or minimum detectable signal, across a range of sample thicknesses extending to ~250 µm. The entire uncertainty associated with FTIR may however be more significant based on the quality of the analyzed spectrum, which is influenced by spectral scans, spectral resolution, and aperture size. The spectra utilized to quantify detection limits were all acquired at 256 collection scans at 4 cm^{-1} spectral resolution with the Thermo Scientific Nicolet iN10 MX FTIR Spectrometer with a liquid nitrogen-cooled MCT-A detector. Aperture sizes were maximized for each sample, with areas ranging from 150-10^4 μm^2. Spectra of lower quality or with more significant noise likely will see different detection limits. Future work may seek to further quantify the impacts of spectrum quality or aperture size on detection limits.

3.7 Chemistry Modulating the Baseline

Here, we explore the relationships between glass chemistry and the baseline in our validation and test datasets. Correlation coefficients between chemical components and the best-fit baseline-generating coefficients from `PyIRoGlass` are low, but these weak correlations may still provide some information. The coefficients impacting the baseline most significantly modify melt polymerization and structure—including the network-forming tetrahedrally coordinated cations in SiO_2, Al_2O_3, and minor components Fe_2O_3 and TiO_2; the network-modifying alkali earth cations in CaO and MgO; silicate network-depolymerizing H_2O (Figure 15). Notably, the commonly used parameter of non-bridging oxygens to tetrahedral cations (NBO/T) does not strongly correlate with any baseline-generating coefficient, while the individual network-forming, network-modifying, and network-depolymerizing components do [Mysen and Richet 2018; Mysen 2022].

Increasing the proportion of network-forming cations in SiO_2 and Al_2O_3, sitting near the center of oxygen tetrahedra, appears to increase x_0 or the best-fit coefficient altering the amplitude of the $\overline{\text{Baseline}}$. The curvature of the baseline captured by x_1, which scales $\overline{\text{Baseline}}_{PC1}$, can be described by the depolymerization of the melt and the decrease of NBO/T due to the increase of minor network-forming cations Fe_2O_3 and TiO_2 and depolymerizer H_2O. Variations in oxidation state alter the extent of polymerization, with the oxidation of Fe^{2+} to Fe^{3+} increasing the NBO/T of the melt [Dingwell and

