Determining the pressure-temperature-composition (P-T-X) conditions of magma storage

Penny E Wieser[a], Matthew LM Gleeson[a], Simon Matthews[b,c], Charlotte DeVitre[a], and Esteban Gazel[d], [a]Earth and Planetary Sciences, UC Berkeley, Berkeley, CA, United States; [b]Earth Science Department, University of Cambridge, Cambridge, United Kingdom; [c]Institute of Earth Sciences, University of Iceland, Reykjavík, Iceland; [d]Department of Earth and Atmospheric Sciences, Cornell University, Ithaca, NY, United States

This chapter has been reviewed by the section editor Matthew Kohn.

Abstract

Determining the pressures and temperatures at which melts are stored in the crust and upper mantle, and the major element composition, redox state and volatile contents of these melts, is vital to constrain the structure and dynamics of magmatic plumbing systems. In turn, constraining these parameters helps when interpretating periods of unrest at active volcanoes, and when deciphering the geochemical and structural evolution of the Earth's lithosphere. We review common thermobarometers, hygrometers and chemometers based on mineral and/or liquid compositions, before discussing recent advances in melt and fluid inclusion barometry, Raman-based elastic thermobarometry, and thermodynamic modeling methods. Where possible, we investigate the accuracy and precision of each technique, and the implications for the application of each method to different research questions.

Keywords

Amphibole; Chemometry; Experimental petrology; FAIR research framework; Fluid inclusions; Hygrometry; Melt inclusions; MELTS modeling; Mineral thermobarometry; Plagioclase; Pyroxene; Raman spectroscopy; Thermodynamics

Key points

- Many mineral-melt barometers have standard errors of ±2–3 kbar (±7–11 km), which limits their ability to provide detailed constraints on magma storage in relatively thin-crusted settings (MORB and OIB).
- Different models/equations can yield very different PT conditions; model choice must be considered as one of the largest sources of systematic error in a given study.
- The paucity of independent test datasets for many methods makes it difficult to truly assess their performance; statistics calculated using the calibration dataset where only one variable is treated as an unknown are overly optimistic of performance in natural systems.
- It has become increasingly clear that melt inclusion studies which did not account for the vapor bubble may have underestimated storage depths by more than a factor of 2, requiring re-evaluation of melt inclusion volatile contents and storage depths globally.

Introduction

Determining the pressures, and therefore depths, at which magmas are stored and evolve in the crust and upper mantle is vital to understand the chemical and structural evolution of volcanic plumbing systems, with implications for our understanding of the formation of mineral deposits, and the evolution of the Earth's lithosphere (Lee and Anderson, 2015). Precisely constraining magma storage depths at a specific volcano using past eruptive deposits can also provide vital context to help inform the interpretation of monitoring signals during periods of volcanic unrest (e.g., distinguishing magmatic and hydrothermal signals, Pritchard et al., 2019).

Constraining the temperatures of different magmatic processes can reveal the thermal evolution of magmatic systems and are vital inputs for many other common workflows in igneous petrology, such as calculations of timescales from elemental diffusion in erupted crystals (termed "diffusion chronometry" or "geospeedometry"). In fact, because diffusion rates are strongly sensitive to temperature (following an Arrhenius relationship), uncertainty in temperature is one of the largest sources of error when obtaining timescales using these chronometers (Chakraborty and Dohmen, 2022; Costa et al., 2020). For example, Mutch et al. (2019a) show that timescales calculated from Cr diffusion in spinel change from ~4000 years at 1190 °C to ~1000 years at 1230 °C. Thus, uncertainty in temperatures affects interpretations of timescales of crustal residence (Mutch et al., 2019a), re-awakening from quiescence to eruption (Shamloo and Till, 2019), and calculations of magma ascent rates (Mutch et al., 2019b). Of course, accurately constraining magma storage depths is also vital for magma ascent rate calculations, where the speed is calculated by dividing the depth to the inferred magma body by the time calculated from diffusion chronometry (Barth et al., 2019; Klügel, 1998).

Here, we describe a multitude of approaches to determine the pressures (P) and temperatures (T) at which magmas were stored in the lithosphere, as well as the chemistry of these stored melts (Fig. 1). We pay particular attention to differences between models (Section "Assessing and comparing models") and the precision and accuracy of calculations (Section "Statistics to compare models"). For barometers, we first evaluate variations in crustal thickness in a wide variety of tectonic settings to provide context when interpreting the precision of different barometry methods (Section "What accuracy and precision are required?"). Then, we review the methods used to convert measured mineral compositions with or without an equilibrium melt into magma storage pressures and temperatures (mineral barometry and thermometry), and to calculate the chemistry and H_2O contents of the melts from which a specific mineral composition grew (chemometry and hygrometry respectively, Section "Mineral thermobarometry and chemometry").

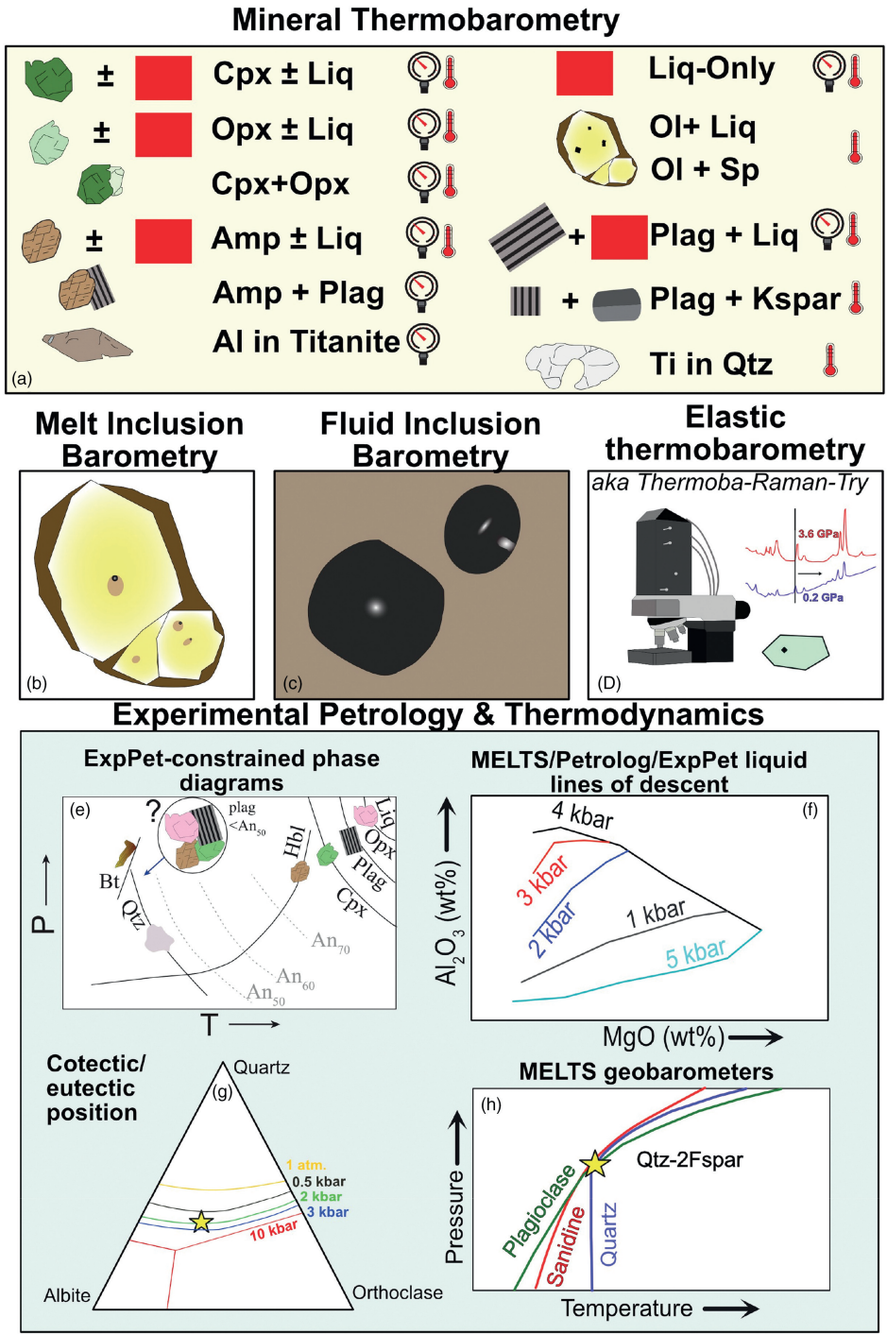

Fig. 1 Summary of some of the methods used to determine pressures and temperatures of magma storage. Panels (e) Adapted from Rutherford (2003), (f) from Cashman and Edmonds (2019), (g) and (h) adapted from Gualda et al. (2012) and Gualda and Ghiorso (2014), respectively.

We also evaluate the methods and uncertainties involved in melt inclusion barometry, with a specific focus on the influence of CO_2 stored within melt inclusion vapor bubbles on calculated pressures (Section "Melt inclusion barometry"). The general proliferation of Raman spectroscopy in the Earth Sciences (Dubessy et al., 2012) has not only been important for performing direct measurements of melt inclusion vapor bubbles; it has also pushed forwards two other barometric methods. First, Raman spectroscopic measurements of fluid inclusions can provide a substantially faster method of calculating storage depths than traditional microthermometric techniques (Section "Fluid inclusion barometry"). Second, Raman measurements of mineral inclusions in-situ can also be used to calculate pressures and temperatures (Thermoba-Raman-try, Section "Elastic thermobarometry/Thermoba-Raman-try"). Experimental petrology not only underpins the formulation and calibration of mineral-melt

thermobarometers and volatile solubility models, it can also be used to provide direct constraints on magma storage in a specific volcanic system (Section "Experimental petrology"). Finally, we discuss approaches using thermodynamic constraints on phase stability in igneous systems (Section "Thermobarometers based on thermodynamic modeling"), methods not covered in detail in this review (Section "Other methods") and summarize the general proliferation of community data repositories and open-source tools to help inform model calibration, and aid calculations of magma storage conditions (Section "Future developments should be guided by FAIR principles").

Assessing and comparing models

When calculating pressure, temperature, and/or melt chemistry from the composition of erupted melts, crystals and their inclusions, the result can be highly sensitive to the choice of model/equation/parameterization, relating the measured quantity (e.g., Cpx, Amp, melt inclusion composition) to the intensive variable (e.g., P, T). Below, we discuss the origin of all these different models, and best practices for model intercomparison.

Mineral thermobarometry (Section "Mineral thermobarometry and chemometry") utilizes chemical reactions involving the crystallization or re-equilibration of minerals which are accompanied by a change in volume (sensitive to P) and/or a change in entropy (sensitive to T, e.g., Putirka, 2008). Chemometry (and hygrometry) relies on the fact that the composition of crystallizing/re-equilibrating phases is sensitive to the composition of the melt phase. Some thermobarometry and chemometry expressions are firmly rooted in thermodynamics, with compositional terms and the functional form of the equation determined from a specific reaction. Typically, these expressions have coefficients attached to thermodynamic terms which are calibrated empirically using the measured compositions of minerals in experiments conducted at well-constrained P-T-fO_2-H2O conditions (e.g., the Plagioclase-Liquid hygrometer of Waters and Lange, 2015).

More commonly, available equations have a general form based on thermodynamics, with the addition of empirically derived terms to improve the fit to the calibration dataset (e.g., Neave and Putirka, 2017). For example, many of the Cpx-Liq barometers of Putirka (2008) have a functional form based on Jadeite exchange between the liquid and Cpx, which is associated with a change in volume, so is P sensitive (Putirka et al., 1996):

$$P = a + \frac{bT}{10^4} + \frac{cT}{10^4}\ \ln\left(\mathrm{Jd}^{\mathrm{Cpx-Liq}}\right) + \ \ldots \tag{1}$$

Empirical terms help to improve the fit between the calculated parameter and the known experimental conditions. The inclusion of these terms may be informed at least in part by thermodynamic reasoning (e.g., Masotta et al., 2013; Neave and Putirka, 2017; Putirka, 2008). For example, Putirka et al. (1996) EqP1 has a term for the cation fractions of Na and Al:

$$\mathrm{EqP1} = \text{Equation 1 terms} \ldots + 367\ \left(X_{Na}^{Liq} * X_{Al}^{Liq}\right) \tag{2}$$

The fact that the best fit involved the multiplication of Na and Al cation fractions implies that these two components have identical activity coefficients (Putirka et al., 1996). However, the thermodynamic interpretation is often obscured by decisions to reduce the complexity of equations, such as wanting to avoid introducing multiple temperature-dependent terms (see equation 12 of Putirka et al., 1996). In general, as calibration datasets have grown larger, more empirical terms have been added that have a less clear thermodynamic basis. For example, Putirka (2008, eq. 31) has many additional terms relative to EqP1, including terms for the Ca, Na, K, Si, Mg and Fe_t liquid cation fractions, H_2O in the liquid, and for EnFs, DiHd, and Al cation fractions in the Cpx:

$$\begin{aligned}\mathrm{Eq31} = {} & \text{Equation 1 terms} \ldots \\ & + 105.7\, X_{Ca}^{Liq} - 165.5\left(X_{Na}^{Liq} + X_{K}^{Liq}\right)^2 - 50.15\, X_{Si}^{Liq} * \left(X_{Mg}^{Liq} + X_{Fe}^{Liq}\right) - 3.178\log(DiHd) - 2.205\log(EnFs) \\ & + 0.864\log\left(X_{Al,6\ cat}^{Cpx}\right) + 0.3962 * \mathrm{H2O}\end{aligned} \tag{3}$$

New models not only arise through the addition of such empirical terms, but also through the recalibration of the coefficients of older equations. For example, Masotta et al. (2013) recalibrate the coefficients of the Putirka (2008) equations for a new calibration dataset of alkaline magmas. Other thermobarometers are purely empirical, including terms and mathematic expressions like logs, exponentials, and different powers which improve the fit to the calibration dataset without a firm tie to the thermodynamics of a specific chemical reaction. For example, the liquid-only thermometer of Helz and Thornber (1987) was calibrated based on the linear relationship between MgO and T in melting experiments on samples from the Kīlauea Iki lava lake. The amphibole-only barometer of Ridolfi and Renzulli (2012) uses multivariate least squares regression to determine the relationship between P and amphibole composition across a number of different P ranges, with highly variable functional forms (Eq1a, Eq1b and Eq1e take the exponential of amphibole cation fractions, Eq1c and Eq1d do not). The results from these equations are selected or combined using an algorithm to give a single P. This algorithm was then tweaked by Ridolfi (2021), generating yet another model.

Most recently, machine learning approaches using regression trees have been used to parameterize the relationship between mineral and melt composition and experimental conditions (e.g., Higgins et al., 2022; Jorgenson et al., 2022; Petrelli et al., 2020). These approaches have no underlying thermodynamic basis. They are trained using measured oxide contents, rather than thermodynamically informed mineral components such as Jadeite, Diopside, etc. The lack of thermodynamic basis may affect

their ability to extrapolate beyond the compositions and P-T conditions for which they are calibrated. One advantage is that these models can be easily updated as new experimental data becomes available if the code for model training is released.

The diversity of empirical, thermodynamic, and machine-learning models for mineral-based thermobarometry/hygrometry/chemometry has resulted in a somewhat overwhelming choice of models to calculate magma storage conditions. Concerningly, these different models can return vastly different results. Wieser et al. (2023b) show that different Cpx-Liq and Cpx-only thermobarometers applied to the same Cpx-Liq pair passing equilibrium tests yield pressures spanning >10 kbar and temperatures spanning >100 °C. Clearly, model choice is one of the largest sources of uncertainty when performing such calculations.

Melt inclusion barometry uses volatile solubility models to calculate the pressure at which a specified melt composition (major + volatile elements) is volatile saturated at a specified temperature. The formulation of these solubility models ranges from being purely empirical (Liu et al., 2005; Shishkina et al., 2014) to purely thermodynamic (Ghiorso and Gualda, 2015; Papale et al., 2006). Some models lie between these end members, having a general form indicated by thermodynamics and some empirical coefficients accounting for silicate melt composition (semi-empirical, e.g., Dixon, 1997; Iacono-Marziano et al., 2012). Even for fully thermodynamical models, the sign and magnitude of many coefficients attached to thermodynamic terms have been criticized as being physically implausible (see Ghiorso and Gualda, 2015 and Wieser et al., 2022b for discussion of the Papale et al., 2006 model). In addition to their different functional forms, the calibration datasets of solubility models are also highly variable with respect to the range of melt compositions, fluid compositions, and P and T. For example, the model of Shishkina et al. (2014) was calibrated on variably alkaline, relatively mafic melts, and expresses volatile solubility empirically using the cation fractions of Ca, K, Na, Mg, Fe, Si and Al for CO_2, and K and Na for H_2O. In contrast, the solubility model of Dixon (1997) was calibrated on a more restricted range of relatively mafic tholeiitic to alkaline melt compositions. This model has a thermodynamic form, with an empirical correction for the effect of melt composition that is only dependent on the concentration of SiO_2 in the melt. The model MagmaSat (Ghiorso and Gualda, 2015) was calibrated on a very wide range of melt compositions (tholeiite and alkaline, from basalts to rhyolites), and is a fully thermodynamic model sensitive to all the commonly measured major oxide species. Unsurprisingly, given these differences in model formulation and calibration datasets, calculated saturation pressures can vary greatly between models, with systematic offsets of at least a factor of two not being uncommon (Wieser et al., 2022b).

In stark contrast to uncertainties associated with model choice when performing mineral thermobarometry or melt inclusion barometry, fluid inclusion barometry (Section "Fluid inclusion barometry") relies on the CO_2 equation of state to convert the measured density of a CO_2-rich fluid into a P for a specified entrapment temperature. Other than an ideal gas law (which does a poor job at high P), different published CO_2 equation of states predict very similar P for a given CO_2 density and T (~1–5% difference, Böttcher et al., 2012; Lamadrid et al., 2017; Span and Wagner, 1996; Sterner and Pitzer, 1994; Wieser and DeVitre, 2023). For example, at 1150 °C for $\rho_{CO_2} = 0.8$ g/cm^3, the relatively simple empirical expression of Sterner and Pitzer (1994) gives 5.008 kbar, while the more complex thermodynamic model of Span and Wagner (1996) gives 4.956 kbar (~3% difference). However, the conversion between measured CO_2 density to depth is substantially less simple when trapped fluids contain other species, such as H_2O or SO_2 (Hansteen and Klugel, 2008; Hurai, 2010). It is difficult to estimate the initial molar ratios of each species, and EOS for these mixed fluids are poorly constrained, if parameterized at all. Fluid inclusion barometry is also not immune to systematic error; re-equilibrium of the host crystal during magma ascent can increase the inclusion volume, reduce the CO_2 density and thus the calculated pressure/depth (Hansteen and Klugel, 2008; Wanamaker and Evans, 1989).