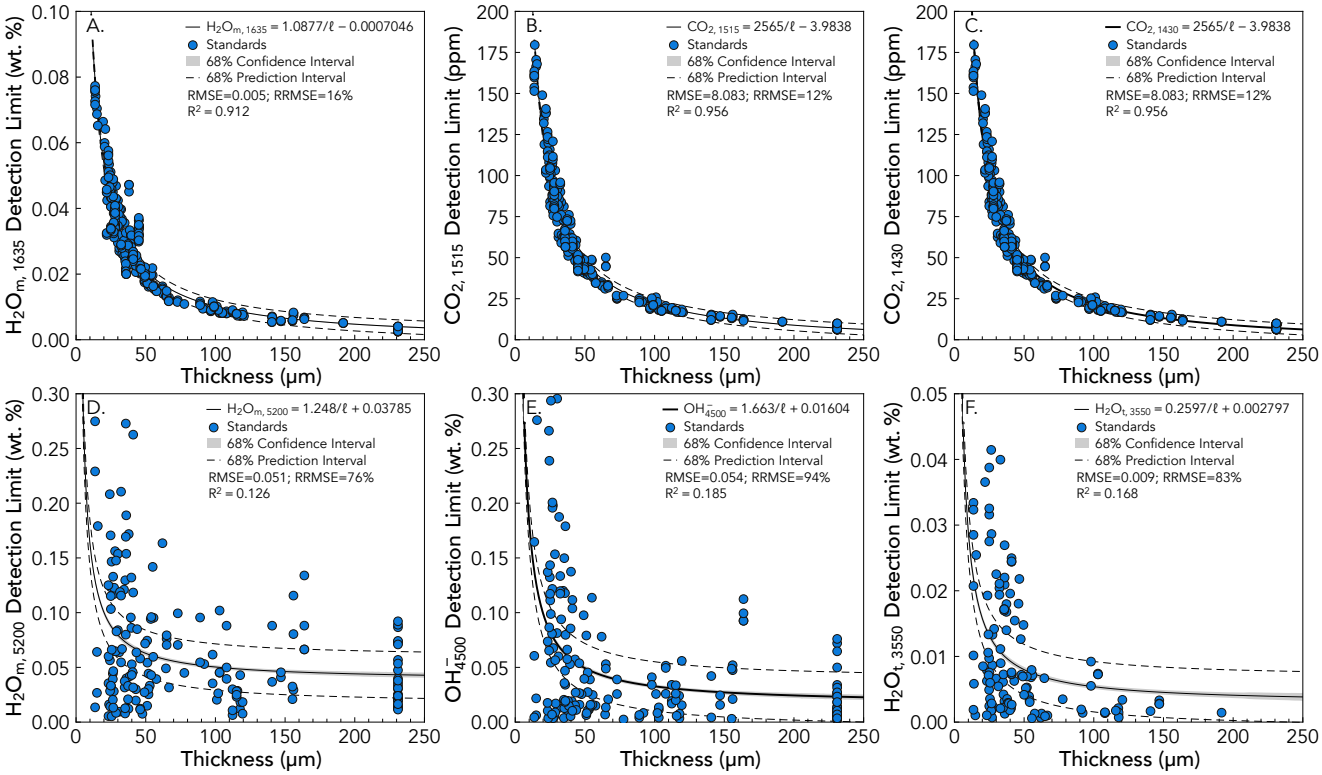

Figure 14: Detection limits across sample thickness for H_2O and CO_2 volatile species in basaltic to andesitic glasses. Three times the standard deviation of the background signal is quantified from the PyIRoGlass sampled baselines, and converted from the signal to concentration domain with the Beer-Lambert Law. The detection limit relationship with thickness takes the form of the reciprocal function, given the presence of thickness in the denominator of the Beer-Lambert Law. Three of the detection limit relationships with thickness are based upon baselines determined from the 10^6 PyIRoGlass MCMC sampled baselines: [A] $H_2O_{m,1635}$; [B] $CO^{2-}_{3,1515}$; [C] $CO^{2-}_{3,1430}$. These detection limits are well fit by the reciprocal function (R^2 greater than 0.9 and RRMSE less than 20 %), given low noise in this mid-IR region with strong constraints upon the baseline. The other three detection limit relationships with thickness are based upon baselines determined from the ALS sampled baselines: [D] $H_2O_{m,5200}$. [E] OH^-_{4500}. [F] $H_2O_{t,3550}$. These detection limits are fit by the reciprocal function (R^2 between 0.1 and 0.2 and RRMSE greater than 75 %), given more significant noise in the higher wavenumber region and lower amplitude near-IR $H_2O_{m,5200}$ and OH^-_{4500} peaks.

Virgo 1987; Mysen 2022]. This relationship between oxidation state, extent of polymerization, and the baseline potentially demonstrates some unintentional consequences for utilizing devolatilized baselines for fitting natural samples, since the experimental process of devolatilizing glasses can alter their oxidation state. x_2 does not offer clear insights into the characteristics of the melt. x_3, scaling $\overline{\text{Baseline}}_{PC3}$, demonstrates the same trends as x_1, but with stronger correlation coefficients for H_2O and density. x_4, scaling $\overline{\text{Baseline}}_{PC4}$, demonstrates the same trends as x_1 and x_3, but with the additional impact of network-modifying cations. Decreasing the proportion of network-modifying cations MgO and CaO, which charge balance tetrahedrally coordinated Al^{3+}, increases the amplitude of x_4 [Mysen and Virgo 1986; Mysen 2022].

While the chemical variability of the glasses in the validation and test dataset provide some insight into the nature of the baseline, the baseline cannot be explicitly generated or empirically fit from solely glass chemistry. The necessity of the ensemble solution offered by PyIRoGlass is underscored by the lack of a straightforward correlation between the baseline and chemistry within these datasets. A full understanding of these relationships remains elusive and is beyond the scope of this paper. The publication and availability of more glass spectra, with a wide range of chemical compositions, will allow for further exploration.

4 Conclusions

PyIRoGlass is a novel Bayesian method with MCMC sampling developed and validated for the purpose of reproducibly processing the transmission FTIR spectra of basaltic to andesitic glasses, where carbon is primarily dissolved as CO^{2-}_3, in the open-source Python language. The utility of PyIRoGlass emerges when devolatilized baselines are not readily available, which is particularly true of melt inclusion studies. All parameters within the Beer-Lambert Law—including the baseline, molar absorptivity, thickness, and density—are closely examined to quantify their associated uncertainties. Inverting for molar absorptivity as a function of compositional parameter properly quantifies both the uncertainty of the inversion and of the composition. PyIRoGlass allows for the sampling of all likely baselines and peaks to iteratively solve for the best-fit coefficients and parameters, in

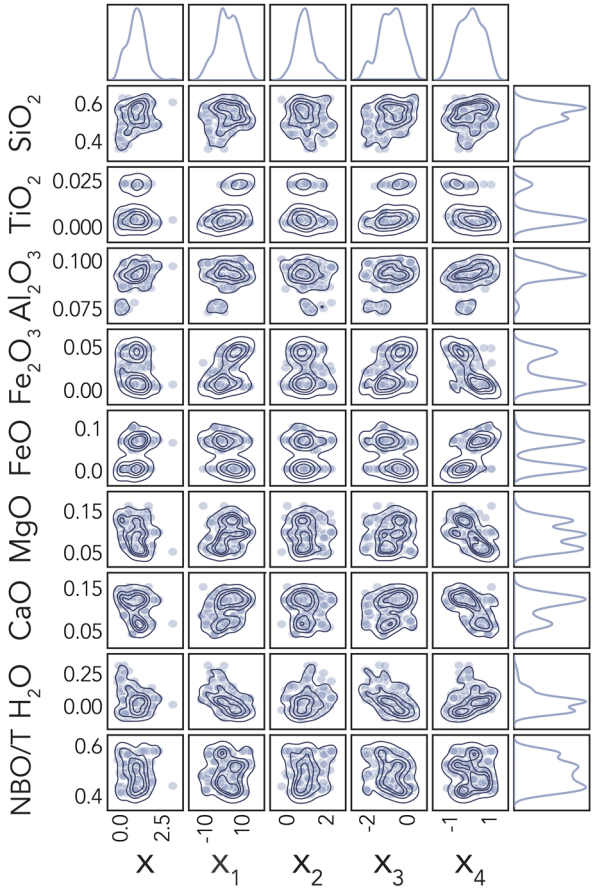

Figure 15: Posterior probability distributions of the baseline-generating coefficients, including x_0 scaling $\overline{\text{Baseline}}$ and x_{1-4} scaling $\overline{\text{Baseline}}_{\text{PC1-4}}$, and chemical composition of the melt in the validation and test datasets. x_0 is weakly correlated with SiO_2 and Al_2O_3. x_1 is correlated with Fe_2O_3, FeO, TiO_2, and H_2O. x_2 is uncorrelated. x_3 is correlated with Fe_2O_3, FeO, and H_2O. x_4 is correlated with Fe_2O_3, FeO, MgO, CaO, and H_2O.

order to determine volatile concentrations with robust estimates of uncertainties.

`PyIRoGlass` is currently tuned to arc basalts and andesites; future work could identify the shapes and variability in the baselines for other compositions such as alkali basalts, dacites, and rhyolites. The performance of `PyIRoGlass` is dependent on the model and calibration inputs, with the input dataset of transmission FTIR spectra with volatiles below detection fundamentally determining the shape and variability of the baseline. As the dataset of spectra with volatiles below detection expands, our understanding of the shape and variability of the baseline might shift to fit a wider array of compositions. Composition-dependent calibrations of molar absorptivity might similarly change as datasets evolve. Open-sourcing this `Python` code encourages iterative development and enhancement. Future validation may target further analysis of the relationships between composition and baseline variability for improved sampling of parameter space, as well as between composition and molar absorptivity. Further development of `PyIRoGlass` will focus on the expansion of this open-source package for processing rhyolitic glasses, where carbon is primarily dissolved as CO_2.

Author Contributions

SCS, WHT, TP, and AB conceived the project. SCS wrote the manuscript, collected the Volcán de Fuego melt inclusion validation dataset and some of the test dataset, developed most of the Python code, documented the code, and interpreted the results. WHT performed the principle component analysis for fitting the CO_3^{2-} doublet and $H_2O_{m,1635}$ peak and wrote the initial MCMC parameter fitting code. TP helped scope the devolatilized baseline approach, provided guidance and support, and edited the manuscript. AB helped with data collection and interpretation, and tested the code. DR analyzed the dataset of melt inclusions with volatiles below detection. YM and HL synthesized, analyzed, and provided glasses for validation and testing. WM developed the inversion code with SCS for determining molar absorptivities and critically contributed to understandings of robust uncertainty and error analysis. All authors provided feedback on the results and manuscript.

Acknowledgements

We thank Renat Almeev, Alexander Lloyd, François Holtz, Katherine Kelley, Sally Newman, and Tatiana Shishkina for providing the glasses analyzed by FTIR in the test dataset. Elizabeth Cottrell generously provided FTIR spectra of the back-arc basin glasses from the Mariana Trough and Lau Basin in the test dataset. Shuo (Echo) Ding assisted with the FTIR analyses of the Brounce et al. [2021] samples and Céline Martin was instrumental in supporting EPMA analyses of both test and validation datasets. Penny Wieser provided invaluable guidance on unit testing and advice on ensuring this package's availability on PyPi and ReadTheDocs. We extend our gratitude to the two anonymous reviewers for their insightful and constructive comments and suggestions. Finally, we acknowledge the editorial work of Jamie Farquharson and the editorial team at Volcanica.

Data Availability

All data and code are available on GitHub (`github.com/sarahshi/PyIRoGlass`), with documentation and examples at Read The Docs (pyiroglass.readthedocs.io). The code (Version 0.6.3) and data from this manuscript are archived on Zenodo [Shi 2024, DOI: 10.5281/zenodo.12735203]. The code can be run through the Google `Colab` cloud server (`colab.research.google.com/github/SarahShi/PyIRoGlass/blob/main/PyIRoGlass_RUN_colab.ipynb`). The Supplementary Material contains information on the code package structure, datasets, inversion techniques, and analytical methodologies.

Copyright Notice

© The Author(s) 2024. 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

Acocella, J., M. Tomozawa, and E. Watson (1984). "The nature of dissolved water in sodium silicate glasses and its effect on various properties". *Journal of Non-Crystalline Solids* 65(2-3), pages 355–372. DOI: 10.1016/0022-3093(84)90058-9.

Allabar, A. and M. Nowak (2020). "(High spatial resolution analysis of H_2O in silicate glass using attenuated total reflection FTIR spectroscopy coupled with a focal plane array detector)". *Chemical Geology* 556, page 119833. DOI: 10.1016/j.chemgeo.2020.119833.

Allison, C. M., K. Roggensack, and A. B. Clarke (2019). "H_2O–CO_2 solubility in alkali-rich mafic magmas: new experiments at mid-crustal pressures". *Contributions to Mineralogy and Petrology* 174(7), pages 1–24. DOI: 10.1007/s00410-019-1592-4.

Analytical Methods Committee (1987). "Recommendations for the definition, estimation and use of the detection limit". *Analyst* 112(2), pages 199–204. DOI: 10.1039/AN9871200199.

Aubaud, C., H. Bureau, C. Raepsaet, H. Khodja, A. C. Withers, M. M. Hirschmann, and D. R. Bell (2009). "Calibration of the infrared molar absorption coefficients for H in olivine, clinopyroxene and rhyolitic glass by elastic recoil detection analysis". *Chemical Geology* 262(1-2), pages 78–86. DOI: 10.1016/j.chemgeo.2009.01.001.