Thermodynamic approaches to determining the conditions of magma storage and evolution are also sensitive to the choice of model (Section "Thermobarometers based on thermodynamic modeling"). Thermodynamic models are typically constructed from two key components: (i) a dataset of standard state properties (e.g., enthalpy of formation, heat capacity, etc.) for all minerals of interest, and (ii) solution models for phases with variable compositions, which describe how composition influences the thermodynamic properties of the phase. Within igneous petrology and volcanology the MELTS "family" of thermodynamic models is the most widely used (Ghiorso et al., 2002; Ghiorso and Sack, 1995; Gualda et al., 2012). The original MELTS model was developed by Ghiorso and Sack (1995), building on the thermodynamic database of Berman (1988) to include thermodynamic models of relevant igneous solid solutions (e.g., Ghiorso, 1990; Ghiorso and Sack, 1991; Sack and Ghiorso, 1994, 1991, 1989). Since the release of this original MELTS model, various updates have been published. pMELTS includes a revised liquid thermodynamic model optimized for mantle-like bulk compositions (Ghiorso et al., 2002). Rhyolite-MELTS v.1.0.2 incorporates changes to the thermodynamic properties of quartz and the potassium endmember of the alkali feldspars to enable the eutectic behavior of high-silica rhyolitic magmas to be recreated (i.e., crystallization over a narrow T range, Gualda et al., 2012). Rhyolite-MELTS v1.2.0 incorporates the mixed H_2O-CO_2 fluid model MagmaSat and is recommended for use away from the granitic ternary minimum. Rhyolite-MELTSv1.1.0 incorporates the updated CO_2 solubility model, but retains the old H_2O solubility model of rhyolite-MELTSv1.0.2 for calculations at the ternary minimum (Ghiorso and Gualda, 2015).

An alternative group of thermodynamic models developed by Tim Holland, Roger Powell and co-workers (Holland and Powell, 2004, 1998, 1990) have traditionally been utilized in metamorphic studies (Tamblyn et al., 2020), but have recently been updated for applications to mantle melting and igneous systems. For example, Jennings and Holland (2015) expand the model system, optimizing its performance for calculations of peridotite melting behavior and the phase relationships of basaltic liquids at crustal to mantle conditions. Most recently, Holland et al. (2018) presented an updated thermodynamic database that is calibrated on a range of compositions, from peridotites through to granites. The Holland dataset can be accessed through a variety of software tools. THERMOCALC calculates the location of known phase boundaries and mineral reactions (Powell et al., 1998), while Perple_X (Connolly, 2009, 2005) and Theriak-Domino (de Capitani and Petrakakis, 2010) use a Gibbs Free Energy minimization approach to calculate the phase assemblage and compositions specified P-T conditions.

Comparisons between the MELTS and Holland-Powell families of models are relatively uncommon; Jennings and Holland (2015) compare the results of mantle melting calculations, and Hernández-Uribe et al. (2022) compare the equilibrium crystallization behavior of an mid oceanic ridge basalt (MORB) magma. Importantly, Hernández-Uribe et al. (2022) do not compare fractional crystallization pathways, noting that these would be "relatively laborious" with current software tools. The recent release of MAGEMin, a Gibbs Free Energy minimization package utilizing the Holland et al. (2018) thermodynamic models written in the programming language C (Riel et al., 2022) allows MELTS-like workflows to be performed including fractional crystallization. MAGEMin has a Julia interface and can be run in Python3 using PetThermoTools (Gleeson and Wieser, 2023), greatly aiding model intercomparison between the Holland and MELTS databases (see Section "Thermobarometers based on thermodynamic modeling"). There are also several more empirical models which have been used as alternatives to MELTS to model fractional crystallization in magmas, and therefore could place constraints on P and T (see Section "Thermobarometers based on thermodynamic modeling"). These include COMAGMAT (Ariskin et al., 1993) and Petrolog3 (Danyushevsky and Plechov, 2011). In Section "Thermobarometers based on thermodynamic modeling," we show that the choice of thermodynamic model to use (and even the version of MELTS used) has a large influence on calculations of magma storage conditions (see also Hernández-Uribe et al., 2022).

Statistics to compare models

When trying to decide which model/equation to use to calculate magma storage conditions, and when assessing whether the chosen method has sufficient resolution to address the science question of interest, it is important to consider both accuracy and precision. Accuracy describes how close the measurement/calculation is to the true value, while precision describes how close repeated measurements/calculations are to one another. For example, if you measure a small, very homogenous region of a Cpx crystal five times using an electron probe microanalyser (EPMA), and calculate Cpx-only pressures from these measurements, you may obtain 5, 5.5, 6, 4.9, and 6.2 kbar. The precision could be quantified using 1 standard deviation of these measurements (e.g., ±0.47 kbar). However, the Cpx may have formed at 8 kbar, in which case these calculations are relatively precise, but inaccurate (mean offset of 2.5 kbar). It is also worth distinguishing between random and systematic errors. Random error describes scatter about the true value (affecting precision, not accuracy), while systematic error describes a constant offset from the true value (affecting accuracy not precision).

Most publications calibrating new mineral-melt thermobarometers or chemometers describe the fit between calculated and experimental values for a given parameter using the root mean square error (RMSE) and the R^2 value (see Table 1, RMSE = standard error estimate, SEE, for linear regressions). When equations are applied to natural systems, these RMSE errors are often quoted as the error on the calculation. This is problematic for several reasons. The RMSE and R^2 alone do not properly distinguish between random and systematic error, so fails to capture different model performances. Many different metrics (Table 1) can be useful to assess model performance.

In Fig. 2, we demonstrate the issues associated with using *only* the root mean square error (RMSE) and R^2 to assess model performance. Instead, we suggest that the RMSE, R^2, gradient, intercept and MBE should all be presented on figures to allow assessment of both random and systematic uncertainty (e.g., Putirka et al., 2003). To demonstrate this, we compare the relative performance of three hypothetical barometers, plotting the experimental P on the x axis, and the calculated P on the y axis (Fig. 2). All barometers have the same RMSE error. However, it is visually apparent that their performance varies greatly. While Fig. 2a has a very high R^2 value, its low gradient and high intercept reveals that it substantially overpredicts P at $P < 12$ kbar, and underpredicts P

Table 1 Metrics used to assess models, where x is the measured parameter (e.g., P, T) and y is the model-predicted parameter.

Metric	*Good at*	*Bad at*
R^2 value: Correlation coefficient of the linear regression between the experimental (X_{Exp}) and calculated (X_{Calc}) value	Assessing precision and random uncertainty. Low precision (lots of random uncertainty) = low R^2 value	Assessing accuracy and systematic uncertainty
Gradient and Intercept of the linear regression	Assessing systematic uncertainty (will generate a gradient different from 1, and an intercept different from zero)	Assessing precision/random uncertainty (as averages all measurements)
Root mean square error (RMSE) aka SEE $RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(X_{Calc} - X_{Exp})^2}$	Describes how concentrated the data is around the linear regression	Struggles to distinguish between low precision and accurate vs. high precision with a systematic offset. Sensitive to outliers
Mean Absolute Error (MAE) $MAE = \frac{1}{N}\sum_{i=1}^{N}\lvert X_{Calc} - X_{Exp}\rvert$	Similar to the RMSE but no squared term. Gives less weight to larger errors than the RMSE	Struggles to distinguish between low precision and accurate vs. high precision with a systematic offset
Mean Bias Error (MBE) $MBE = \frac{1}{N}\sum_{i=1}^{N}(X_{Calc} - X_{Exp})$	Identifies average model bias, as no absolute or squared term	Doesn't identify random error, as + and − errors cancel out

RMSE and SEE are identical for a linear regression, for higher order regressions, they differ as only the SEE accounts for the degrees of freedom.

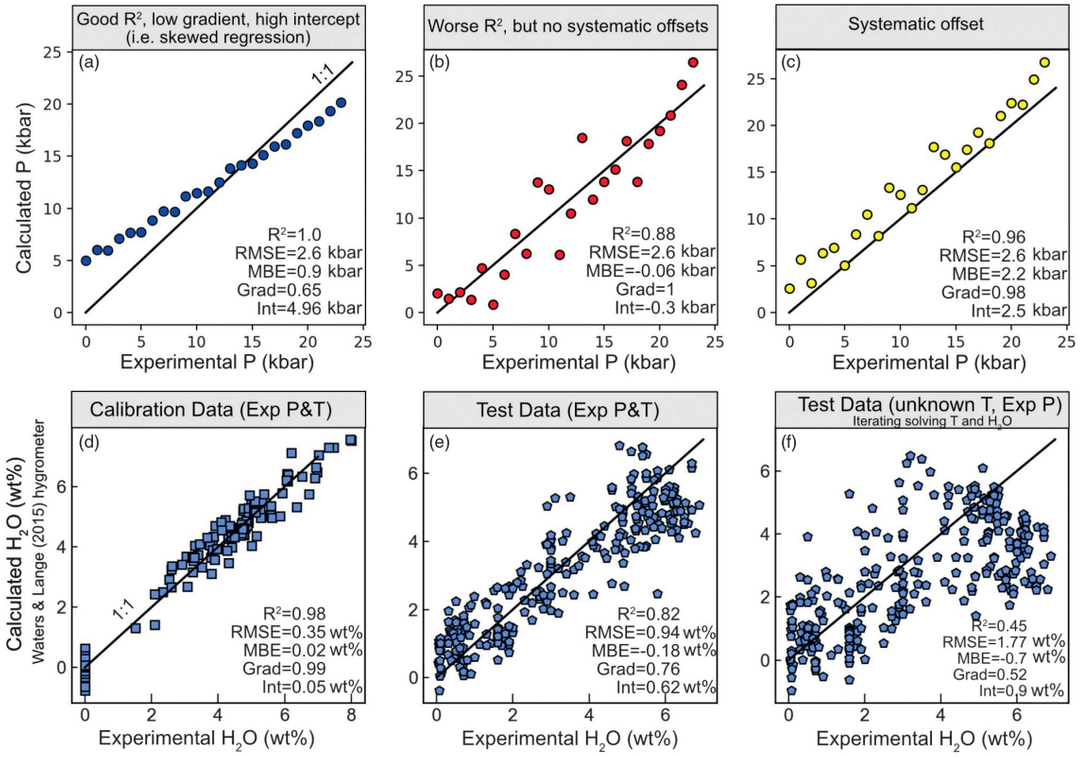

Fig. 2 Schematic figure showing the limitations of using just R^2 and RMSE to assess thermobarometers. (a–c) Three hypothetic barometers with the same RMSE, but vastly different performance. (d–f) Comparing experimental and calculated H_2O using the Waters and Lange (2015) hygrometer for: (d) the calibration dataset using experimental T and P, (e) the ArcPL dataset using experimental T and P, and (f) the ArcPL dataset iteratively solving H_2O and T (using Eq24a of P2008, see Section "Plagioclase-liquid thermometry and hygrometry" for more detail).

at $P > 12$ kbar. While Fig. 2b has a lower R^2 value, it has no substantial systematic offsets, meaning it is a far more accurate barometer across a wide range of pressures than that shown in Fig. 2a (indicated by the gradient close to one, an intercept close to zero, and the low MBE). The barometer in Fig. 2c also has a gradient close to 1, but the higher intercept and MBE indicate that it systematically overpredicts at all pressures. Clearly, if you wish to distinguish the absolute depth of magma storage, perhaps to compare to geophysical inversions of magma storage depths, or signals of unrest, the barometer in Fig. 2b is the best, despite its lower R^2. However, if you only wish to distinguish differences in storage pressures between different crystal populations (without caring about the actual depth), the barometer in Fig. 2a is the best, as it is the most precise. Use of all five metrics in tandem is essential for identifying systematic and random uncertainty.

It is also challenging to directly compare the quoted RMSE and R^2 values from different models, because these statistics are not calculated in an equivalent manner. Many of the statistics reported in the abstracts of papers proposing new models describe the fit to the calibration dataset (e.g., the commonly quoted ±1.7 kbar value for the Cpx-Liq barometer of Putirka et al., 2003; the ±0.35 wt% H_2O value for the Waters and Lange, 2015, Plag-Liq hygrometer). In other papers, the number reported in the abstract is the fit to a test dataset (e.g., Jorgenson et al., 2022; Petrelli et al., 2020), which inevitably makes the equation look "worse" than one assessed using calibration data. Indeed, it is apparent in the main text of many thermobarometry papers that the RMSE is much larger when applied to test data not used in the calibration dataset, or when applied to a global dataset containing a mix of new and calibration data (e.g., RMSE = 4.2 kbar for the Cpx-Liq Putirka et al., 2003 barometer vs. the commonly quoted RMSE = 1.7 kbar, Fig. 3a). The size of the test dataset can also vary greatly, influencing the statistics. For example, the test dataset of Petrelli et al. (2020) only contains 59 experiments conducted at <15 kbar, which come from only 4 studies. Some reported statistics are calculated using only experiments for a select subregion of compositional space (e.g., Neave and Putirka, 2017 for Cpx-Liq equilibria in tholeiites), while others report the fit for a global dataset with a much larger compositional range, again leading to larger apparent errors (e.g., Putirka, 2008). For many equations, a paucity of experimental data resulted in all data being used for calibration, leaving no test dataset (e.g., Waters and Lange, 2015 for Plag-Liq hygrometry, Ghiorso and Gualda, 2015 for volatile solubility). Additionally, some proportion of the error in any quoted statistic (e.g., RMSE) will represent uncertainty in the experimental data used for testing (both analytical and experimental sources of scatter), making it difficult to differentiate failures of the calibration equation vs. the data used to assess it. This is a particular problem for models where higher quality data was saved for calibration, with less strict filters applied to test data (Putirka et al., 2003; Putirka, 2008).

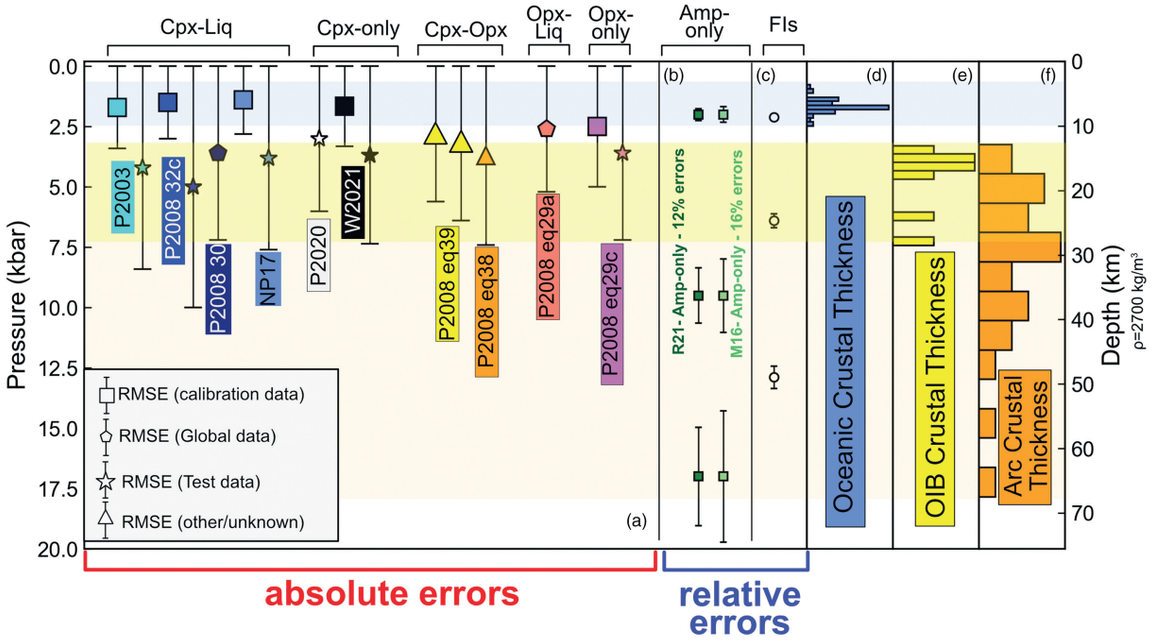

Fig. 3 Comparison of quoted errors on barometers to estimates of crustal thickness (assuming $\rho = 2700$ kg/m^3). (a) We anchor each method with an absolute uncertainty such that the upper part of the 1σ error bar sits at 0 kbar. This visualization shows the range of pressures/depths which cannot be statistically distinguished from storage at the surface. The symbol shape represents whether the quoted SEE/RMSE was for the calibration dataset, a test dataset, or global data. (b and c). Relative (%) uncertainties are shown for different pressures. For (c), we calculate FI errors assuming an uncertainty in CO_2 density of $\pm$0.01 g/cm^3 and $\pm$50 K for entrapment T. (d) Oceanic crustal thickness compilation from Chen (1992), (e) OIB thicknesses compiled for this study (see supporting information), (f) Continental arcs from Profeta et al. (2016). Errors for melt inclusions are highly sensitive to bubble volumes, so a generic example cannot be given (see Fig. 24b). Abbreviations: **P2003**: Putirka et al. (2003), **P2008**: Putirka (2008), **NP17:** Neave and Putirka (2017), **P2020**: Petrelli et al., 2020, **W2021**: Wang et al. (2021), **R21**: Ridolfi (2021), **M16**: Mutch et al. (2016).

Ideally when calibrating any new equation/model (for thermobarometry, chemometry, volatile solubility modeling, etc.), data should be subdivided into a test dataset and a calibration dataset. It is then standard practice for many machine learning workflows to further subdivide the calibration dataset into train and validation datasets (this split may be made repeatedly during model training, e.g., Petrelli et al., 2020). The train-validation split allows investigation of the effect of adding/removing terms, changing the regression algorithm, and tweaking the regression tuning parameters. Only once the model is fully tweaked should it be applied to the test dataset to assess its performance on unseen data. While it is then tempting to continue to change parameters to improve the fit to the test dataset, this strategy is generally criticized in the world of machine learning, as it means the testing dataset has "leaked" into the training of the model, so no longer provides an independent assessment of the model validity. The complete isolation of a testing dataset prior to model tuning and choice of parameters is important for machine learning and other regression workflows, because otherwise it is difficult to assess model performance on samples which are distinct in P-T-X space from those used in model calibration (Lones, 2021). To address this, in some biomedical studies, a test dataset is kept completely separate, and held under "lock and key" by a honest broker to avoid such leakage (Dobbin and Simon, 2011; Shedden, 2008). In igneous petrology, the difficult balance to strike is between having a large test dataset to robustly assess model performance, versus having too small a calibration dataset to adequately capture variation in melt or mineral composition.

It is also worth noting that the RMSE value is an average for the fit across the entire P or T range, while the actual error when applied to natural samples will vary as a function of P and T. RMSE and R^2 values are also not always calculated across the same P and T range (e.g., 0–40 kbar for Putirka, 2008, 0–20 kbar for Petrelli et al., 2020). These pressure ranges are also often far larger than the pressure ranges of interest in volcanic systems, and models tend to perform worse at lower pressures (Putirka, 2008).

Finally, mineral-melt thermobarometers and chemometers are often assessed using experimental data, where T, P (and often H_2O and fO_2) are constrained. In natural systems, it is common that several intensive parameters are poorly constrained. For example, a T and H_2O-sensitive barometer is normally assessed using the test dataset, inputting the experimental temperature and the measured H_2O content in the experimental charge. Similarly, thermometers with P-sensitive terms are normally assessed using the experimental P, and hygrometers are assessed using experimental T and P. In natural systems, the most common scenario is that *both* P and T are unknown, so must be iteratively solved using a thermometer and a barometer. Thus, to estimate a realistic error when applied to natural systems, statistics should be calculated by iteratively solving the same variables that would be unknown in natural samples. Of course, this adds additional uncertainty; when applying a T-sensitive barometer, any deviation from the true T will affect the calculated P. Additional uncertainties related to parameters which cannot be iteratively solved (e.g., uncertainty in

H_2O in Cpx-Liq thermobarometers) should be propagated with Monte-Carlo techniques using a realistic uncertainty for H_2O in a specific system of interest (e.g., Wieser et al., 2022c).

Using the T-sensitive Plag-Liq hygrometer of Waters and Lange (2015) as an example, we demonstrate how the apparent performance of an equation can vary based on the number of constrained intensive parameters, and the dataset used to test it. First, we show H_2O contents for the model calibration dataset using experimental temperatures (Fig. 2d). This comparison yields the RMSE quoted in the abstract of the Waters and Lange (2015) paper that is commonly stated in the literature as the error on the method (±0.35 wt%). Next, we evaluate this hygrometer using a newly compiled test dataset of variably hydrous experiments at 0–17 kbar spanning from basalts to dacites (ArcPL, Wieser et al., 2023b,d). We only include experiments where H_2O was measured using quantitative methods (Fourier Transform infrared spectroscopy—FTIR, secondary ion mass spectrometry—SIMS, Raman, calibrated volatiles-by-difference, solubility laws, see Section "Plagioclase-liquid thermometry and hygrometry"), and discard experiments which were present in the Waters and Lange (2015) calibration dataset. Using experimental T to calculate H_2O for this new test dataset yields statistics that are noticeably worse than those calculated from the calibration dataset (e.g., RMSE of 0.94 wt% vs. 0.35 wt%, Fig. 2e). In a scenario where T is not known (i.e., when this method is applied to natural systems), the statistics decline further: iterating the hygrometer with the Plag-Liq thermometer of Putirka (2008, Eq24a) results in a RMSE of ±1.77 wt% (Fig. 2f, see Section "Plagioclase-liquid thermometry and hygrometry" for detailed discussion). Thus, simply quoting the uncertainty from the original publication may underestimate the error by a factor of ∼5 when applied to natural systems. We observe similarly large declines in model performance moving from calibration to test data to iterative solving for many mineral-melt equilibria (see Section "Mineral thermobarometry and chemometry," and Wieser et al., 2023b).

Thus, at present, it is not possible to select the best thermobarometer for a given application simply based on quoted statistics alone. Instead, to pick the most suitable model, it is worthwhile to check the natural compositions of interest vs. the calibration range of the model of interest (calibration datasets for many models are available in Thermobar, Wieser et al., 2022c). After identifying models calibrated on suitable compositions, we encourage authors to compile a test dataset of experimental compositions most similar to their system. Ideally these experiments would not have used during calibration of the model. If the planned workflow in the natural system will involve iterative calculations, these same iterative workflows should be applied to the test dataset. These tests will provide a more realistic estimate of the uncertainty associated with the calculation than simply quoting RMSE values. If there is very little data to help pick, it may be best to perform calculations using several different equations calibrated on relevant compositional and P-T ranges. An average for all these models could be calculated, with the difference between models giving insight into the error in the calculation. At the moment, only a very small proportion of studies compare the results from multiple equations (e.g., Erdmann et al., 2016; Geiger et al., 2018; Sas et al., 2017; Sheehan and Barclay, 2016 for Cpx-Liq equilibrium, Rasmussen et al., 2022; Wieser et al., 2021 for volatile solubility).

What accuracy and precision are required?

In the 2008 RiMG short course, Putirka described his Cpx barometer as: *"a chisel, not a pen knife"*. Once realistic estimates are made of the precision and accuracy of different thermobarometers using relevant test datasets, it is worth thinking about what scientific questions can be addressed with confidence within the associated uncertainties of each method. Petrological barometers calculate pressure, which can be converted into depth by making assumptions about crustal and upper mantle densities. Converting pressures to depths allows comparisons to geophysical inversions of magma storage locations within the crust (e.g., Rasmussen et al., 2022), and to field observations in exposed crustal and upper mantle sections. Thus, it is useful to think of errors in the context of crustal and lithospheric thicknesses worldwide.

Oceanic crust formed away from the influence of mantle plumes ranges in thickness from ∼5–8.5 km, with an average thickness of ∼7 ± 0.8 km (White et al., 1992). Assuming an average density of 2900–3000 kg/m^3 (Afonso et al., 2007), this corresponds to an average Moho pressure of ∼2 kbar. To be able to conclusively pull apart different crustal storage geometries (e.g., upper vs. lower crust) and distinguish between storage regions at different levels in the crust, a RMSE error of ±0.25–0.5 kbar would be needed. No mineral-based barometers applicable to MORB lavas achieve anything like this precision (Fig. 3a vs. d). In fact, even if magma is stored below the Moho (e.g., in slow spreading ridges, Bennett et al., 2019; Drignon et al., 2019), available mineral-melt models can't confidently differentiate upper mantle vs. crustal storage.

Crust in ocean island basalts (OIBs) is thicker than in MORB settings because the newer volcanic pile rests on top of older oceanic crust (total crustal thickness of 14–24 km in Hawai'i, Leahy et al., 2010; 14–15 km in La Palma, Ranero et al., 1995), and/or because the plume increases melting extents at the mid-oceanic ridge (e.g., ∼11–39 km in Iceland) (Menke, 1999). The elevated crustal thickness in OIBs, corresponding to pressures of 3–4 kbar or greater, means that mineral-based barometers with uncertainties of 2–3 kbar can begin to distinguish storage below the Moho vs. in the crust (e.g., Gleeson et al., 2021). Mineral-based barometers are also aided by the fact that OIB lithosphere can be extremely thick relative to MORB as a result of conductive cooling as the oceanic crust moves away from the ridge (e.g., 45–60 km thick in the Galápagos, Gibson and Geist, 2010; 50–110 km thick in the Hawaiian Islands, Li et al., 2004), and many OIB magmas are stored at sub Moho depths (e.g., Barker et al., 2021; DeVitre et al., 2023a; Gleeson et al., 2021), Again, this makes uncertainties of 2–3 kbar less problematic. It is only in volcanic arcs, and particularly continental arcs with thicker crust (Profeta et al., 2016) that mineral-melt barometers can reliably distinguish between storage in the upper, middle and lower crust.

While mineral-melt thermobarometers often show a reasonably constant error regardless of the pressure, solubility models show a clear increase in error at higher volatile contents, so are best described using percentage errors (e.g., $H_2O = \pm 10\%$ and $CO_2 = \pm 20\%$ for Shishkina et al., 2014, $H_2O = \pm 10\%$ and $CO_2 = \pm 17\%$ for Iacono-Marziano et al., 2012). Systematic offsets

between solubility models also tend to increase reasonably proportionally with increasing P (Wieser et al., 2022b). Even the large errors resulting from volume estimates during vapor bubble reconstructions when performing melt inclusion barometry (more detail in Section "Melt inclusion barometry") are percentage errors. Many uncertainties associated with fluid inclusion barometry scale with pressure (Fig. 3, see Section "Fluid inclusion barometry"). In general, methods with percentage errors that scale with the value of the quantity, rather than absolute errors, are better suited to distinguishing subtle variations in magma storage in relatively thin-crusted tectonic settings (Fig. 3b and c). Interesting, Ridolfi (2021) and Mutch et al. (2016) express the errors on their amphibole barometers as percent errors (e.g., 12–16%, see Fig. 3b), although, these errors only describe the fit to the calibration data, and their magnitude is highly debated (see Erdmann et al., 2016, and Section "Amphibole thermobarometry and chemometry").

Influence of analytical error on precision and accuracy

Measurement of any quantity in igneous systems is subject to analytical error. Mineral and melt compositions are typically measured using an electron microprobe (EPMA), which is associated with random and systematic uncertainties (see Wieser et al., 2023d for a more detailed discussion). Pressure estimates from melt inclusions rely on volatile measurements by FTIR or SIMS, EPMA measurements of the host and melt phase, Raman spectroscopy measurements of the vapor bubble, and estimates of the relative volume of the vapor bubble (see Section "Melt inclusion barometry" for a detailed discussion). Pressure estimates from fluid inclusions rely on measurements of the fluid using Raman Spectroscopy/Microthermometry, and an independent estimate of T using mineral or mineral-melt thermometry (e.g., using EPMA analyses). Thermodynamic methods of inverting liquid compositions or liquid lines of descent rely on EPMA or energy-dispersive spectroscopy (EDS) measurements of glasses, or whole-rock X-ray fluorescence (XRF) measurements.

As many of the methods discussed in this review rely on EPMA analysis, we briefly discuss the uncertainties relating to this method. The fundamentally random process of x-ray generation, and subsequent detection by spectrometers is a significant source of uncertainty, termed "counting statistics." The magnitude of this uncertainty depends on the concentration in the sample and the beam current and voltage (affecting x-ray production), and the analysis time and spectrometer efficiency (affecting x-ray detection). Most simply, unless long count times and higher beam currents are used, low concentration elements (<1 wt%) tend to be associated with relatively large percentage errors. When analytical errors on all measured oxides are propagated through thermobarometry equations using Monte-Carlo methods, they can generate a large range in calculated P and T.

For example, Wieser et al. (2023d) show that measurement of Na_2O in crustal Cpx using popular analytical conditions (e.g., 10 s, 10 nA) yields a 1σ error of ~10–40%. Most importantly, the magnitude of this error is often underestimated by factors of 2–4× because secondary standards with higher Na_2O concentrations than the sample are routinely used to assess precision. Secondary standards should only be used to assess accuracy, not precision, as precision of an EPMA analysis is very closely related to the concentration of the element of interest, so varies greatly depending on the individual mineral composition. The large imprecision of many Na_2O in Cpx measurements is problematic because this oxide is often used to calculate the abundance of the P-sensitive Jadeite component. Propagating typical analytical errors through popular expressions for Cpx-Liq thermobarometry can generate highly correlated PT arrays spanning 3–5 kbar, which could be incorrectly interpreted as transcrustal storage (Wieser et al., 2023d). Analytical errors also affect the experimental data used to calibrate thermobarometers. As many experimental studies perform <5 measurements on each phase, random analytical error does not get sufficiently averaged out, so the reported phase composition may not be the true phase composition (affecting model calibration, and calculated statistics on test data, see Wieser et al., 2023d, see also Section "Clinopyroxene ± liquid").

To visually demonstrate the effect of analytical precision on different mineral-melt thermobarometers, we perform Monte-Carlo simulations using Thermobar (Wieser et al., 2022c, Fig. 4). We use analytical errors estimated by the EPMA based on counting statistics for experimental phase compositions reported by Krawczynski et al. (2012). For each phase, we produce 500 synthetic compositions with each element normally distributed about a measured value in the experiment, with the standard deviation equal to that estimated by the EPMA. We also include ±5% relative error for H_2O in the liquid, which is a very conservative estimate of the error associated with FTIR and SIMS measurements. For two-phase equilibrium (e.g., Cpx-Liq), we consider all possible matches between the 500 synthetic compositions for each phase (250,000 pairs total). We then perform thermobarometry calculations on all synthetic compositions, showing the spread of calculated P and T relative to the median of the distribution (Fig. 4).

In general, the influence of analytical error on calculated T is relatively minor ($1\sigma < 20\,^{\circ}C$). The broadest distribution of calculated temperatures comes from equations that are very sensitive to melt H_2O content, such as Liq-only T from P2008 Eq14 and Eq22 (Fig. 4a), Opx-Liq T from P2008 Eq28a (Fig. 4b), and Cpx-Liq T from P2008 Eq30 (Fig. 4e). For calculated P, the spread of simulations is highly variable, with barometry methods sensitive to a component with a relatively low concentration showing a very wide spread of pressure (e.g., Na in Cpx, Fig. 4j). In contrast, barometers sensitive to high concentration elements show very narrow distributions (e.g., Al in Amphibole, Ridolfi, 2021; Médard and Le Pennec, 2022, Fig. 4h).

We also test the OPAM liquid barometer, where pressures are calculated from the composition of melts co-saturated in olivine, plagioclase, and augite (see Section "Liquid-only barometry"). As the experiments of Krawczynski et al. (2012) are too H_2O-rich for OPAM, we use estimates of analytical precision for measurements of submarine basaltic glasses from the Galápagos Spreading Centre (Gleeson and Gibson, 2021). Notably, the analytical errors have a very similar magnitude to those from Krawczynski et al. (2012). This method is sensitive to analytical uncertainty, returning a similar range of calculated pressures to Cpx-based mineral-based barometers. The model of Voigt et al. (2017) is particularly sensitive, because of the inclusion of a term for the Cr content of the liquid, which has a low concentration, so is associated with poor analytical precision.

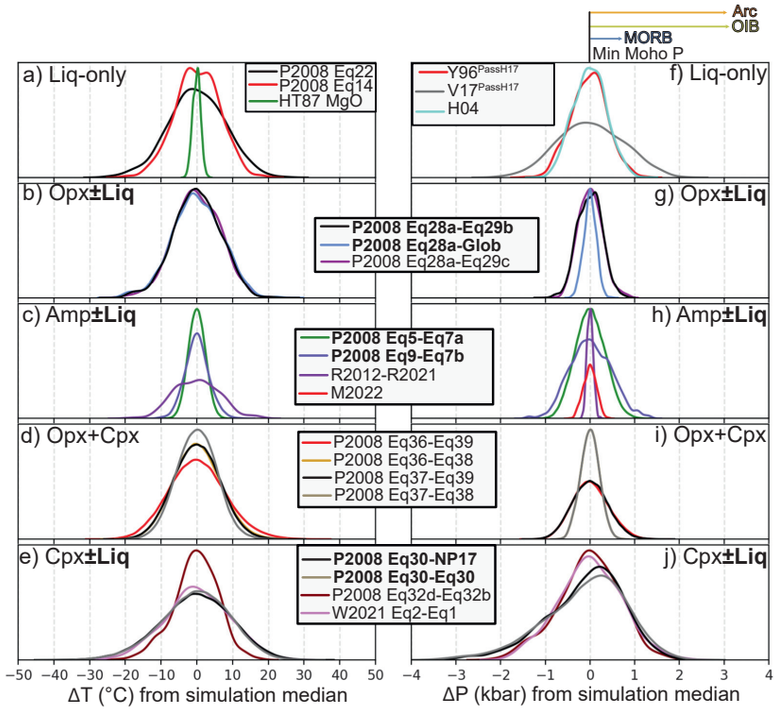

Fig. 4 Monte Carlo simulations of the influence of analytical precision on thermometry (a–e) and barometry (f–j) calculations. All panels except f show analytical uncertainties from EPMA counting statistics from Krawczynski et al. (2012) with H_2O of ±5%. As OPAM (panel f) is not applied to arc magmas, we propagate EPMA errors from natural Galápagos glasses from Gleeson and Gibson (2021). The bars at the top show approximate minimum Moho pressures from each tectonic setting for comparison (see also Fig. 3). Abbreviations: **P2008**: Putirka (2008), **HT87**: Helz and Thornber (1987), **R2021**: Ridolfi (2021), **R2012**: Ridolfi and Renzulli (2012), **M2022**: Médard and Le Pennec (2022), **W2021**: Wang et al. (2021), **Y96**: Yang et al. (1996), **V17**: Voigt et al. (2017), **H04**: Herzberg (2004).

Overall, the effect of propagated analytical error on P and T depends greatly on the selected equation (e.g., Fig. 4f and i), as well as the analytical conditions used. Thus, studies should propagate EPMA-estimates of counting statistic errors from each analysis not secondary standards using their chosen equation. Analytical errors should be propagated for all thermobarometric techniques, including melt inclusion and fluid inclusion barometry (see Section "Can hygrometers be used as barometers?" for an example). If propagated analytical errors are of similar magnitude to the resolution required to investigate the geological process of interest, time must be spent optimizing analysis (See Section "Clinopyroxene ± liquid"). Counting statistics are of course, only one source of analytical uncertainty. There are also notable offsets between different EPMA laboratories (e.g., Wieser et al., 2023b,d), and even between different operators on the same EPMA (see Kohn and Spear, 1991). These systematic uncertainties will increase the spread in calculated P and T for a given liquid, or crystal vs. that shown on Fig. 4.

Mineral thermobarometry and chemometry

The seminal review of Putirka (2008, hereafter P2008) summarizes several decades of thermobarometric work, and proposes a large number of new empirical expressions linking the compositions of Cpx ± Liq, Opx ± Liq, Cpx + Opx, Plag ± Liq, and Plag + Kspar to P and T. These new equations have similar forms to those published in older papers, but have recalibrated coefficients, and additional terms to improve performance when applied to the large compilation of experiments published in 2008 (LEPR; Library of Experimental Phase Relationships, Hirschmann et al., 2008). In the same issue, Anderson et al. (2008) summarized available P-T constraints on granitic rocks. To avoid repetition, we largely focus on new methods developed since 2008, as well as assessing the performance of these 2008 models using new experiments published in the last 15 years. Specifically, we assess models using a new test dataset of experiments conducted on arc magma compositions (ArcPL—Arc post LEPR, Wieser et al., 2023b,d) at variably hydrous (H_2O = 0–16.6 wt%, median = 4.7 wt%) crustal conditions (0–17 kbar). Importantly, the experiments in this new dataset were not used during the calibration of most models, and any overlaps were removed when testing a specific model. Our focus on arcs is partially a result of the experimental data available not used in model calibration, but also reflects the fact there has been very little focus on this tectonic setting, despite the fact it is one of the few places where the crust is thick enough that mineral-based

barometers may be able to resolve different crustal storage regions within uncertainty (Fig. 3). For methods not applicable to arc magmas (e.g., OPAM, Section "Liquid-only barometry"), we compile other available experiments. For many methods, there are not sufficient "test" experiments available for MORB, OIB and alkali lavas to build on the comparisons presented by Masotta et al. (2013), Masotta and Mollo (2019), Neave et al. (2019) and Neave and Putirka (2017), so we refer readers to these papers.

Liquid-only thermometry

A huge variety of liquid-only thermometers exist, with varying complexity and calibration ranges. For example, the thermometer of Helz and Thornber (1987) uses just the MgO content of the liquid, calibrated on compositions from the Kīlauea Iki lava lake. In contrast, P2008 eq14 has terms for the MgO, FeO_t, Na_2O, K_2O, H_2O and Mg# in the liquid, and was calibrated/tested on 1536 olivine-saturated liquids. There are also a number of olivine-liquid thermometers with a term for the partitioning of Mg between olivine and liquid (D_{Mg}^{Ol-Liq}). In their supporting spreadsheet, Putirka (2008) replace this term with the theoretical value of D_{Mg} predicted by Beattie (1993) to produce a liquid-only thermometer. Here, we discuss the best-performing thermometers for our new test dataset.

As P and H_2O may be poorly constrained in many systems, we first assess the sensitivity of various thermometers to these parameters. We randomly select 40 experiments—for each liquid composition, we make 100 duplicates with the experimental H_2O content perturbed by ±3 wt% (Fig. 5a), and separately perturb the experimental P by ±5 kbar (Fig. 5b). We take the calculated T at these perturbed values and subtract the calculated T at the measured P and H_2O (ΔT, ΔP). Thermometers with a H_2O term show variable, but relatively high sensitivity to H_2O, with the average calculated T dropping by ~9–16 °C per wt% H_2O added (Fig. 5a). For comparison, we show the quoted RMSE of each thermometer. Uncertainty in H_2O of 3 wt% introduces a systematic error comparable in size to the stated RMSE (Fig. 5a). Thermometers are substantially less sensitive to P; uncertainty in P of 5 kbar results in <30 °C of variation. Even a full 10 kbar of uncertainty in P doesn't exceed the quoted RMSE (Fig. 5b).