Barth, A. (2021). "Ascent rates and volatiles of explosive basaltic volcanism". PhD thesis. Columbia University in the City of New York. DOI: 10.7916/d8-zkzt-a130.

Behrens, H. (1995). "Determination of water solubilities in high-viscosity melts: an experimental study on $NaAlSi_3O_8$ and $KAlSi_3O_8$ melts". *European Journal of Mineralogy*, pages 905–920. DOI: 10.1127/ejm/7/4/0905.

Behrens, H., V. Misiti, C. Freda, F. Vetere, R. E. Botcharnikov, and P. Scarlato (2009). "Solubility of H_2O and CO_2 in ultrapotassic melts at 1200 and 1250°C and pressure from 50 to 500 MPa". *American Mineralogist* 94(1), pages 105–120. DOI: 10.2138/am.2009.2796.

Behrens, H., C. Romano, M. Nowak, F. Holtz, and D. B. Dingwell (1996). "Near-infrared spectroscopic determination of water species in glasses of the system MAlSi3O8 (M= Li, Na, K): an interlaboratory study". *Chemical Geology* 128(1-4), pages 41–63. DOI: 10.1016/0009-2541(95)00162-X.

Blundy, J. and K. Cashman (2001). "Ascent-driven crystallisation of dacite magmas at Mount St Helens, 1980–1986". *Contributions to Mineralogy and Petrology* 140(6), pages 631–650. DOI: doi.org/10.1007/s004100000219.

Branch, M. A., T. F. Coleman, and Y. Li (1999). "A Subspace, Interior, and Conjugate Gradient Method for Large-Scale Bound-Constrained Minimization Problems". *SIAM Journal on Scientific Computing* 21(1), pages 1–23. DOI: 10.1137/S1064827595289108.

Brooker, R., S. Kohn, J. Holloway, and P. McMillan (2001a). "Structural controls on the solubility of CO_2 in silicate melts: part I: bulk solubility data". *Chemical Geology* 174(1-3), pages 225–239. DOI: 10.1016/S0009-2541(00)00353-3.

– (2001b). "Structural controls on the solubility of CO_2 in silicate melts: part II: IR characteristics of carbonate groups in silicate glasses". *Chemical Geology* 174(1-3), pages 241–254. DOI: 10.1016/S0009-2541(00)00318-1.

Brounce, M., M. K. Reagan, K. A. Kelley, E. Cottrell, K. Shimizu, and R. Almeev (2021). "Covariation of slab tracers, volatiles, and oxidation during subduction initiation". *Geochemistry, Geophysics, Geosystems* 22(6), e2021GC009823. DOI: 10.1029/2021GC009823.

Carroll, M. R. and J. G. Blank (1997). "The solubility of H_2O in phonolitic melts". *American Mineralogist* 82(5-6), pages 549–556. DOI: 10.2138/am-1997-5-615.

Carvajal, R. C., L. E. Arias, H. O. Garces, and D. G. Sbarbaro (2016). "Comparative analysis of a principal component analysis-based and an artificial neural network-based method for baseline removal". *Applied Spectroscopy* 70(4), pages 604–617. DOI: 10.1177/0003702816631293.

Cressie, N. (1990). "The origins of kriging". *Mathematical geology* 22, pages 239–252. DOI: 10.1007/BF00889887.

Cubillos, P., J. Harrington, T. J. Loredo, N. B. Lust, J. Blecic, and M. Stemm (2017). "On Correlated-noise Analyses Applied to Exoplanet Light Curves". *The Astronomical Journal* 153(1), page 3. DOI: 10.3847/1538-3881/153/1/3.

Delgado, R. (2022). "Misuse of Beer–Lambert Law and other calibration curves". *Royal Society Open Science* 9(2), page 211103. DOI: 10.1098/rsos.211103.

Di Matteo, V., M. Carroll, H. Behrens, F. Vetere, and R. Brooker (2004). "Water solubility in trachytic melts". *Chemical Geology* 213(1-3), pages 187–196. DOI: 10.1016/j.chemgeo.2004.08.042.

Di Matteo, V., A. Mangiacapra, D. Dingwell, and G. Orsi (2006). "Water solubility and speciation in shoshonitic and latitic melt composition from Campi Flegrei Caldera (Italy)". *Chemical Geology* 229(1-3), pages 113–124. DOI: 10.1016/j.chemgeo.2006.01.015.

Dingwell, D. B. and D. Virgo (1987). "The effect of oxidation state on the viscosity of melts in the system Na2O-FeO-Fe2O3-SiO2". *Geochimica et Cosmochimica Acta* 51(2), pages 195–205. DOI: 10.1016/0016-7037(87)90231-6.

Dixon, J. E. and D. A. Clague (2001). "Volatiles in basaltic glasses from Loihi Seamount, Hawaii: Evidence for a relatively dry plume component". *Journal of Petrology* 42(3), pages 627–654. DOI: 10.1093/petrology/42.3.627.

Dixon, J. E. and V. Pan (1995). "Determination of the molar absorptivity of dissolved carbonate in basanitic glass". *American Mineralogist* 80(11-12), pages 1339–1342. DOI: 10.2138/am-1995-11-1224.

Dixon, J. E., E. Stolper, and J. R. Delaney (1988). "Infrared spectroscopic measurements of CO_2 and H_2O in Juan de Fuca Ridge basaltic glasses". *Earth and Planetary Science Letters* 90(1), pages 87–104. DOI: 10.1016/0012-821X(88)90114-8.

Dixon, J. E. and E. M. Stolper (1995). "An experimental study of water and carbon dioxide solubilities in mid-ocean ridge basaltic liquids. Part II: applications to degassing". *Journal of Petrology* 36(6), pages 1633–1646. DOI: 10.1093/oxfordjournals.petrology.a037268.

Dixon, J. E., E. M. Stolper, and J. R. Holloway (1995). "An experimental study of water and carbon dioxide solubilities in mid-ocean ridge basaltic liquids. Part I: calibration and solubility models". *Journal of Petrology* 36(6), pages 1607–1631. DOI: 10.1093/oxfordjournals.petrology.a037267.

Dobson, P. F., S. Epstein, and E. M. Stolper (1989). "Hydrogen isotope fractionation between coexisting vapor and silicate glasses and melts at low pressure". *Geochimica et Cosmochimica Acta* 53(10), pages 2723–2730. DOI: 10.1016/0016-7037(89)90143-9.

Duncan, M. S. and R. Dasgupta (2015). "Pressure and temperature dependence of CO_2 solubility in hydrous rhyolitic melt: implications for carbon transfer to mantle source of volcanic arcs via partial melt of subducting crustal lithologies". *Contributions to Mineralogy and Petrology* 169(6), pages 1–19. DOI: 10.1007/s00410-015-1144-5.

Eilers, P. H. (2004). "Parametric time warping". *Analytical Chemistry* 76(2), pages 404–411. DOI: 10.1021/ac034800e.

Fiege, A., F. Holtz, H. Behrens, C. W. Mandeville, N. Shimizu, L. S. Crede, and J. Göttlicher (2015). "Experimental investigation of the S and S-isotope distribution between H_2O–S±Cl fluids and basaltic melts during decompression". *Chemical Geology* 393, pages 36–54. DOI: 10.1016/j.chemgeo.2014.11.012.

Fine, G. and E. Stolper (1985). "The speciation of carbon dioxide in sodium aluminosilicate glasses". *Contributions to Mineralogy and Petrology* 91(2), pages 105–121. DOI: 10.1007/BF00377759.

– (1986). "Dissolved carbon dioxide in basaltic glasses: concentrations and speciation". *Earth and Planetary Science Letters* 76(3-4), pages 263–278. DOI: 10.1016/0012-821X(86)90078-6.

Gelman, A. and D. B. Rubin (1992). "Inference from iterative simulation using multiple sequences". *Statistical science*, pages 457–472. DOI: 10.1214/ss/1177011136.

Hauri, E., J. Wang, J. E. Dixon, P. L. King, C. Mandeville, and S. Newman (2002). "SIMS analysis of volatiles in silicate glasses: 1. Calibration, matrix effects and comparisons with FTIR". *Chemical Geology* 183(1-4), pages 99–114. DOI: 10.1016/S0009-2541(01)00375-8.

Helo, C., M.-A. Longpré, N. Shimizu, D. A. Clague, and J. Stix (2011). "Explosive eruptions at mid-ocean ridges driven by CO2-rich magmas". *Nature Geoscience* 4(4), pages 260–263. DOI: 10.1038/ngeo1104.

Hess, K. and D. B. Dingwell (1996). "Viscosities of hydrous leucogranitic melts: A non-Arrhenian model". *American Mineralogist: Journal of Earth and Planetary Materials* 81(9-10), pages 1297–1300. DOI: 10.2138/am-1996-9-1031.

Howie, R., J. Zussman, and W. Deer (1992). *An introduction to the rock-forming minerals*. Longman London, UK. DOI: 10.1180/DHZ.

Iacovino, K. and C. B. Till (2019). "DensityX: A program for calculating the densities of magmatic liquids up to 1,627 C and 30 kbar". *Volcanica* 2(1), pages 1–10. DOI: doi.org/10.30909/vol.02.01.0110.

Ihinger, P., R. Hervig, and P. McMillan (1994). "Analytical Methods for Volatiles in Glasses". *Reviews in Mineralogy and Geochemistry* 30, pages 67–121. DOI: 10.1515/9781501509674-008.

Jakobsson, S. (1997). "Solubility of water and carbon dioxide in an Icelandite at 1400°C and 10 kilobars". *Contributions to Mineralogy and Petrology* 127(1-2), pages 129–135. DOI: 10.1007/s004100050270.

Jendrzejewski, N., T. W. Trull, F. Pineau, and M. Javoy (1997). "Carbon solubility in Mid-Ocean Ridge basaltic melt at low pressures (250–1950 bar)". *Chemical Geology* 138(1-2), pages 81–92. DOI: 10.1016/S0009-2541(96)00176-3.

King, P., T. Vennemann, J. Holloway, R. Hervig, J. Lowenstern, and J. Forneris (2002). "Analytical techniques for volatiles: A case study using intermediate (andesitic) glasses". *American Mineralogist* 87(8-9), pages 1077–1089. DOI: 10.2138/am-2002-8-904.

Krekel, H., B. Oliveira, R. Pfannschmidt, F. Bruynooghe, B. Laugher, and F. Bruhin (2004). *pytest 8.2.0*.

Krige, D. G. (1951). "A statistical approach to some basic mine valuation problems on the Witwatersrand". *Journal of the Southern African Institute of Mining and Metallurgy* 52(6), pages 119–139. DOI: 10.10520/AJA0038223X_4792.