Given the sensitivity of calculated temperatures to H_2O, we only use experiments where H_2O was determined using quantitative methods (calibrated volatile-by-difference, FTIR, SIMS, Raman spectroscopy). We perform calculations using experimental P and measured H_2O, because there are no suitable liquid barometers or hygrometers to iteratively solve all unknowns. For the ArcPL dataset, the best liquid-only thermometer is an adapted version of the Ol-Liq thermometer of P2008 Eq22, where the olivine D_{Mg} term is replaced with the calculated D_{Mg} value from Beattie (1993), converting it into a Liq-only thermometer (Fig. 5c). Surprisingly, this adapted Ol-Liq thermometer does a very good job of predicting T in experiments without olivine, including melts with high SiO_2 contents (square symbols).

We also calculate the liquidus T using rhyolite-MELTS v1.2.0, v.1.0.2 and pMELTS (Ghiorso and Gualda, 2015; Gualda et al., 2012; Ghiorso et al., 2002, Fig. 5d,e, Supporting Fig. 2). All three versions of MELTS give similar statistics; pMELTS has the higher R^2 and lowest MBE and RMSE ($R^2 = 0.89$, MBE = 45 °C, RMSE = 58 °C, Fig. 5d), but shows worse performance than v1.0.2 for the lowest T liquids (Fig. 5e). It is notable that the MBE for all 3 versions is relatively high (44–56 °C vs. –0.12 °C for eq22DMg), which demonstrates that MELTS consistently overestimates liquidus temperatures (Fig. 5d,e). The other disadvantage of MELTS calculations vs. empirical approaches is the additional computational expense; 2000 liquidus calculations take 47 min using PetThermoTools through alphaMELTS for Python compared with 0.015 s using eq22DMg through Thermobar on a desktop computer with 128 Gb RAM and a 10–core 10th generation Intel i-9 processor. Liquidus temperatures can also be calculated using the Holland et al. (2018) thermodynamic model, implemented through the Gibbs Free Energy minimization software MAGEMin (Riel et al., 2022). The fit between the experimental and calculated T is considerably worse than that achieved by rhyolite-MELTS v1.2.0, with temperatures often overestimated by 100–200 °C, particularly for experiments performed at <1100 °C (Fig. 5f). These calculations take similar times to MELTS calculations. In summary, if melt H_2O contents are well constrained, empirical liquid-only thermometers such as eq22DMg of Putirka (2008) perform surprisingly well across a surprisingly wide range of melt compositions.

Liquid-only barometry

Barometric methods based on the composition of silicate melts rely on the fact that the thermodynamic variance of the system (i.e., the degrees of freedom) is lower in multi-phase saturated systems, so the melt composition contains information about the conditions of magma storage. This is typically discussed with regards to magma evolution along mineral cotectics, or at invariant points such as eutectics, where the degrees of freedom is close to zero (i.e., high-silica rhyolites, Ludden, 1978; Vogt, 1931). The positions of these cotectics and eutectics are sensitive to pressure because of differences in the volume of common igneous minerals. Thus, if the influence of P and melt composition on the location of cotectics/eutectics can be determined by thermodynamic or empirical parameterizations, measured compositions of lavas saturated in the required phases can be used to calculate P (e.g., Grove et al., 1992; Gualda et al., 2012; Wilke et al., 2017; Yang et al., 1996, Fig. 1g).

Thermodynamic approaches to liquid-only barometry have become increasing common in the last decade, with the rhyolite-MELTS thermodynamic models used to assess the storage pressure of evolved, rhyolitic magmas from several locations worldwide (Gualda and Ghiorso, 2014; Harmon et al., 2018; Pamukcu et al., 2015). The application of thermodynamic methods to liquid-only barometry is addressed in detail in Section "Thermobarometers based on thermodynamic modeling." Here, we instead focus on empirical methods calibrated using experimental data.

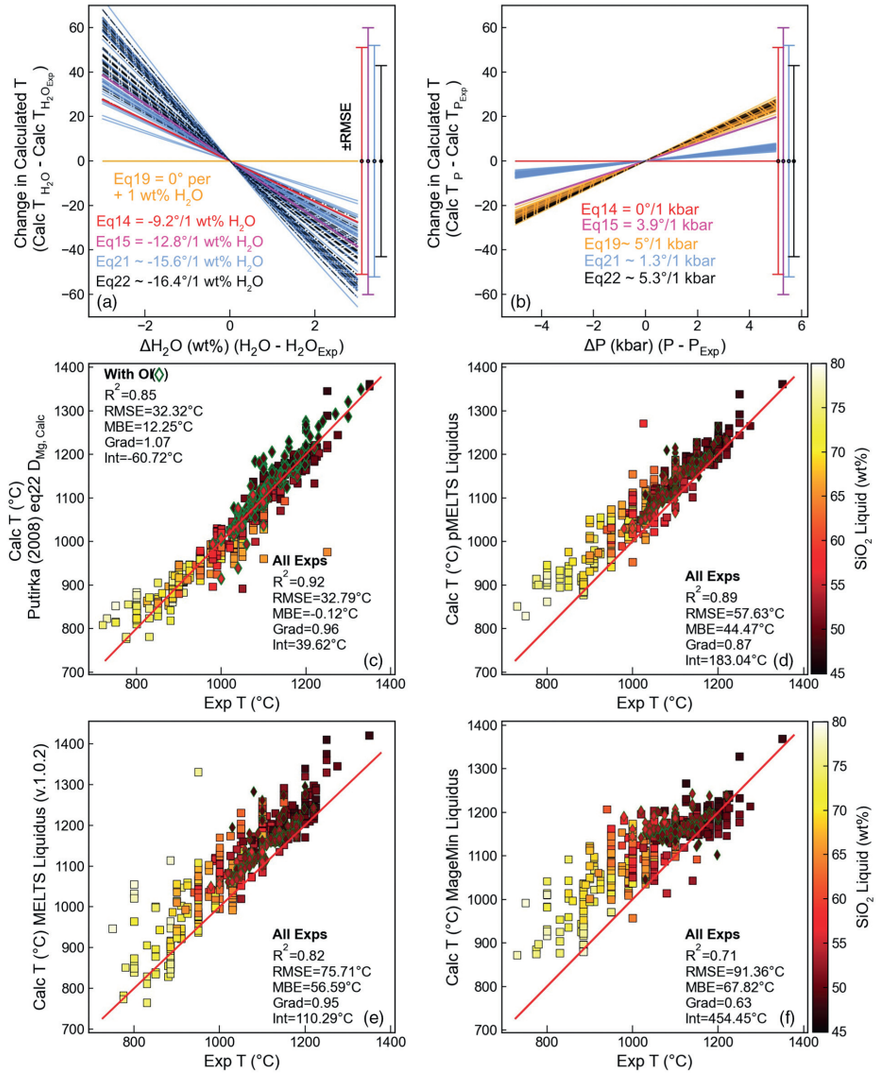

Fig. 5 Assessing Liquid-only thermometers. (a and b) Sensitivity of calculated T to H_2O (a) and pressure (b). For 40 randomly selected experiments (each represented by a line), we pertur the experimental H_2O content by $\pm$3 wt% (a), and pressure by $\pm$5 kbar (b). These plots show that increasing H_2O causes a drop in calculated T, and an increase in P causes a smaller increase in T. (c–f) Experimental T vs. calculated T for experiments where H_2O was measured by quantitative methods (FTIR, SIMS, Raman, calibrated volatiles by difference). Experiments containing olivine are shown as green-outlined diamonds, and experiments without olivine as black-outlined squares. All symbols are colored by the SiO_2 content of the melt. MELTS (part e) and MAGEMin (part f; using the (Holland et al., 2018) thermodynamic model) liquidus T calculated using $Fe^{3+}/Fe_T = 0.15$ (calculated T are not very sensitive to Fe^{3+} but having $Fe^{3+} > 0$ is necessary for the algorithm to converge). Calculations performed using MAGEMin_C v1.3.2.

OPAM barometry

Experimental work on the location of the olivine-plagioclase-augite-melt (OPAM) cotectic has been ongoing for at least three decades. Early work constrained the position of the OPAM cotectic at 0.001 kbar in the simplified CaO-MgO-Al_2O_3–SiO_2 (CMAS) + FeO (Shi, 1992) and CMAS + FeO + Na_2O systems (Shi, 1993). Further experiments on natural MORB compositions demonstrated that minor components in the liquid phase (e.g., Ti, Na, and K) and the expansion of the clinopyroxene stability field at higher pressures strongly influences the cotectic position (Grove et al., 1992; Yang et al., 1996). Using these observations, Grove et al. (1992) provided a set of empirical equations to locate the position of the olivine + plagioclase + melt and OPAM cotectics as a function of melt chemistry and pressure. Yang et al. (1996, hereafter Y96) built on this work with an updated parameterization for the location of the OPAM cotectic, expressing the Mg, Ca, and Al molar fractions (X_{Mg}, X_{Ca}, and X_{Al}) of an OPAM-saturated melt as a function of P and the remaining molar fractions (e.g., X_{Si}, X_{Ti}). For example, their expression for the molar fraction of Ca has the following form:

$$X_{Ca} = 1.133 - 0.00339 * P(kbar) - 0.569 * X_{Na} - 0.776 * X_K - 0.672 * X_{Ti} - 0.214 * X_{Fe} - 3.355 * X_{Si} + 2.830 * {X_{Si}}^2 \quad (4)$$

These equations were updated by Voigt et al. (2017, hereafter V17), who used new experimental data to show that Cr_2O_3 has a strong influence on clinopyroxene phase stability. Their equations have the same functional form, with an addition of a term for the molar fraction of Cr, separate terms for ferrous and ferric Fe molar fractions, and updated coefficients for the other parameters:

$$X_{Ca} = 1.07 - 0.02707 * P(GPa) - 0.634 * X_{Na} - 0.618 * X_K - 0.515 * X_{Ti} - 0.188 * X_{Fe^{2+}} - 0.597 * X_{Fe^{3+}} - 3.044 * X_{Si} + 2.477 * {X_{Si}}^2 - 9.367 * X_{Cr} \quad (5)$$

Herzberg (2004) (hereafter H04), develop an alternative approach, relating P to liquid components projected onto an Anorthite, Diopside, Enstatite ternary diagram from Olivine:

$$P(GPa) = -9.1168 + 0.2446 * (0.4645 * En + An) - 0.001368 * (0.4645 * En + An)^2 \quad (6)$$

Despite the greater simplicity of pressure calculations using the H04 OPAM parameterization, the Y96 and V17 molar fraction equations have formed the basis of most OPAM barometry over the last decade (V17: Bell et al., 2021; Stock et al., 2018, Y96: Baxter et al., 2023; Caracciolo et al., 2022, 2020; Halldórsson et al., 2022). The methods used to extract pressure information from these expressions vary. Early studies used the equations of Y96 to calculate the position of the OPAM cotectic at multiple pressures and plotted these positions on an olivine-clinopyroxene-quartz pseudoternary diagram (projected from plagioclase). Visual comparison of the position of the calculated OPAM cotectic to natural melt compositions projected onto the same pseudoternary diagram were used to estimate the pressure of magma storage (Geist et al., 1998; Maclennan et al., 2001). Kelley and Barton (2008) use a complex approach calculating theoretical and measured normative mineral components at different pressures, and regressing these components against pressure to determine the best fit for each measured sample. More recent studies have determined the pressure of the minimum misfit between calculated X_{Mg}, X_{Ca}, and X_{Al} using the equations of Y96 or V17, and measured X_{Mg}, X_{Ca}, and X_{Al} in their samples without requiring the conversion to normative mineral components (Hartley et al., 2018).

Regardless of the calculation method used to obtain P, OPAM barometry requires input melt compositions to be saturated in all three solid phases (olivine, clinopyroxene, and plagioclase). Solving for pressure using the equations of Y96 or V17 for samples that are not saturated in all three-phases could lead to erroneous results and may explain the wide range in calculated P estimated for some systems (e.g., 1.4–7.7 kbar at Laki in Iceland; Kelley and Barton, 2008). Therefore, it is important to determine which samples are co-saturated before interpreting OPAM calculations. One way to ensure three-phase saturation in the target melts is via textural observations, such as petrographic observations of euhedral crystals with no disequilibrium textures. Alternatively, in systems with abundant whole-rock or matrix glass data spanning a range of MgO contents, comparing liquid lines of descent with mass balance approaches or fractional crystallization models (e.g., MELTS, Petrolog3, Danyushevsky and Plechov, 2011; Gualda et al., 2012) can be used to confirm three-phase saturation.

An approach that can be applied to individual liquid compositions without textural or LLD context to calculate P and check for cosaturation was developed by Hartley et al. (2018). They determined the probability of three-phase saturation using a Chi-squared distribution to calculate the misfit between the predicted (X_i^{Pred}) and measured (X_i^{Meas}) molar fractions for Mg, Ca and Al:

$$\chi^2 = \sum_{i=1}^{3} \left[\frac{X_i^{Meas} - X_i^{Pred}}{\frac{\sigma_{X_i}}{100} X_i^{Meas}} \right]^2 \quad (7)$$

Here, σ_{X_i} is an estimate of the relative analytical uncertainty for each cation under consideration (expressed as a percentage). As X_i^{Pred} is a function of the system pressure, users can determine the best-fit model pressure by locating the position of the minimum possible χ^2 value ($\chi^2{}_{Min}$). To assess whether the melt is truly three-phase saturated, Hartley et al. (2018) compare $\chi^2{}_{Min}$ to the cumulative distribution function of the Chi-squared distribution with two degrees of freedom. A probability is selected as a cut off value (e.g., $P = 0.05$, $P = 0.2$); Hartley et al. use a probability of fit, P_f, which can be thought of as 1 minus the P value. If $\chi^2{}_{Min}$ is less than the critical value of the chi squared cumulative distribution function for this chosen cut off (i.e., P-value and $n = 2$ degrees of freedom), the liquid is assumed to be three-phase saturated. By considering the results of OPAM calculations on experimental melt compositions (including the Yang et al., 1996 calibration data), Hartley et al. (2018) determine that a "probability of fit" filter of $P \sim 0.2$ ($P_f = 0.8$) results in a good fit between the experimental and calculated pressures, while minimizing the number of "false negatives" (i.e., melt compositions which are OPAM saturated, but would fail this test). The minimization method and co-saturation check of Hartley et al. (2018) (hereafter H18) has been utilized by several studies investigating the storage conditions of MORBs and OIBs (Bell et al., 2021; Stock et al., 2018; Baxter et al., 2023; Caracciolo et al., 2022, 2020; Halldórsson et al., 2022).

Baxter et al. (2023) evaluated the performance of the Y96 equations using 315 OPAM saturated experimental melt compositions, ranging from 1 atm to 12.5 kbar pressure. They filter these experiments using the H18 filter and exclude those with MgO < 4 wt%, leaving 95 experiments. Comparing calculated and experimental pressures for these filtered experiments yields a RMSE of ±1.51 kbar, which is similar to the "random error" of ±1.32 kbar estimated by Hartley et al. (2018, definition uncertain). However, both studies evaluated the performance of the Y96 equations using a dataset which included some experiments used in the Y96 calibration. For example, 57% of the 95 experiments assessed by Baxter et al. (2023) are part of the Y96 calibration dataset. If we rerun their comparison using only experimental data that was not used in the Y96 calibration the RMSE increases from

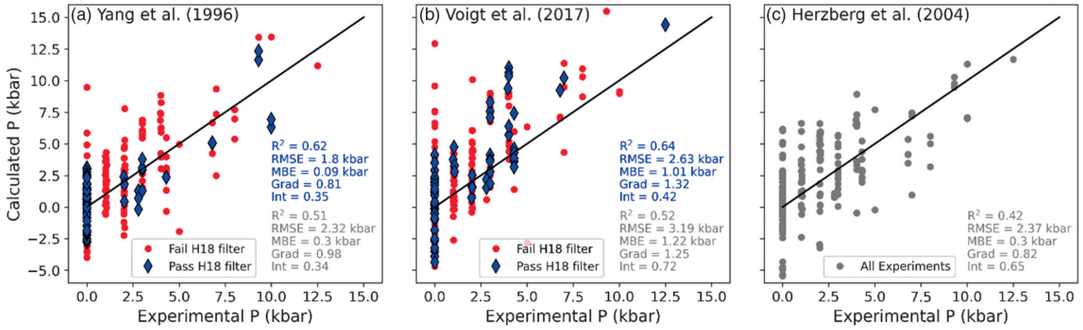

Fig. 6 Experimental vs. calculated P using the OPAM expressions of Yang et al. (1996), Voigt et al. (2017) and Herzberg (2004) for OPAM-saturated experiments on MORB-like compositions. The calculated pressure was determined using the Chi-squared method of Hartley et al. (2018, H18) for the Yang et al. (1996) and Voigt et al. (2017) equations (a,b). Samples passing the co-saturation filter of Hartley et al. (2018) are shown as blue diamonds, while those that fail are shown as red circles. The filter is not applicable to the method of Herzberg (2004, part c). Statistics using all experiments are shown in gray text, statistics for experiments passing the H18 filter are shown in blue text.

±1.51 kbar to ±1.89 kbar. A particular issue is that only 7/41 of the remaining experiments were conducted at >0.001 kbar, making it very hard to assess the performance of the Y96 barometer without new test data at higher pressures.

To test the performance of all three OPAM barometers listed here (Y96, V17, and H04), and the success of the H18 filter, we compile an experimental dataset of OPAM saturated experiments not used to calibrate either Y96 or V17. $N = 188$ overlap with the dataset of Baxter et al. (2023), with $N = 38$ experiments not in that compilation. Fig. 6 demonstrates that there are strong correlations between the calculated and experimental pressures for all OPAM methods when the H18 saturation check is not applied (R^2 of 0.51, 0.52, and 0.42 for Y96, V17, and H04 respectively, all symbols, gray statistics on Fig. 6). Despite having the highest R^2 value of the three methods tested here, calculated pressures determined using the V17 expressions display the largest uncertainties (RMSE = 3.19 kbar) and a systematic offset to high pressures (MBE = 1.22 kbar, Fig. 6b). The Y96 expression has a smaller systematic offset (gradient = 0.98, MBE = 0.3 kbar), as well as a lower RMSE (2.32 kbar, Fig. 6a). The parameterization of H04 also returns a relatively low RMSE (2.37 kbar) and similar MBE to the Y96 method (0.3 kbar), but has a lower gradient of the correlation (0.82; Fig. 6c).