Lee, H. J., Y. Moussallam, E. F. Rose-Koga, L. Piani, J. Villeneuve, N. Bouden, A. A. Gurenko, B. D. Monteleone, G. A. Gaetani, and É. Deloule (2024). "High–precision determination of carbon stable isotope in silicate glasses by secondary ion mass spectrometry: Evaluation of international standards". *EarthArXiv*. DOI: 10.31223/X5R970.

Lee, L. C., C. Liong, O. Khairul, and A. A. Jemain (2017). "Effects of baseline correction algorithms on forensic classification of paper based on ATR-FTIR spectrum and principal component analysis (PCA)". *Pertanika Journal of Science and Technology* 25(3), pages 767–774.

Leschik, M., G. Heide, G. Frischat, H. Behrens, M. Wiedenbeck, N. Wagner, K. Heide, H. Geißler, and U. Reinholz (2004). "Determination of H_2O and D_2O contents in rhyolitic glasses". *Physics and Chemistry of Glasses* 45(4), pages 238–251.

Lesher, C. E. and F. J. Spera (2015). "Thermodynamic and transport properties of silicate melts and magma". *The Encyclopedia of Volcanoes*. Elsevier, pages 113–141. DOI: 10.1016/B978-0-12-385938-9.00005-5.

Lesne, P., B. Scaillet, M. Pichavant, G. Iacono-Marziano, and J.-M. Beny (2011). "The H_2O solubility of alkali basaltic melts: an experimental study". *Contributions to Mineralogy and Petrology* 162(1), pages 133–151. DOI: 10.1007/s00410-010-0588-x.

Lin, L. I.-K. (1989). "A concordance correlation coefficient to evaluate reproducibility". *Biometrics*, pages 255–268. DOI: 10.2307/2532051.

Lowenstern, J. B. (2001). "Carbon dioxide in magmas and implications for hydrothermal systems". *Mineralium Deposita* 36(6), pages 490–502. DOI: 10.1007/s001260100185.

Lowenstern, J. B. and B. W. Pitcher (2013). "Analysis of H_2O in silicate glass using attenuated total reflectance (ATR) micro-FTIR spectroscopy". *American Mineralogist* 98(10), pages 1660–1668. DOI: 10.2138/am.2013.4466.

Lytle, M. L., E. H. Hauri, and K. A. Kelley (2023). *Volatile and trace elements in submarine glasses from global back-arc spreading centers*. Interdisciplinary Earth Data Alliance (IEDA). DOI: 10.26022/IEDA/112829.

Maclennan, J. (2017). "Bubble formation and decrepitation control the CO_2 content of olivine-hosted melt inclusions". *Geochemistry, Geophysics, Geosystems* 18(2), pages 597–616. DOI: 10.1002/2016GC006633.

Mandeville, C. W., J. D. Webster, M. J. Rutherford, B. E. Taylor, A. Timbal, and K. Faure (2002). "Determination of molar absorptivities for infrared absorption bands of H_2O in andesitic glasses". *American Mineralogist* 87(7), pages 813–821. DOI: 10.2138/am-2002-0702.

McIntosh, I. M., A. R. Nichols, K. Tani, and E. W. Llewellin (2017). "Accounting for the species-dependence of the 3500 cm^{-1} H_2O_t infrared molar absorptivity coefficient: Implications for hydrated volcanic glasses". *American Mineralogist: Journal of Earth and Planetary Materials* 102(8), pages 1677–1689. DOI: 10.2138/am-2017-5952CCBY.

Menke, W. (2018). *Geophysical data analysis: Discrete inverse theory*. Academic Press. DOI: 10.1016/B978-0-12-490920-5.X5001-7.

Mercier, M., A. Di Muro, N. Métrich, D. Giordano, O. Belhadj, and C. W. Mandeville (2010). "Spectroscopic analysis (FTIR, Raman) of water in mafic and intermediate glasses and glass inclusions". *Geochimica et Cosmochimica Acta* 74(19), pages 5641–5656. DOI: 10.1016/j.gca.2010.06.020.

Moussallam, Y., W. Tobbin, T. Plank, H. Bureau, H. Khodja, Y. Guan, C. Ma, M. B. Baker, E. M. Stolper, F. U. Naab, B. D. Monteleone, G. A. Gaetani, K. Shimizu, U. Takayuki, H. J. Lee, S. Ding, S. C. Shi, and E. F. Rose-Koga (2024). "ND70 series basaltic glass reference materials for volatile element (H_2O, CO_2, S, Cl, F) analysis and the C ionisation efficiency suppression effect of water in silicate glasses in SIMS analysis". *Geostandards and Geoanalytical Research*. DOI: 10.31223/X5QQ4P.

Mysen, B. (2022). *Mass Transport in Magmatic Systems*. Elsevier. DOI: 10.1016/C2019-0-03232-3.

Mysen, B. O. and P. Richet (2018). *Silicate glasses and melts*. Elsevier. DOI: 10.1016/C2018-0-00864-6.

Mysen, B. O. and D. Virgo (1980a). "Solubility mechanisms of carbon dioxide in silicate melts: a Raman spectroscopic study". *American Mineralogist* 65(9-10), pages 885–899.

– (1980b). "The solubility behavior of CO_2 in melts on the join $NaAlSi_3O_8$- $CaAl_2Si_2O_8$-CO_2 at high pressures and temperatures: a Raman spectroscopic study". *American Mineralogist* 65(11-12), pages 1166–1175.