The larger uncertainty and systematic offset in pressures estimated using V17 likely originates from the sensitivity of these expressions to the Cr_2O_3 content of the melt phase. Only 53 of the 225 experiments used in this analysis (~24%) report melt Cr_2O_3 contents, despite most experiments using natural basalt compositions that likely contain at least trace amounts of Cr_2O_3. Adding Cr_2O_3 into the melt phase changes the stability of pyroxene (see Fig. 3, Onuma and Tohara, 1983), and thus the pressure estimated by OPAM barometry (Voigt et al., 2017). Consequently, the systematic offset to higher pressures obtained by applying the V17 expressions to our compiled experimental database might result from the absence of reported Cr_2O_3 in most experiments, and uncertainty as to whether Cr was even present in starting materials (Wieser et al., 2023d). Additionally, it is unclear how precise the analyses of Cr_2O_3 in these experiments are; typical analytical conditions can result in large errors for this low concentration oxide (propagating into a much wider spread of pressures than using Y96 or H04, Fig. 4f). Despite its poor performance on our test dataset, we cannot rule out the possibility that V17 may be the best parameterization in natural Cr containing systems, even if it behaves poorly on Cr-free experiments. Further experimental work on Cr-bearing experiments where the Cr content of the melt was characterized with high precision at a range of pressures are required to assess the performance of the V17 expressions as an OPAM barometer in natural systems.

Interestingly, although all $N = 225$ experiments were OPAM saturated, only 27% ($N = 61$) pass the H18 cosaturation filter when the equations of Y96 are used (blue diamonds, Fig. 6a). While the statistics for calculated vs. experimental pressures are greatly improved using only experiments which pass the filter, it is concerning how many false negatives this filter produces, particularly as only 18 of the experiments which passed the filter were performed at pressures above 0.001 kbar. To further assess the performance of the H18 saturation filter, we isolated experiments on MORB-like compositions that are saturated in only one or two ($N = 172$) of the three key mineral phases in OPAM barometry. Using the Y96 expression, 19% of these non-OPAM saturated experiments pass the H18 co-saturation filter (false positives). These results demonstrate that the H18 co-saturation filter should not be used in isolation to identify which samples are suitable for OPAM barometry. Therefore, petrological observations of three-phase saturation remain critical for determining pressures via OPAM barometry.

To date, OPAM barometry has mostly been used to evaluate the storage pressure of MORBs, or tholeiitic to transitional OIBs (these compositions broadly overlap with the calibration dataset of Yang et al., 1996). Unfortunately, the uncertainty associated with OPAM barometry assessed using our independent test dataset (RMSE = 1.8–2.32 kbar, Fig. 6) indicates that this method does not have the resolution to distinguish between upper vs lower crustal storage at most mid-ocean ridges. However, at hotspot influenced ridges (e.g., Iceland) and regions of plume derived volcanism (e.g., ocean islands), the thickened crust (Fig. 3e) means that OPAM can provide insights into the characteristics of magma storage in these locations if independent checks for three-phase saturation are performed prior to pressure calculation (via petrographic observations or liquid line of descent analysis).

For example, melt inclusion and matrix glass data from Barðarbunga, Iceland, indicates that there is a slight difference in the entrapment pressure of melt inclusions and the final equilibration pressure of the matrix glasses (Hartley et al. (2018); Caracciolo et al., 2022).

Other liquid-based barometers

The effect of H_2O on phase boundaries in the OPAM system has not been investigated in detail, meaning this method cannot be applied to arc magmas with confidence. To provide a liquid-based thermobarometer with application to more hydrous systems like volcanic arcs, Blundy (2022) create parametrizations for the stability of clinopyroxene-hornblende-orthopyroxene-magnetite-plagioclase-ilmenite (CHOMPI) as a function of P, T, and the ratio of H_2O and CO_2 in the fluid, including a cosaturation check to establish whether CHOMPI phases were present based on just a liquid composition. However, Wieser et al. (2023a) show this cosaturation test classifies ~44% of ArcPL experiments lacking the CHOMPI assemblage as CHOMPI-saturated (false positives), and unsurprisingly these experiments show a poor correspondence between calculated and experimental pressure (as the system is not low variance). The logic behind the CHOMPI approach has been generalized to liquids saturated in a wider variety of phases by Weber and Blundy (2023), who apply a regression tree machine learning algorithm to a training dataset of liquids ranging from basalt to rhyolite. As this model is still being tweaked and is currently a preprint, we do not comment on it further.

Clinopyroxene ± liquid

The majority of Cpx-Liq thermobarometers rely on the exchange of diopside hedenbergite (DiHd, $CaFeSi_2O_6$) and Jadeite (Jd, $NaAlSi_2O_6$) between Cpx and liquid. Here, we briefly summarize the most recent parameterizations of Cpx equilibrium, and workflows that have been developed for calculating P and T in natural systems. A detailed discussion of pre-2008 models can be found in Putirka (2008, and refs within), and a description of the best Cpx-Liq models for our test dataset and sensitivity to P and T as well as the best equilibrium tests to use can be found in Wieser et al. (2023b,d).

Putirka (2008) propose a number of equations based on Cpx-Liq equilibria (Eq30, Eq31, Eq32c for P, Eq33 for T), expanding and recalibrating the models of Nimis (1995) and Putirka et al. (1996, 2003) to account for an increase in experimental data, particularly hydrous experiments. More recently, several new equations have been calibrated which focus on a more limited subset of compositional space than these global regressions (Fig. 7). Neave and Putirka (2017) calibrate a Cpx-Liq barometer for application to tholeiitic basalts from ocean island settings (specifically Iceland) using six experimental studies ($N = 113$ experiments) with mafic-intermediate compositions, and pressures ranging from atmospheric to 20 kbar (Neave and Putirka, 2017, Fig. 7, red diamonds). Masotta et al. (2013) conduct experiments and compile available literature data on trachytic to phonolitic compositions (Fig. 7, blue pentagons), to calibrate a Cpx-Liq barometer applicable to more alkaline compositions, which are poorly represented in the calibration dataset of P2008. Their expressions have a very similar/identical form to those of Putirka (2008), with recalibrated coefficients. Brugman and Till (2019) note that in high-silica melts, there is very little Al_2O_3 in Cpx (<2 wt %), and the calculated Jd component is very low or even zero. Along with the fact that high Si compositions are very poorly represented in P2008, they showed that this results in existing Cpx-Liq thermometers overestimated temperatures in evolved melts by up to 170 °C. Thus, Brugman and Till (2019) compiled recent studies from the literature on rhyolitic compositions (e.g., Almeev et al., 2012; Bolte et al., 2015; Gardner et al., 2014) supplemented with their own experiments on the Scaup Lake Rhyolite (Fig. 7, cyan squares), to calibrate a new thermometer for highly evolved melts which is independent of Jadeite, P, and H_2O content.

Another key shift in the use of Cpx-based barometry since 2008 has accompanied the rapid rise in programming literacy among petrologists, allowing the development of new workflows based on existing thermobarometry equations. For example, while experimental studies have clearly paired Cpx and liquid compositions, it is significantly more challenging in natural systems to identify such pairs. Typically, a given volcanic eruption will erupt a narrow range of liquid compositions along with a much more chemically diverse crystal cargo, incorporating Cpx crystals grown from a wide variety of melt compositions. Additionally, many whole rock compositions are a mix of crystals and melts, so are not representative of true liquids (Reubi and Blundy, 2009; Ubide et al., 2022). A variety of algorithms have been developed to try to identify matched pairs, by combining erupted Cpx compositions with liquids erupted over a longer time period at a given edifice, or even compilations of whole-rock data from an entire volcanic region.

For example, Maclennan et al. (2001) assessed magma storage pressures of basalts from the Icelandic Northern Volcanic Zone by filtering possible Cpx-Liq matches using the K_D equilibrium test of Putirka (1999). Winpenny and Maclennan (2011) devised a more complex test for Cpx-Liq equilibrium, using equation 35 of Wood and Blundy (1997) for K_D, and trace element equilibrium tests. Specifically, they use trace elements measured in Cpx, and then use the partition coefficients from Wood and Blundy (1997) to calculate the expected trace element contents of equilibrium liquid compositions. By comparing these predicted liquid compositions to available trace element data for compiled liquids (for Ce or La, Nd or Sm, and Yb, Dy and Y depending on data availability), they assess all possible matches between Cpx erupted at Borgarhraun, and ~1000 whole-rock and glass analyses from Icelandic basalts for equilibrium. However, this method requires trace element data in both Cpx and liquids/whole-rock, which isn't widely available in the literature for different volcanic systems (and is rarely collected as part of thermobarometric studies). Furthermore, the trace element partition coefficients used in these calculations are highly sensitive to the temperature of the system, complicating calculations (Sun and Liang, 2012; Wood and Blundy, 1997).

Neave and Putirka (2017) and Neave et al. (2019) develop an R code (also used by Gleeson et al., 2021 and Stock et al., 2018) to identify equilibrium Cpx-Liq pairs using components which can be calculated from elements measured routinely by electron microprobe (Diopside-Hedenbergite, DiHd; Enstatite-Ferrosilite, EnFs; Calcium Tschermak, CaTs). Specifically, they assess the

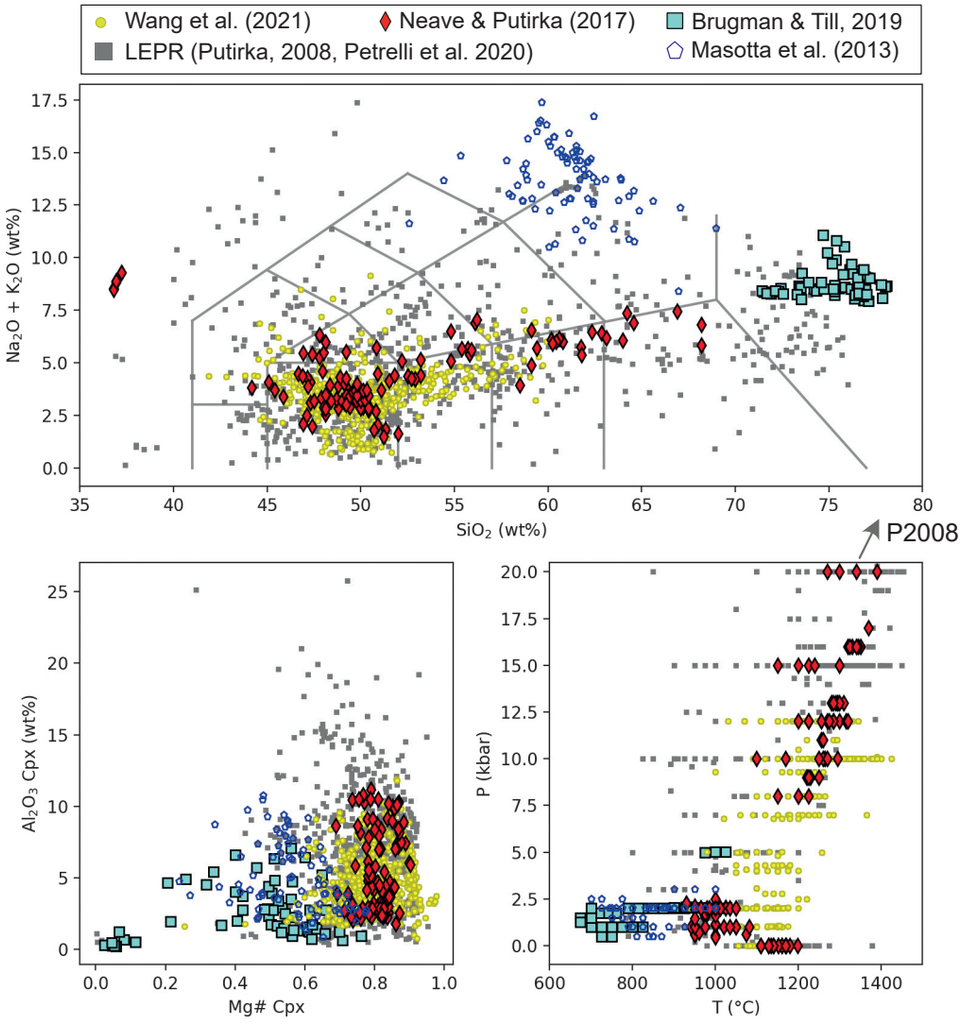

Fig. 7 Calibration range of Cpx-Liq and Cpx-only thermobarometers. Since the compilation of the LEPR database in 2008 (Hirschmann et al., 2008) used to calibrate Putirka (2008) and Petrelli et al. (2020, gray squares), many new thermobarometers have focused on specific regions of compositional space: Brugman and Till (2019, cyan squares), Masotta et al. (2013, blue pentagons), Neave and Putirka (2017, red diamonds). TAS lines drawn using Stevenson (2015), https://bitbucket.org/jsteven5/tasplot/src/master/.

difference in the measured components in the Cpx and the equilibrium components calculated from the liquid composition using Eq35 of P2008 for K_D, equations from Mollo et al. (2013) for DiHd and EnFs, and from Putirka (1999) for CaTs. As the selected K_D equilibrium equation is sensitive to T, and the DiHd, EnFs and CaTs equations sensitive to P and T, equilibrium tests must be calculated concurrently with P and T estimates. A computationally optimized algorithm loosely based on the Neave R algorithm (which was never publicly released) was incorporated in the Open-Source Python3 tool Thermobar (Wieser et al., 2022c). This tool allows comparison of 100 s of Cpx and Liq within seconds, with highly customizable equilibrium filters. Even with fast matching codes, identifying Cpx-Liq pairs is associated with substantial uncertainty regarding choice of equilibrium tests, and the chosen cut-off parameter (see Wieser et al., 2023b for arc magmas). In many systems, very few to no matches are found with erupted liquids, requiring generation of synthetic liquids to match to measured Cpx (e.g., Scruggs and Putirka, 2018). Another approach is to pair Cpx rims with matrix glass or groundmass separates (Klügel et al., 2020), although because this method focuses on rims, it may only reveal the depth of the uppermost storage chambers (neglecting deeper pressure information preserved in core compositions).

Issues associated with identifying equilibrium liquids can be circumvented using Cpx-only barometers. Putirka (2008) present two Cpx-only barometers (eq32a-b) and a Cpx-only thermometer (eq32d). Wang et al. (2021) show that these (and other existing) Cpx-only thermobarometers perform poorly in mafic and intermediate melt compositions. To address this, Wang et al. (2021) select 559 experiments conducted at 0–12 kbar with liquid SiO_2 contents between 42 and 60 wt% to calibrate a new barometer which is independent of T and H_2O (Fig. 7, yellow circles). They use the PyTorch python library (Paszke et al., 2019) to implement a

gradient descent algorithm to select Cpx cation components to include in a non-linear term for a Cpx-only barometer, rather than selecting components based on thermodynamic reasoning. They also present a thermometer, which is independent of P but requires knowledge of the H_2O content of the liquid.

Most recently, supervised machine learning methods have been applied to Cpx-based equilibrium to develop thermobarometry expressions (e.g., random forests, Higgins et al., 2022, extra trees, Jorgenson et al., 2022; Petrelli et al., 2020). Unlike more traditional empirical thermobarometers which are underlain by thermodynamic principles, measured oxide data are fed into these algorithms, rather than calculated components such as Jadeite. Wieser et al. (2023b) show that the Cpx-only barometers of Wang et al. (2021) and Jorgenson et al. (2022) show similar performance to popular Cpx-Liq barometers for the ArcPL dataset, justifying the use of Cpx-only barometers as an alternative to trying to identify equilibrium liquids. That said, just because Cpx-Liq and Cpx-only barometry show similar performance, it is worth nothing that neither approach performs particularly well, with neither method yielding $R^2 > 0.74$, or RMSE <2.1 kbar. Additionally, all expressions have low gradients, overpredicting at lower pressures, and underpredicting at higher pressures.

The relatively poor performance of Cpx-based barometry on independent test datasets (e.g., high RMSE, low gradients, inability to distinguish magma storage within ~10–15 km) is at least partially the result of low analytical precision when measuring Na_2O in experiments used to calibrate and test barometers (see Wieser et al., 2023d, and Section "Influence of analytical error on precision and accuracy"). Of course, analytical precision also affects measurements of natural samples and can result in an anomalously large spread of calculated P and T, which can be incorrectly attributed to transcrustal storage because of spurious trends in P-T space. For example, Wieser et al. (2023c) show that a the entire P-T range of the lunar Cpx analyzed by Luo et al. (2023) can be explained by the propagation of analytical uncertainty. To further demonstrate the importance of error propagation, we calculate pressures using Cpx and melt inclusion compositions from the first 2 weeks of the 2018 eruption of Kīlauea Volcano (Wieser et al., 2022a). Cpx-hosted melt inclusions yield relatively shallow saturation pressures, which overlap very well with the pressure at which a highly evolved magma was drilled in this region in 2005 (Teplow et al., 2009, yellow star, Fig. 8a). We use the PEC-corrected melt inclusion compositions and analyses of the Cpx composition close to the melt inclusion to calculate Cpx-Liq and Cpx-only pressures. These barometers return a much greater spread of pressures than melt inclusion saturation pressures, with a notable offset to significantly higher values (Fig. 8). In fact, these calculated depths are similar to the 3–4 kbar estimated by Putirka (2008) for Cpx erupted during the Pu'u O'o eruption of Kīlauea. However, there are no other geological or geophysical indications that Cpx-saturated magmas are stored at these pressures.

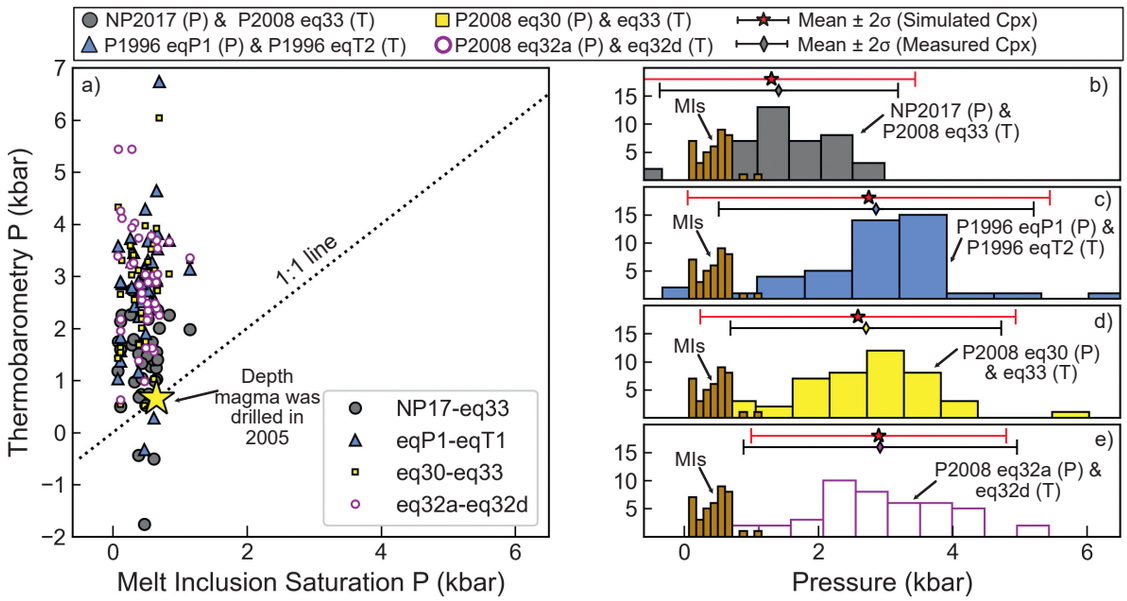

Fig. 8 Comparison of melt inclusion saturation pressures (Wieser et al., 2022a), and Cpx-Liq and Cpx-only barometry for samples erupted during the first two weeks of the 2018 Kīlauea eruption. (a) Melt inclusion saturation pressures calculated in MagmaSat (Ghiorso and Gualda, 2015) using VESIcal (Iacovino et al., 2021) cluster tightly at shallow pressures. These pressures are very similar to the pressure at which an evolved magma body was drilled in 2005 in the same area (yellow star, Teplow et al., 2009). Cpx-Liq and Cpx-only pressures calculating using PEC-corrected melt inclusion-host pairs show substantial scatter, extending to substantially deeper pressures. (b–e) Histograms showing calculated pressures for each iterative mineral barometry calculation. Error bars with black lines show $\pm 2\sigma$ for these calculations. Red error bars show $\pm 2\sigma$ for Monte-Carlo simulations for the mean composition of each Cpx and Liq, with 5000 Cpx compositions simulated using estimates of analytical precision for these Cpx measurements.