– (1986). "Volatiles in silicate melts at high pressure and temperature: 1. Interaction between OH groups and Si^{4+}, Al^{3+}, Ca^{2+}, Na^{+} and H^{+}". *Chemical Geology* 57(3-4), pages 303–331. DOI: 10.1016/0009-2541(86)90056-2.

Newman, S., E. Stolper, and R. Stern (2000). "H_2O and CO_2 in magmas from the Mariana arc and back arc systems". *Geochemistry, Geophysics, Geosystems* 1(5). DOI: 10.1029/1999GC000027.

Newman, S., E. M. Stolper, and S. Epstein (1986). "Measurement of water in rhyolitic glasses; calibration of an infrared spectroscopic technique". *American Mineralogist* 71(11-12), pages 1527–1541.

Nichols, A. R. and R. Wysoczanski (2007). "Using micro-FTIR spectroscopy to measure volatile contents in small and unexposed inclusions hosted in olivine crystals". *Chemical Geology* 242(3-4), pages 371–384. DOI: 10.1016/j.chemgeo.2007.04.007.

Nishikida, K., E. Nishio, and R. W. Hannah (1996). *Selected applications of modern FTIR techniques*. CRC Press. DOI: 10.1201/9780203739983.

Nowak, M. and H. Behrens (1995). "The speciation of water in haplogranitic glasses and melts determined by in situ near-infrared spectroscopy". *Geochimica et Cosmochimica Acta* 59(16), pages 3445–3450. DOI: 10.1016/0016-7037(95)00237-T.

Ochs III, F. A. and R. A. Lange (1999). "The density of hydrous magmatic liquids". *Science* 283(5406), pages 1314–1317. DOI: 10.1126/science.283.5406.1314.

Ohlhorst, S., H. Behrens, and F. Holtz (2001). "Compositional dependence of molar absorptivities of near-infrared OH- and H_2O bands in rhyolitic to basaltic glasses". *Chemical Geology* 174(1-3), pages 5–20. DOI: 10.1016/S0009-2541(00)00303-X.

Okumura, S., M. Nakamura, and S. Nakashima (2003). "Determination of molar absorptivity of IR fundamental OH-stretching vibration in rhyolitic glasses". *American Mineralogist* 88(11-12), pages 1657–1662. DOI: 10.2138/am-2003-11-1204.

Okumura, S. and S. Nakashima (2005). "Molar absorptivities of OH and H_2O in rhyolitic glass at room temperature and at 400–600°C". *American Mineralogist* 90(2-3), pages 441–447. DOI: 10.2138/am.2005.1740.

Pandya, N., D. W. Muenow, and S. K. Sharma (1992). "The effect of bulk composition on the speciation of water in submarine volcanic glasses". *Geochimica et Cosmochimica Acta* 56(5), pages 1875–1883. DOI: 10.1016/0016-7037(92)90317-C.

Pearce, J. A., M. Ernewein, S. H. Bloomer, L. M. Parson, B. J. Murton, and L. E. Johnson (1994). "Geochemistry of Lau Basin volcanic rocks: influence of ridge segmentation and arc proximity". *Geological Society, London, Special Publications* 81(1), pages 53–75. DOI: 10.1144/GSL.SP.1994.081.01.04.

Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al. (2011). "Scikit-learn: Machine learning in Python". *Journal of Machine Learning Research* 12, pages 2825–2830. DOI: 10.48550/arXiv.1201.0490.

Peng, J., S. Peng, A. Jiang, J. Wei, C. Li, and J. Tan (2010). "Asymmetric least squares for multiple spectra baseline correction". *Analytica Chimica Acta* 683(1), pages 63–68. DOI: 10.1016/j.aca.2010.08.033.

Rasmussen, D. J. (2019). *The Aleutian Arc through and through: Subduction dynamics and the generation, storage, and eruption of hydrous magmas*. Columbia University. DOI: 10.7916/d8-vmrb-pf70.

Savitzky, A. and M. J. Golay (1964). "Smoothing and differentiation of data by simplified least squares procedures". *Analytical Chemistry* 36(8), pages 1627–1639. DOI: 10.1021/ac60214a047.

Shi, S. (2024). *sarahshi/PyIRoGlass: PyIRoGlass v0.6.3*. Version v0.6.3. DOI: 10.5281/zenodo.12735203.

Shi, S., A. Barth, T. Plank, W. Towbin, O. Flores, and C. Arias (2021). "Magma stalling weakens eruption". *AGU Fall Meeting Abstracts*. Volume 2021, V25E–09.

Shishkina, T. A., R. E. Botcharnikov, F. Holtz, R. R. Almeev, A. M. Jazwa, and A. A. Jakubiak (2014). "Compositional and pressure effects on the solubility of H_2O and CO_2 in mafic melts". *Chemical Geology* 388, pages 112–129. DOI: 10.1016/j.chemgeo.2014.09.001.

Shishkina, T. A., R. E. Botcharnikov, F. Holtz, R. R. Almeev, and M. V. Portnyagin (2010). "Solubility of H_2O-and CO_2-bearing fluids in tholeiitic basalts at pressures up to 500 MPa". *Chemical Geology* 277(1-2), pages 115–125. DOI: 10.1016/j.chemgeo.2010.07.014.

Silver, L. A., P. D. Ihinger, and E. Stolper (1990). "The influence of bulk composition on the speciation of water in silicate glasses". *Contributions to Mineralogy and Petrology* 104, pages 142–162. DOI: 10.1007/BF00306439.