To interpret the large spread of calculated pressures, we simulate 5000 Cpx compositions, with the variability of each oxide following a normal distribution centered on the mean measured Cpx composition, with a standard deviation equal to the average counting statistic precision estimated from the EPMA output. We calculate pressures using these 5000 synthetic Cpx paired with the average liquid composition, and show $\pm 2\sigma$ of these simulations as a red error bar (Fig. 8b–e). The simulated spread in calculated pressures resulting from analysis of an entirely homogenous Cpx using typical EPMA operating conditions almost exactly matches the observed spread in calculated pressures (2σ error bars for each overlain on each histogram). This illustrates how, without sufficient averaging of individual analyses, it would be easy to incorrectly invoke storage at a range of pressures spread throughout the crust. While iteration of Putirka (2008, eq30) for T and Neave and Putirka (2017) for P produces calculated pressures that align within error of the melt inclusion saturation pressures, the other Cpx-based barometers estimate far too high pressures compared to melt inclusions and drilling depths. This demonstrates both the lack of precision and accuracy of Cpx-based methods. Even without analytical uncertainty, the RMSE on these barometers mean there is no way to reliably distinguish storage at ~0.5 kbar from storage at 4 kbar; limiting the applicability of these methods in systems where magma is stored in the upper part of a relatively thin crust.

The relatively poor performance of many Cpx-based barometers can also be attributed to the fact that many expressions are parameterized in terms of Cpx components (e.g., Jd, DiHd) calculated from EPMA oxide data (e.g., Neave and Putirka, 2017; Putirka, 2008). An alternative technique calculates site occupancy and cell volumes using a combination of single-crystal X-ray diffraction and measured oxide data (Dal-Negro et al., 1989). This technique was used extensively in the 80s–90s, and was recently applied by Tommasini et al. (2022) to natural Cpx crystals from Popocatépetl Volcano. They show that XRD-informed site assignments and the resulting calculated mineral components differ greatly from the routines used by Neave and Putirka (2017) and Putirka (2008) using EPMA data alone. This discrepancy between components calculated using EPMA data and the true crystal structure may reflect an additional source of uncertainty plaguing Cpx-based barometers. Single crystal XRD may also be vital to help us determine where Na is going (pairing with Al, Fe^{3+} or Cr^{3+}), which will further enhance our understanding of how Cpx composition relates to pressure. However, while single-crystal X-ray diffraction has been applied to natural crystals, the small size of experimental Cpx means that it is very difficult to obtain these parameters for experiments in order to calibrate expressions using these parameters (Dal-Negro et al., 1989; Tommasini et al., 2022). Additionally, a workflow involving XRD would be far more time consuming from a sample preparation perspective than existing EPMA-based techniques, which may severely restrict its uptake by the community, even if it results in improved barometer performance.

A number of papers have also shown the importance of considering the petrogenetic history of a Cpx population when interpreting thermobarometry calculations. For example, Hammer et al. (2016) assessed the influence of disequilibrium crystal growth on Cpx barometry using X-ray mapping and quantitative spot analyses of clinopyroxene crystals from a post-shield ankaramite erupted from Haleakalā volcano, Hawai'i. Their work demonstrates the complexity of interpreting Cpx compositions in regions where disequilibrium growth processes occur. Specifically, texturally and compositionally distinct domains within individual crystals thought to have formed by high degrees of undercooling during crystallization have distinct distributions of Cpx Jd contents, so return very different P estimates. They find offsets within individual crystals up to ~3.5 kbar. In contrast, Ubide et al. (2019a,b) calculate P-T-H_2O for different domains in sector zoned Cpx from Stromboli, and show remarkably constant results regardless of the textural context (although without individual calculated PTs, it is not possible to directly compare the statistical differences between these different studies). Clearly, further work is required to understand how the formation of different chemical domains in magmatic systems affect the distribution of P- and T-sensitive Cpx components, and thus the results from thermobarometric calculations.

The relatively recent appreciation that most magmas are stored in predominantly crystal-rich "mushy" environments also adds additional complexity to the interpretation of clinopyroxene-based barometric estimates, because clinopyroxene chemistry can be influenced by chemical processing in these crystal-rich regions. For example, residual melt compositions can be modified by reactive porous flow, driven by disequilibrium between a percolating melt phase and the surrounding crystal framework (Boulanger et al., 2020; Gleeson et al., 2021; Lissenberg and MacLeod, 2016; Sanfilippo et al., 2020). Critically, the melts formed from this process are rarely observed at the surface, which makes it difficult to quantify their chemistry to pair with erupted Cpx compositions. Cpx-Liq and Cpx-only barometry is also affected by the fact that reactive porous flow might drive interstitial melt compositions (and thus the Cpx compositions) outside the compositional range used to calibrate existing expressions (Fig. 7). Gleeson et al. (2021) demonstrate the importance of accounting for mush processes in their study of Cpx from wehrlite xenoliths found on Isla Floreana in the southern Galápagos. These crystals have Na_2O contents up to ~1.3 wt% and returned apparent crystallization pressures up to ~18 kbar using the clinopyroxene-only thermobarometers of Putirka (2008; eq. 32b–32d). In contrast, Cpx-Opx barometry (using the same crystals), melt inclusion and clinopyroxene-liquid thermobarometry from neighboring scoria cones on the same island indicate that magma storage was dominantly located at around ~7 kbar (Gleeson et al., 2022, 2021). Based on rare earth element signatures in these pyroxenes (elevated LREE contents relative to MREE and HREE contents) the authors concluded that the anomalously high pressure estimates of the wehrlitic clinopyroxene crystals result from chemical modification of Na_2O (and REE) by reactive porous flow in olivine ± clinopyroxene mush zones. Comparison with other thermobarometric techniques (e.g., Opx-Cpx, melt inclusions, Gao et al., 2022; Gleeson et al., 2021), as well as identifying trace element patterns indicative of reactive flow, provide one method to interpret the spuriously high pressure estimates returned from these cumulate Cpx.

In summary, to push Cpx-based thermobarometry forward, we need to improve the experimental datasets available to calibrate and test barometers (Wieser et al., 2023b,d), develop new methods to relate chemical components to parameters we can easily measure in natural samples and experimental charges (Tommasini et al., 2022), and increase our understanding of natural processes causing variation in Cpx compositions (Gleeson et al., 2021; Hammer et al., 2016; Neave et al., 2019; Ubide et al., 2019a,b).

Orthopyroxene ± Liq

Beattie (1993) presented the first Opx-Liq thermometer based on the use of non-regular solution models to calculate saturation temperatures. Putirka (2008) then calibrated two different equations (Eq28a, 28b) using the LEPR dataset, both of which are sensitive to P and H_2O. Eq28a uses the composition of the Opx and Liq, while Eq28b only uses the composition of the liquid to calculate the temperature of Opx saturation. Wood (1974) published the first Opx barometer, which used the Al content of Opx in equilibrium with Garnet. Putirka (2008) calibrated a more widely applicable Opx-only barometer (Eq29c) and two Opx-Liq barometers (Eq29a, Eq29b) using the LEPR database. The main pressure-sensitive component in Eq29a is a Jadeite-like component in the Opx, while Eq29b uses the (Fe, Mg) Al_2SiO_6 Opx component. Both equations also contain terms for other liquid cation fractions, and Opx cation fractions on the basis of 6 oxygens. Eq29c (Opx-only) uses the cation fractions of Ca, Cr and Al in the Opx, as well as terms for T. It is worth noting that several of the Putirka (2008) Opx thermobarometers are prone to numerical issues resulting from the presence of logarithmic terms. For example, Eq29c (Opx-only) has a term for the logarithm of the cation fraction of Cr in Opx. If Cr was not measured or reported in Opx, this term cannot be evaluated (the case for 63% of the experiments in our test dataset). Similarly, Eq28b has a term calculated from the log of the (Fe, $Mg)_2Al_2SiO_6$ component in the Opx divided by terms involving the cation fractions of Si, Al, Fe, Mn, and Mg in the liquid. In ~9% of experiments (mostly at <3 kbar, but a few at 5 and 8 kbar), the (Fe, $Mg)_2Al_2SiO_6$ component in the Opx is 0, the log of which yields infinity. A value of zero for the (Fe, $Mg)_2Al_2SiO_6$ component can also occur in Opx where the calculated Al_{VI} component is zero, which is particularly common in low pressure Opx ($Al_{VI} = 0$ in 7% of the Opx considered here, all at <2.5 kbar). To address the numerical issue associated with the (Fe, $Mg)_2Al_2SiO_6$ component, in later versions of the spreadsheets from Putirka (2008), two additional Opx-Liq barometers were included ("Global" and "Felsic"). These equations simply use the ratio of the measured Al_2O_3 content in Opx and Liq. The global model was calibrated on 795 Opx-Liq pairs, and the felsic model on 40 pairs.

In natural systems, it is even more difficult to assess the equilibrium relationships between erupted Opx and Liquid compositions than for Cpx-Liq, because only $K_{D,Fe-Mg}^{Opx-Liq}$ has been widely explored as an equilibrium test. Putirka (2008) find that K_D in experimental charges roughly correlates with the cation fraction of Si in the liquid (X_{Si}) ($1\sigma = 0.06$). In the ArcPL dataset, there is no significant correlation between K_D and X_{Si} ($R^2 = 0.01$, Fig. 9a), although ~67% lie within the given error bound of the predicted value. In natural systems is it very plausible that liquids and orthopyroxenes that are not chemically related could yield K_D values passing these relatively broad equilibrium tests (particularly if 2σ is used as the cut off). One way to avoid this problem is to use an Opx-only barometer (e.g., Eq29c of P2008), However, this equation contains a temperature-sensitive term, and given no Opx-only thermometer exists to our knowledge, a liquid composition will still need to be used to estimate (or iteratively calculate) the temperature.

Using experimental P and H_2O (that was measured by quantitative methods), Eq28a (Opx and Liq comps) and Eq28b (only Liq comps) are remarkably good thermometers. Eq28a has a slightly higher R^2 (0.90 vs 0.85), and lower RMSE (39 °C vs. 48 °C, Supporting Fig. 3). Using experimental T and H_2O measured by quantitative methods, Eq29b is the best Opx-Liq barometer ($R^2 = 0.58$, RMSE = 3.01 kbar, Supporting Fig. 4), and has a decent gradient and intercept relative to other mineral-melt barometers we evaluate in this review (grad = 0.67, int = 1.2). Eq29a performs slightly worse in terms of R^2 and RMSE but has a slightly better gradient and intercept ($R^2 = 0.50$, RMSE = 3.65 Gradient = 0.74, Int = 0.3 kbar). The global barometer is noticeably worse ($R^2 = 0.47$, RMSE = 3.47 kbar, grad = 0.59, int = 1.4). The Opx-only barometer (Eq29c) is difficult to compare because of the presence of many experiments without reported Cr measurements in Opx that return numerical errors; for available data, $R^2 = 0.04$, RMSE = 4.63 (Supporting Fig. 4d).

As discussed, it is probable in natural systems that neither T nor P is known, and melt H_2O is relatively uncertain. We investigate the sensitivity of calculated P to T, and calculated T to P (Fig. 10) by perturbing the experimental T by ±100 °C and P by ±5 kbar. The Putirka (2008) Global and Felsic barometers have no T term, so show no changes in P with T. Eq29a (dark blue lines, Fig. 10a) shows a relatively small increase in calculated P with increasing T, while Eq29b shows a larger increase (light blue lines, Fig. 10a). Eq29c, the Opx-only barometer, shows variable sensitivity to T, with P increasing in some samples, and decreasing in others as T is increased (gray lines, Fig. 10a). The strong change in calculated P as a function of T for Eq29c, a drop of up to 10 kbar for +100 °C for some samples, is problematic, as there is no way to determine temperature simply from Opx compositions at present. The three Opx-Liq thermometers also show variable changes with P (some increasing, some decreasing, Fig. 10b). However, the uncertainty in calculated T for changes in P of 10 kbar (equivalent to the crustal thickness in most arcs) is comparable to the stated RMSE on these thermometers. Overall, the relatively large change in calculated P (and to a lesser extent T) with relatively small changes in the other parameter means that we can expect Opx-based thermobarometers to show far worse statistics when solved iteratively than when tested with only a single unknown.

We also investigate the change in P when using iterative calculations when we perturb the experimental H_2O content by ±3 wt%. When T from Eq28a is iterated with all barometers, calculated T decreases with increasing H_2O (Fig. 10d). Iteration of Eq28a (T)-Eq29a (P) (dark blue line) and Eq28a (T)-Eq29b (P) (light blue line) show a small increase in calculated P with increasing H_2O content (Fig. 10c). Both these barometers have terms multiplying H_2O by a constant, resulting in a consistent change in P for varying H_2O contents in non-iterative calculations (+0.748 kbar/1 wt% H_2O for Eq29a, and +0.784 for Eq29b, dashed blue lines, Fig. 10c). However, because increased H_2O contents cause T to drop using Eq28a (Fig. 10b), which causes a decrease in the calculated P (Fig. 10a), the influence of H_2O on calculated P in iterative calculations is slightly less than that obtained in calculations using the experimental T. This complex feedback between equations during iterative calculations results in a greater variability in the influence

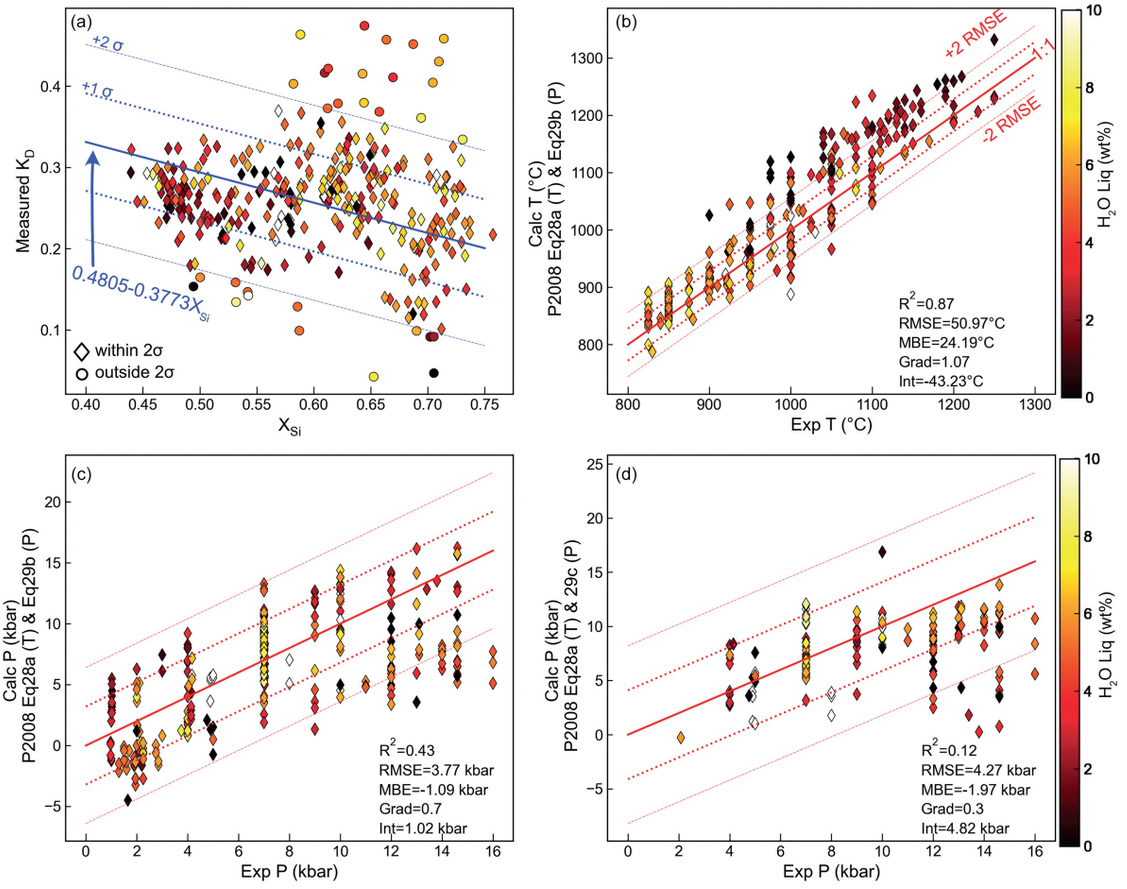

Fig. 9 Opx-Liq and Opx-only thermobarometry. (a) Comparison of measured K_D with that predicted from the X_{Si} content of the liquid using Putirka (2008). (b and c) Comparison of calculated and experimental T and P by iterating Eq28a(T) and Eq29b(P). (d) Iteration of Eq28a (for T, Opx-Liq) with Eq29c (Opx-only, P) of Putirka (2008). Dotted lines around the 1:1 line show ± the quoted 1σ, with dashed lines showing ±2 σ. All symbols colored by H_2O content, with $H_2O > 10$ wt% white, to emphasize the scale at lower H_2O contents.

of H_2O on calculated P for different samples. Eq29c has no H_2O term but shows the strongest sensitivity to T (Fig. 10a), resulting in a large drop in P with increasing H_2O when iterated with Eq28a (Fig. 10c). The high sensitivity of Eq29c to H_2O is concerning because an Opx-only barometer is most likely to be applied in systems where liquid compositions (and thus, H_2O contents) are poorly constrained. As the Global and Felsic barometers have no T or H_2O term, they show no sensitivity to H_2O.

Ideally, given the effects of H_2O on calculated P and T (Fig. 10c and d), we would only assess experiments where H_2O is known by quantitative methods (as for Liq-only). However, only 147/324 experiments from 10/23 studies have reported Opx-Liq compositions where H_2O was measured quantitatively. Thus, all experiments are shown in Fig. 9, with the same calculations using only experiments with quantitative H_2O measurements shown in Supporting Fig. 5 (the statistics for which are actually less favorable).

Iterating the best thermometer (Eq28a) and barometer (Eq29b) yields a very good match to experimental T in ArcPL using experimental H_2O contents (Fig. 9b). The thermometer performs worse for experiments with low H_2O contents (darker symbols, <2–3 wt%). Based on a similar observation for the LEPR dataset, Putirka (2008) suggest that Opx may re-equilibrate faster in hydrous liquids. Calculated P from iteration of Eq28a-Eq29b is reasonably accurate (Fig. 9c, grad = 0.7, MBE = −1.09 kbar, int = −1 kbar), but very imprecise (RMSE = 3.77 kbar), so this barometer will be of limited utility in relatively thin-crusted settings such as OIBs and MORBs (Fig. 3). Unlike for thermometers, there is no clear relationship between barometer performance and H_2O content. Although we cannot perform very many calculations for Eq29c because of the absence of Cr data, calculated Opx-only pressures (Eq29c) iterated with Eq28a shows a reasonable correspondence to the 1:1 line (Fig. 9d) at <12 kbar, although substantially more experiments with reported Cr contents are required to robustly assess this barometer.

Overall, Opx-Liq thermobarometry has been relatively neglected given its performance; no new calibrations have been proposed since 2008. In particular, the relative absence of systematic uncertainty when Opx-based barometers are tested on ArcPL contrast strongly with other mineral barometers discussed here. We suggest that future experimental and theoretical work to further develop

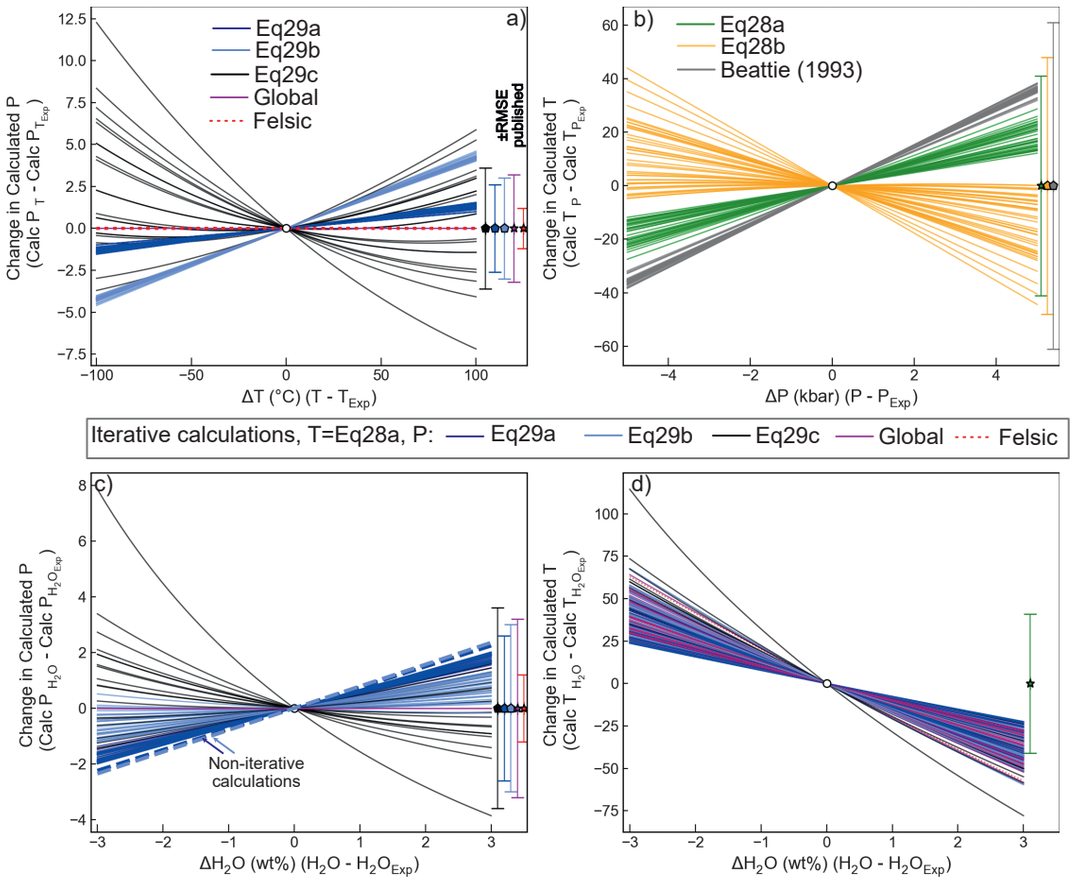

Fig. 10 Assessing the sensitivity of Opx thermobarometry to variation in P, T and H_2O for $N = 40$ randomly selected experiments. (a) Change in calculated P for perturbing the experimental T by ±100 °C. (b) Change in calculated T by perturbing P by ±5 kbar. (c and d) Change in calculated P and T for iterative calculations perturbing experimental H_2O content by ±3 wt%. For all plots, the calculated P or T at experimental conditions is subtracted from the calculation performed at perturbed P, T or H_2O conditions. Calculations performed in Thermobar (Wieser et al., 2022c).

Opx-based thermobarometry is warranted, with a particular focus on the development of more robust equilibrium tests. Ensuring starting materials contain Cr contents comparable to natural systems, and ensuring all elements are measured at high precision within experimental Opx (particularly Cr) may lead to improved thermobarometry calibrations. This may provide a particularly promising barometer in arc magmas, where Opx is relatively common, and magmas are relatively H_2O-rich.

Two pyroxene (Cpx-Opx) thermobarometry

Two pyroxene (Cpx-Opx) thermobarometry is widely used in the literature, likely reflecting the fact that these phases crystallize together across a wide range of P, T, and H_2O contents in a variety of tectonic settings. For example, Cpx-Opx thermobarometry has been applied to OIB basalts from the Galápagos by Gleeson et al. (2021), to mafic arc basalts from SW Japan by Zellmer et al. (2014), and to arc dacites from the Aucanquilcha Volcanic Cluster by Walker et al. (2013). Additionally, it is common that Cpx and Opx form crystal clusters, and calculations from these touching pairs are easier to justify than trying to select equilibrium liquids for each crystal (c.f. Opx-Liq and Cpx-Liq thermobarometry). Opx-Cpx thermobarometry can also be applied in systems where it is difficult to analyze the composition of the liquid (e.g., highly crystalline lavas or xenoliths, Gao et al., 2022; Gleeson et al., 2021).

As discussed above for Opx-Liq equilibrium, one of the main limitations of Cpx-Opx thermobarometry is the paucity of equilibrium tests to help filter out pairs which are not in equilibrium (in experiments and natural samples). The only established equilibrium test compares the exchange of Fe-Mg in Cpx and Opx ($K_{D,\,Fe\text{–}Mg}^{Cpx\text{–}Opx}$). Putirka (2008) suggest that $K_D = 1.09 \pm 0.14$ in high temperature systems, while $K_D = 0.7 \pm 0.2$ in subsolidus systems. However, it is difficult to know where to draw the line between these different filters; what value would be correct to use in a supra-solidus dacitic-rhyolitic melt that is a similar temperature to a subsolidus mafic cumulate? K_D values for our compiled experiments lie mostly within the high T bracket (±1σ shown in pink), although a number extend to higher values (Fig. 11a). There is no strong relationship between T and measured K_D. Another way to assess the equilibrium value would assume that the Cpx and Opx are each in equilibrium with the liquid, so the Opx-Cpx K_D value can be obtained from the Cpx-Liq and Opx-Liq K_D values from Putirka (2008):

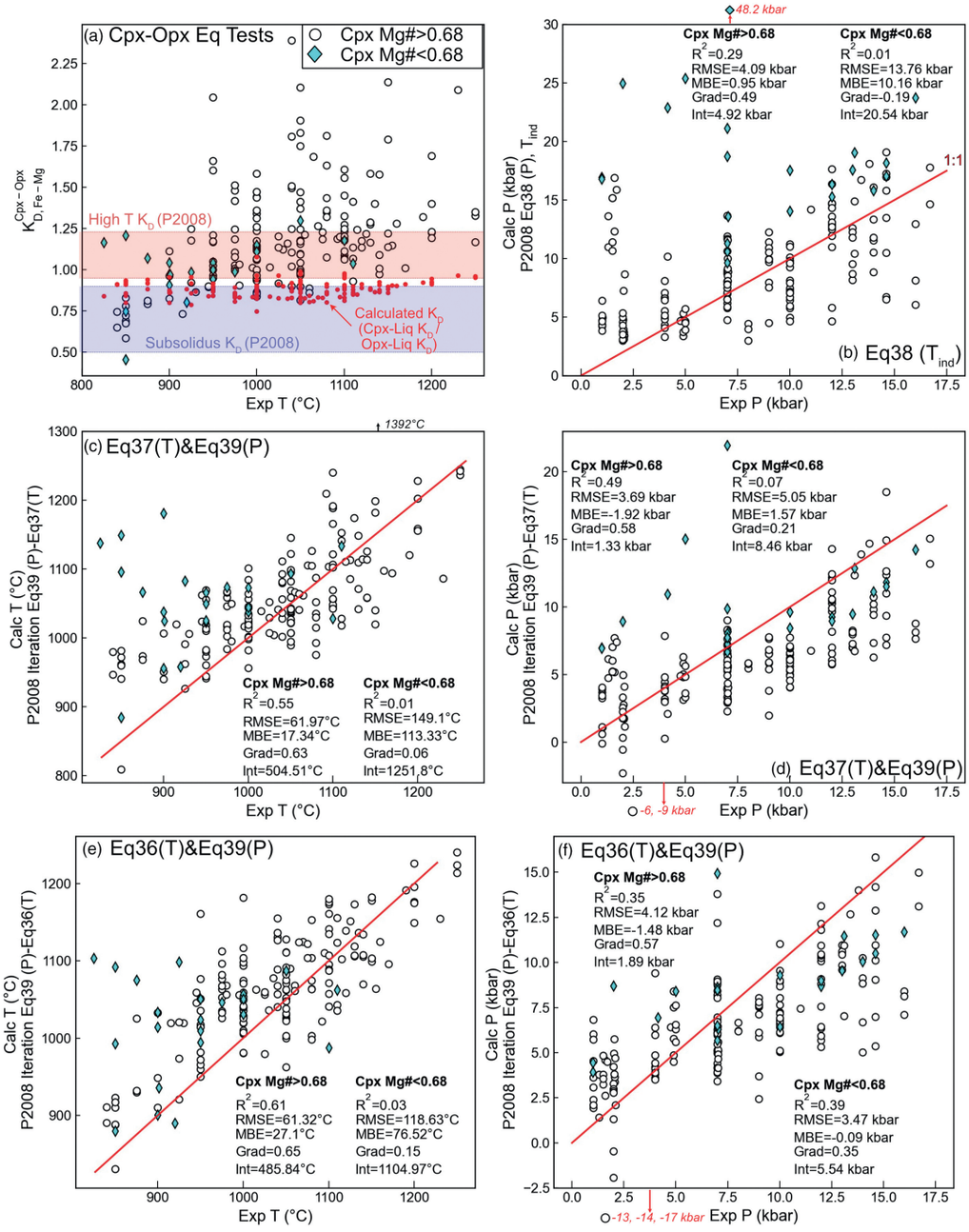

Fig. 11 Assessing two-pyroxene thermobarometers. (a) Measured $K_{D,Fe-Mg}^{Cpx-Opx}$ values for the ArcPL dataset against experimental T. The range of equilibrium values from Putirka (2008) for "HighT" and "Subsolidus" systems are shown as colored bars. K_D values calculated from the Cpx-Liq (T-dependent) and Opx-Liq (Si-dependent) K_D values (Eq. 8) are shown as small red dots. (b–f) Comparison of predicted and experimental P and T for different combinations of equations from P2008. 1:1 line shown in red. In all plots, Cpx-Opx pairs with Cpx Mg#<0.68 are colored cyan, and Mg#>0.68 colored white. To maintain scale, a few experiments returning extreme values are excluded from plots. Their y coordinate is labeled with a red arrow.

$$K_{D,Fe-Mg}^{Cpx-Opx} = \frac{K_{D,Fe-Mg}^{Cpx-Liq}}{K_{D,Fe-Mg}^{Opx-Liq}} = \frac{e^{-0.107-\frac{1719}{T\ (K)}}}{0.4805 - 0.3733 * X_{Si,Liq}} \quad (8)$$

This approach could be advantageous, because there are far more Cpx-Liq and Opx-Liq experiments than Cpx-Opx experiments, so the individual Px-Liq equilibrium values are better constrained. Values calculated using this expression lay between the subsolidus and high T value (red dots, Fig. 11a), with very little variation with temperature. This is because higher T drives up the Cpx-Liq K_D,

but higher T liquids tend to have lower Si contents, which drives up Opx-Liq K_D. Unfortunately, Eq. (8) appears to underestimate the K_D value for many of the experimental samples given here, which is perhaps unsurprising as we find the Cpx-Liq K_D expression underestimates many of the Cpx-Liq K_D values (by up to ~0.5), while the Opx-Liq expression tends to overestimate many of the Opx-Liq K_D values. Dividing these two relatively uncertain quantities compounds errors. The fact so many of our experimental Cpx-Opx pairs sit outside any estimate of equilibrium could suggest that they should not be used to assess thermobarometers. However, we find no discrepancy between K_D and the offset between calculated and predicted P and T. This suggests instead, that more experimental and theoretical work is needed to robustly assess K_D. Thus, we proceed to assess thermobarometers using all Cpx and Opx pairs in the ArcPL compilation.

The T-independent barometer of P2008 (Eq38) performs very poorly for high Mg# Cpx ($R^2 = 0.29$, RMSE = 4.1 kbar) and even worse for low Mg# Cpx (<0.68, cyan diamonds, Fig. 11b, $R^2 = 0.01$, RMSE = 13.7 kbar). Iteration of P2008 Eq37–39 and Eq36–Eq39 show very similar statistics to one another for calculated P and T, with Eq36–Eq39 showing slightly better behavior. Neither thermometer is very promising (Fig. 11c and e) producing more of a data cloud than a meaningful correlation. At lower temperatures, two pyroxene thermometers tend to overestimate temperatures (also noted by Ziberna, 2021). Iterated barometers are similarly disappointing; even for Cpx with Mg#>0.68, the RMSE is 3.7 kbar (Eq37–Eq39, Fig. 11d) and 4.1 kbar (Eq3 6–Eq39, Fig. 11c).

Overall, we suggest that substantially more experiments where Opx and Cpx are stabilized at a variety of pressures and melt compositions are required to improve Cpx-Opx thermobarometers, particularly in more evolved systems with lower Cpx Mg#s. Additional experiments would also help to better constrain controls over the equilibrium value at a variety of temperatures and melt compositions.

Amphibole thermobarometry and chemometry

Amphibole (Amp)-only and Amp-Liq thermobarometry have been used extensively to calculate P and T in volcanic and plutonic igneous systems (e.g., Higgins et al., 2022; Scruggs and Putirka, 2018). Amphibole chemometers are also becoming widely used to probe the compositions of melts present at depth within plumbing systems which are not always well represented at the surface (e.g., Humphreys et al., 2019; Zhang et al., 2017).

Amphibole thermobarometry

Amphibole barometry stems from the seminal work of Hammestrom and Zen (1986), who showed that the Al^{VI} and Al^{Tot} contents of hornblendes from calc-alkaline plutons emplaced at different depths correlate with estimates of pressure, and that the same compositional-pressure relationships were seen in experimental products. Hollister et al. (1987) analyzed rim compositions from plutons where the pressure of emplacement could be estimated from phase assemblages in the surrounding country rock. They confirmed the Al-pressure relationship of Hammestrom and Zen (1986) and proposed an updated calibration. Additional calibrations have also been developed for Amp-Plag, Amp-Garnet, and Amp-Plag-Qtz (see Molina et al., 2021b and refs. within).

More recently, Mutch et al. (2016) present an Amp-only barometer parameterized in terms of Al^{tot} calibrated using a dataset comprising: (i) their new experiments on 3 different bulk compositions, (ii) the experiments of Johnson and Rutherford (1989), Schmidt (1992), Thomas and Ernst' (1990), (iii) published analyses from plutons with independent depth constraints (Ague, 1997; Hammestrom and Zen, 1986), and (iv) their new analyses on amphiboles from the Yerrington Batholith in Nevada. Mutch et al. (2016) exclude experiments where garnet or phengite was stabilized, as this alters the Al^T-pressure relationship. Importantly, their compiled dataset shows a curvature in Al^T vs. pressure space relationship at 0.5–3 kbar not seen by previous workers. They parameterize this empirically relating pressure to a second order polynomial of Al^T.

While Al^T in amphibole is controlled by both T and P, so can only be used as a barometer on an isotherm or near a solidus (Médard and Le Pennec, 2022), correlations between P and Al^{VI} have been noted by Krawczynski et al. (2012), (Helz, 1982), Larocque and Canil (2010) and Schmidt (1992) in a wide variety of systems. Médard and Le Pennec (2022) present a T-independent barometer using a simple linear regression of P and Al^{VI}, calibrated on 47 published experimental compositions with Si-rich silicate melts in equilibrium with Biotite, Plagioclase, and Magnetite. They state that this barometer has a RMSE of 0.86 kbar (or 0.72 kbar for <4 kbar). They also test their expression on 22 biotite-bearing experiments not used in calibration, which lie within the stated RMSE window. However, all their tests are at <4 kbar, which makes it difficult to evaluate the precision of the barometer at higher pressures.

Importantly, all the Al-based amphibole-only barometers discussed so far are only applicable in the presence of certain phases. For example, Hammestrom and Zen (1986) warned about the applicability of their regressions in rocks without quartz, where Al^T can be significantly higher at a given P. Mutch et al. (2016) emphasize that their barometer should only be applied to amphibole rims in equilibrium with melts saturated in plagioclase (An_{15-80}), biotite, quartz, alkali feldspar, Fe-Ti oxides, and apatite (e.g., near solidus, low variance volcanic systems). In higher T melts with fewer co-crystalizing phases and higher thermodynamic variance, the equilibria controlling amphibole composition are still poorly understood (Putirka, 2016). Médard and Le Pennec (2022) test their barometer using 7 Bt-free experiments, resulting in a RMSE of 1.4 kbar, which is a large % error given these experiments were conducted at ~1–2 kbar. Thus, they suggest that this barometer should not be used in biotite-free rocks. While limited in their applicability to silicic systems saturated in a large number of phases, the amphibole-only barometers discussed thus far are extremely useful for determining pluton emplacement depths and crustal exhumation rates in orogenic belts, or when investigating the formation and evolution of Porphyry copper deposits (Anderson, 1996; Hollister et al., 1987; Mutch et al., 2016).

In contrast to these studies focused on low variance silicic systems, Ridolfi et al. (2010) compile calcic amphibole analyses from experiments conducted on a wide range of melt compositions and phase assemblages to generate an empirical amphibole-only thermometer (Eq1), oxybarometer (Eq2) and hygrometer (Eq3). Ridolfi et al. (2010) justify the application of Amp-based methods in less evolved systems with fewer phases (and therefore higher variance) by pointing out that most volcanic amphiboles exist close to their stability limits (shown by abundant breakdown textures). They suggest that close to the amphibole stability curve, the variance of the system is lower so amphibole composition is more closely related to P, T and fO_2. Ridolfi et al. (2010) also present an Al^T barometer calibrated on 9 amphiboles (Eq4) but find that attempts to perform a similar calibration on their larger experimental dataset performs poorly, particularly for magnesiohastingsite amphiboles which are common in nature, but sparse in available experiments.