Silver, L. A. and E. Stolper (1989). "Water in albitic glasses". *Journal of Petrology* 30(3), pages 667–709. DOI: 10.1093/petrology/30.3.667.

Sisson, T. and T. Grove (1993). "Experimental investigations of the role of H_2O in calc-alkaline differentiation and subduction zone magmatism". *Contributions to mineralogy and petrology* 113(2), pages 143–166. DOI: 10.1007/BF00283225.

Stabile, P., E. Appiah, M. Bello, G. Giuli, E. Paris, and M. R. Carroll (2020). "New IR spectroscopic data for determination of water abundances in hydrous pantelleritic glasses". *American Mineralogist* 105(7), pages 1060–1068. DOI: 10.2138/am-2020-7363.

Stolper, E. (1982). "Water in silicate glasses: an infrared spectroscopic study". *Contributions to Mineralogy and Petrology* 81(1), pages 1–17. DOI: 10.1007/BF00371154.

– (1989). "Temperature dependence of the speciation of water in rhyolitic melts and glasses". *American Mineralogist* 74(11-12), pages 1247–1257.

Sun, W., R. Binns, A. Fan, V. S. Kamenetsky, R. Wysoczanski, G. Wei, Y. Hu, and R. Arculus (2007). "Chlorine in submarine volcanic glasses from the eastern Manus basin". *Geochimica et Cosmochimica Acta* 71(6), pages 1542–1552. DOI: 10.1016/j.gca.2006.12.003.

Tamic, N., H. Behrens, and F. Holtz (2001). "The solubility of H_2O and CO_2 in rhyolitic melts in equilibrium with a mixed CO_2–H_2O fluid phase". *Chemical Geology* 174(1-3), pages 333–347. DOI: 10.1016/S0009-2541(00)00324-7.

Ter Braak, C. J. (2006). "A Markov Chain Monte Carlo version of the genetic algorithm Differential Evolution: easy Bayesian computing for real parameter spaces". *Statistics and Computing* 16, pages 239–249. DOI: 10.1007/s11222-006-8769-1.

Ter Braak, C. J. and J. A. Vrugt (2008). "Differential Evolution Markov Chain with snooker updater and fewer chains". *Statistics and Computing* 18(435-446). DOI: 10.1007/s11222-008-9104-9.

Thibault, Y. and J. R. Holloway (1994). "Solubility of CO_2 in a Ca-rich leucitite: effects of pressure, temperature, and oxygen fugacity". *Contributions to Mineralogy and Petrology* 116, pages 216–224. DOI: 10.1007/BF00310701.

Vetere, F., H. Behrens, F. Holtz, and D. R. Neuville (2006). "Viscosity of andesitic melts—new experimental data and a revised calculation model". *Chemical Geology* 228(4), pages 233–245. DOI: 10.1016/j.chemgeo.2005.10.009.

Vetere, F., R. E. Botcharnikov, F. Holtz, H. Behrens, and R. De Rosa (2011). "Solubility of H_2O and CO_2 in shoshonitic melts at 1250°C and pressures from 50 to 400 MPa: implications for Campi Flegrei magmatic systems". *Journal of Volcanology and Geothermal Research* 202(3-4), pages 251–261. DOI: 10.1016/j.jvolgeores.2011.03.002.

Vetere, F., F. Holtz, H. Behrens, R. E. Botcharnikov, and S. Fanara (2014). "The effect of alkalis and polymerization on the solubility of H_2O and CO_2 in alkali-rich silicate melts". *Contributions to Mineralogy and Petrology* 167, pages 1–17. DOI: 10.1007/s00410-014-1014-6.

Virtanen, P., R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors (2020). "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python". *Nature Methods* 17, pages 261–272. DOI: 10.1038/s41592-019-0686-2.

Von Aulock, F. W., B. M. Kennedy, C. I. Schipper, J. Castro, D. Martin, C. Oze, J. M. Watkins, P. J. Wallace, L. Puskar, F. Bégué, et al. (2014). "Advances in Fourier transform infrared spectroscopy of natural glasses: From sample preparation to data analysisdvances in Fourier transform infrared spectroscopy of natural glasses: From sample preparation to data analysis". *Lithos* 206, pages 52–64. DOI: 10.1016/j.lithos.2014.07.017.

Withers, A. C. and H. Behrens (1999). "Temperature-induced changes in the NIR spectra of hydrous albitic and rhyolitic glasses between 300 and 100 K". *Physics and Chemistry of Minerals* 27(2), pages 119–132. DOI: 10.1007/s002690050248.

Wysoczanski, R. and K. Tani (2006). "Spectroscopic FTIR imaging of water species in silicic volcanic glasses and melt inclusions: An example from the Izu-Bonin arc". *Journal of Volcanology and Geothermal Research* 156(3-4), pages 302–314. DOI: 10.1016/j.jvolgeores.2006.03.024.

Xue, X. (2009). "Water speciation in hydrous silicate and aluminosilicate glasses: Direct evidence from ^{29}Si-^{1}H and ^{27}Al-^{1}H double-resonance NMR". *American Mineralogist* 94(2-3), pages 395–398. DOI: 10.2138/am.2009.3088.

Yamashita, S., T. Kitamura, and M. Kusakabe (1997). "Infrared spectroscopy of hydrous glasses of arc magma compositions". *Geochemical Journal* 31(3), pages 169–174. DOI: 10.2343/geochemj.31.169.