Ridolfi and Renzulli (2012) compile additional experiments and apply more stringent filters to produce an amphibole-only barometer (e.g., removing amphiboles with 1σ for $SiO_2 > 0.9$ wt%, Piston cylinder experiments conducted at <6 kbar, etc.). Of the 61 experimental amphiboles remaining, 19 (31%) are magnesiohastingsite, which is close to the proportion of magnesiohastingsites (42%) in their natural compilation. This is in stark contrast to the dataset of Ridolfi et al. (2010), which contained no magnesiohastingsites (or pargasites or kaersutites). Ridolfi and Renzulli (2012) create several different regressions to calculate P using amphibole cation fractions calculated on the basis of 13 oxygens. For example, Eq1a was calibrated on $N = 61$ experiments, and expresses P in terms of the exponential of compositional terms (Si, Ti, Al, Fe, Mg, Ca, Na, K). However, using comparisons to seismic depths from Ridolfi et al. (2010) for a variety of natural systems, they show that this equation overestimates at low P, and underestimates at high P. They calibrate 4 additional expressions using a smaller subset of analyses ($N = 20$–41) in different pressure ranges (Eq1b and 1e use exponentials, 1c and 1d use multilinear regressions). They present an algorithm where the user calculates P for each equation, and these different values are averaged/combined in a variety of different ways to construct a final P.

To test the expressions of Ridolfi and Renzulli (2012), Erdmann et al. (2014) compile a series of experiments not used in their calibration dataset. While T, melt SiO_2 content and fO_2 are reasonably well predicted in their new dataset, they demonstrate a very poor correspondence between calculated and experimental P, and calculated and experimental H_2O contents. Erdmann et al. (2014) suggest the P discrepancy results from the fact that the Si-Al content of an amphibole is more strongly related to the liquid composition and T than P. In particular, they note that the calibration dataset of Ridolfi and Renzulli (2012) is skewed, with amphiboles from felsic and intermediate melts clustered at lower P, and mafic to intermediate melts at higher P. Molina et al. (2021a) also publish a short note stating that they test the Ridolfi and Renzulli (2012) barometer using a dataset of experiments compiled by Molina et al. (2021b), and obtain "*unsustainable pressure estimates.*"

Putirka (2016) further examine Amp-only and Amp-Liq equilibrium, presenting two P-independent Amp-only thermometers (Eq5 and SiHbl), two P-dependent Amp-only thermometers (Eq6 and Eq8), two P-independent Amp-Liq thermometer (Eq4b and Eq9), and a P-independent Liq-only amphibole saturation thermometer (Eq4a). He also presents three T-independent Amp-Liq barometers (Eq7a, b, c). Using an extensive test dataset, Putirka (2016) show that Amp-only and Amp-Liq equilibrium do a reasonably good job of predicting pressure when averaging multiple amphiboles, but conclude that P estimates from individual amphibole grains are "*nearly useless for understanding crustal processes.*" They hypothesize that the main limitation of Amp-based barometry is the fact that none of the common amphibole components have particularly large changes in volume when they precipitate from the liquid (ΔV fusion), or large molar volume contrasts between different components in the amphibole itself. For example, even though Jadeite in pyroxene is not a particularly sensitive barometer at crustal conditions, relative to Diopside it has a significantly more negative ΔV fusion (Fig. 12a), and a smaller molar volume (8% smaller, Fig. 12b, Putirka, 2016). As Jadeite has both a smaller molar volume and a more negative ΔV fusion (Fig. 12c), the amount of Jadeite vs. Diopside in Cpx are sensitive to P. In contrast, there is a pretty weak

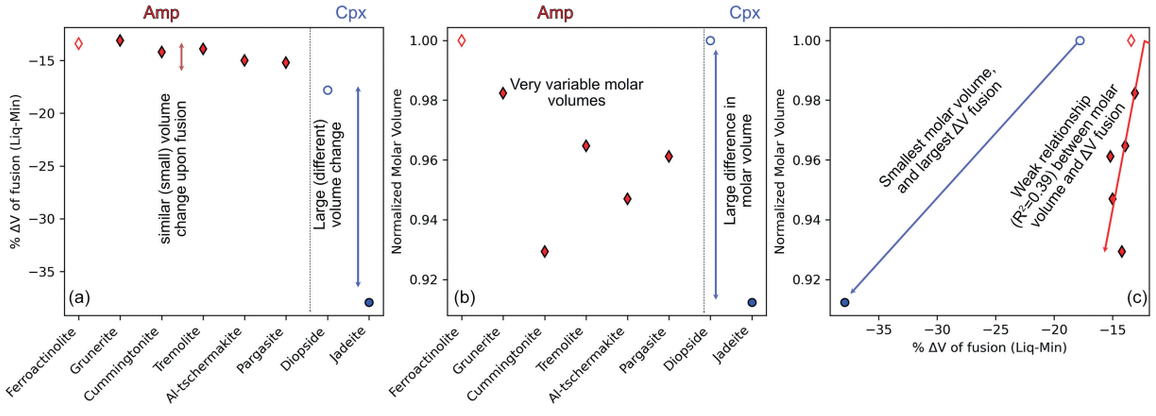

Fig. 12 Comparison of molar volumes and volumes of fusion for Amp vs. Cpx components. (a) Changes in volume following precipitation of different mineral components from the liquid (ΔV fusion). Amphibole components show significantly smaller changes than jadeite, and the different Amp components show very similar ΔV fusion. (b) Molar volume normalized to the component of each phase with the highest molar volume (ferroactinolite for Amp, diopside for Cpx). (c) Correlation between normalized molar volume and ΔV fusion. Thermodynamic data from table 1 of Putirka (2016).

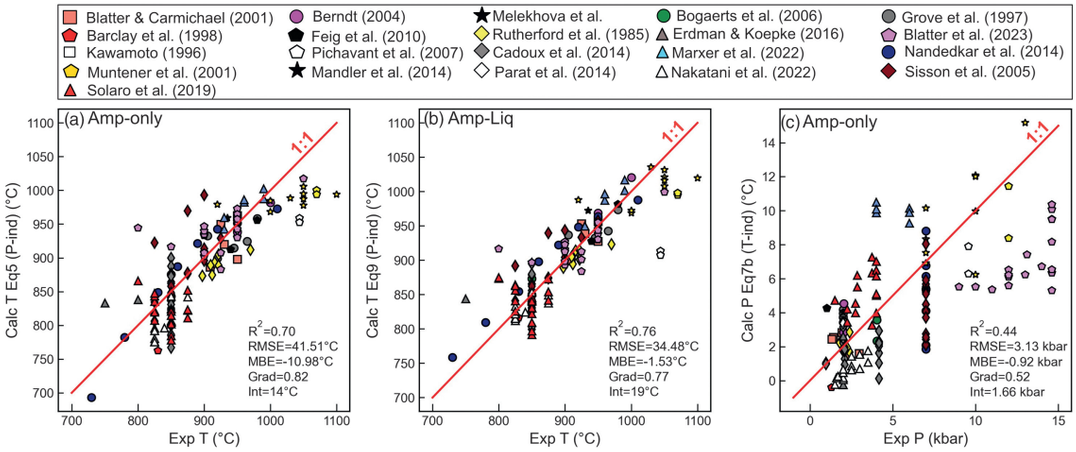

Fig. 13 Assessing Amp-Liq and Amp-only thermobarometers from Putirka (2016) using experiments not used during model calibration. Only experiments passing the K_D filter of Putirka (2016) are shown.

relationship between the ΔV fusion and the molar volume for different amphibole components (Fig. 12b and c, Putirka, 2016). While the exchange of different amphibole components is not particularly P-sensitive, Putirka (2016) do show that the partition coefficient of Al between amphibole and liquid is correlated to P, explaining why amphibole barometers are normally parameterized in terms of Al and other oxides, rather than explicitly calculated mineral components (e.g., Jadeite in Cpx).

We use ArcPL data not used for calibration of the Putirka (2016) equations to test their performance and sensitivity to other terms present in the regression (see Supporting Figs. 6–8). The best performing Amp-only thermometer for this dataset is Eq5 ($R^2 = 0.7$, RMSE = 41.5 °C, Fig. 13a, Supporting Fig. 6), which has the advantage of being independent of P and H_2O in the liquid. The best Amp-Liq thermometer is Eq9 ($R^2 = 0.76$, RMSE = 34 °C, Fig. 13b, Supporting Fig. 8), which is P independent and not very sensitive to H_2O in the liquid (Supporting Fig. 7, ~1 °C change in T per 1 wt% H_2O). Like Putirka (2016), we find that all three Amp-Liq barometers show disappointing statistics (RMSE = 3–4 kbar, Grad = 0.3–0.5, Supporting Fig. 8, e.g., Fig. 13c). In particular, the pressures for the higher P experiments of Blatter et al. (2023), and Müntener et al. (2001) are greatly underestimated, although the barometer does a reasonable job of the lower P data, which clusters around the 1:1 line at ~2–5 kbar. The skew at high P means that even after the averaging suggested by Putirka (2016), Amp-Liq barometry can give misleading results.

Ridolfi (2021) updated the Ridolfi and Renzulli (2012) amphibole-only barometer, perhaps partially in response to the criticism of Erdmann et al. (2014), Molina et al. (2021a) and Putirka (2016), and partially because of an increase in the number of amphibole-bearing experiments. Using an expanded experimental dataset, they tweak the algorithms used to select between different P equations and add more stringent filters to give users warning of when the equations are being applied to amphiboles failing quality tests or outside the model calibration range (based on totals, unbalanced charge, low B cations, etc., low and high Ca cations, low Mg cations). We tested this new algorithm and quality filter on experiments from ArcPL not used during calibration (Alonso-Perez et al., 2009; Berndt, 2004; Blatter and Carmichael, 2001; Bogaerts et al., 2006; Cadoux et al., 2014; Erdmann and Koepke, 2016; Grove et al., 1997; Kawamoto, 1996; Marxer et al., 2022; Nakatani et al., 2022; Parat et al., 2014; Rutherford et al., 1985; Sisson et al., 2005; Sisson and Grove, 1993; Solaro et al., 2019; Blatter et al., 2023). Of our compiled experiments ($N = 193$), 112 pass the quality check filters of Ridolfi (2021, Fig. 14a). For those passing the quality filters, calculated P show a moderate correspondence to experimental P, although calculations clearly lie well outside the stated ±12% error in the abstract of Ridolfi (2021, red dashed lines, Fig. 14b). The overall fit to these experiments yields $R^2 = 0.36$ and RMSE = 3.6 kbar. The experiments conducted at 2 kbar by Sisson and Grove (1993) show particularly poor results, returning pressures which are up to 15 kbar too high. Excluding these experiments yields $R^2 = 0.67$ and RMSE = 2.7 kbar. Compared to the Putirka (2016) Amp-Liq barometers, this Amp-only barometer does a much better job of recreating the high pressures of Blatter et al. (2023), although it does still underestimate to a degree. Despite the criticism of this method in the literature, it is noteworthy that Amp-only barometry doesn't perform any worse than Opx-Liq and Opx-Cpx based on our dataset, but none of these methods are precise enough to be useful for many volcanological questions. As amphibole-only T from Ridolfi and Renzulli (2012) require a P to be entered, we also test how effective this thermometer is using pressures obtained from the 2021 barometer. The correlation between calculated and experimental T is reasonably good (Fig. 14a, $R^2 = 0.53$, RMSE = 53 °C), although the RMSE is twice that stated in the abstract of Ridolfi (2021, 22 °C, shown with dashed red lines, Fig. 14a).

We examined the Sisson and Grove (1993) experiments in the context of the calibration dataset of Ridolfi (2021) to try to understand the poor performance of the barometer. These experiments plot to substantially higher amphibole Na_2O contents at lower pressures than any of the calibration experiments (Fig. 14d). To investigate whether this offset to higher Na_2O contents could explain the anomalously high calculated P, for each Sisson and Grove (1993) amphibole, we perturb the Na_2O content by ±2 wt%. Pressures for measured amphibole compositions are shown as squares (Fig. 14e), with a line stretching from each square showing

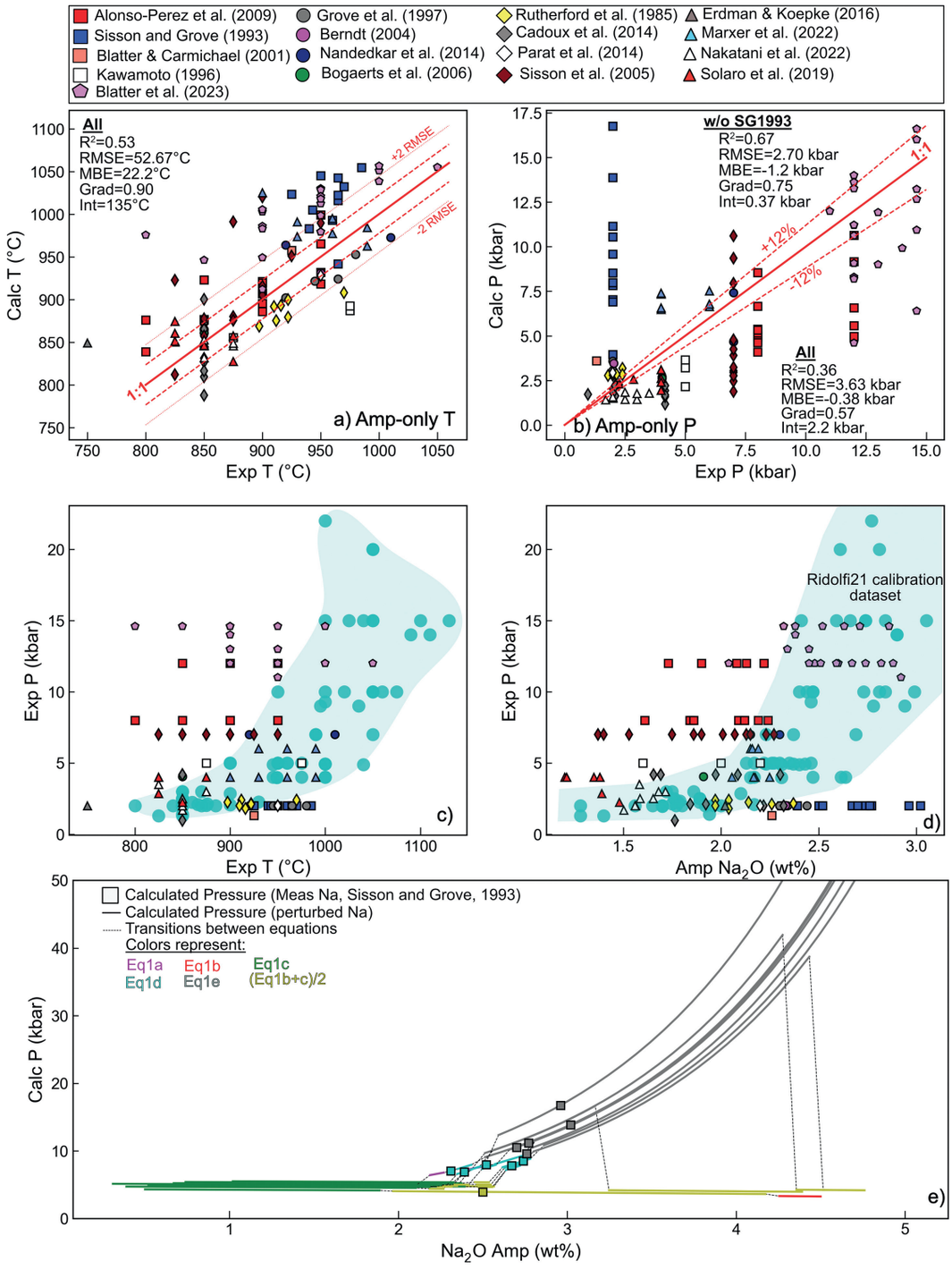

Fig. 14 Assessing Amp-only thermometers and barometers. (a) Comparing experimental and calculated T using the thermometer of Ridolfi and Renzulli (2012) iterated with Ridolfi (2021). (b) Comparing experimental and calculated P using the new barometer of Ridolfi (2021, T-independent). (c and d) Visualizing experimental compositions relative to the calibration dataset (cyan dots and field) of Ridolfi (2021). Many of the experiments showing the largest discrepancies in calculated pressure (e.g., red and blue squares, magenta diamonds) lie outside the calibration range of Ridolfi (2021) in Pressure-Na_2O space. (e) To visualize the effect of Na_2O on calculated pressure, we perturb the Na_2O content of the experiments of Sisson and Grove (1993) by ± 2 wt%. The lines are not continuous, with jumps to a different pressure as the algorithm flips between different "root" equations (indicated by the color of the line). Experimental Na_2O (and calculated pressures) are indicated with squares, with the calculations following the perturbation shown as lines.

the change in pressure as Na_2O is changed. The colors indicate the root equation from Ridolfi (2021) algorithm used to determine P, with the rapid jumps in pressure reflect a flip to a different equation selected by the algorithm. For the samples with high calculated P (>7.5 kbar), it is very clear that Eq1e (and to a lesser extent 1d and 1a) is highly sensitive to the Na_2O content, rapidly shooting up to extremely high pressures for very small changes in Na_2O. It may well be that these natural samples lie outside the Na_2O range used to calibrate Eq1e (the calibration data for each specific equation is not available). The Alonso-Perez et al. (2009) experiments (red squares) where pressure is underestimated are also clearly offset from the calibration dataset in P—Na_2O space (Fig. 14c). This comparison emphasizes the importance of ensuring that sample compositions are well represented in the calibration dataset of the chosen model, not just in terms of P-T space, but also compositional space. It also shows that having the correct functional form for a barometer routed in thermodynamics is essential to minimize extrapolation issues commonly seen with empirical fits.

Amphibole chemometry

Ridolfi and Renzulli (2012) also present equations to calculate the contents of SiO_2, TiO_2, Al_2O_3, FeO_t, MgO, CaO, and K_2O in the melt from which amphiboles crystallized. These equations are parameterized in terms of amphibole composition and pressure. However, given the uncertainties discussed above calculating pressures from amphibole compositions, these P-sensitive parameterizations can be problematic to implement in natural systems. Erdmann et al. (2014) show that the predicted SiO_2 content is a reasonable match above 65 wt% SiO_2, but tends to overpredict SiO_2 for more mafic melts. Putirka (2016) use their newly compiled dataset to produce an updated expression (Eq10) for melt SiO_2 from the amphibole using the temperature of the melt and the cation fraction of Al in the amphibole.

Zhang et al. (2017) compile a reasonably similar calibration dataset to Putirka (2016), and produce multiple regressions to calculate melt oxide components from a variety of amphibole site positions (e.g., Si-Ti-Mg-Fe-Ca in amphibole for predicting SiO_2 in the melt). These expressions are P-independent, and only Eq3 (for SiO_2) and Eq5 (for TiO_2) are T-sensitive. Zhang et al. (2017) calibrate multiple equations for some melt oxide contents (e.g., four equations for SiO_2, two equations for FeO_t). We test the 133 ArcPL amphiboles which pass the Zhang et al. (2017) equilibrium filter ($K_{D,\ Fe\text{-}Mg} = 0.28 \pm 0.11$) and do not appear in their calibration dataset. The calculated statistics are similar if a K_D filter isn't used. Overall, unlike many of the barometers discussed in this review, these chemometers perform well for experiments they were not calibrated on (Fig. 15). Eq2 of Zhang et al. (2017) does a good job of predicting melt SiO_2 content across a wide range, showing much better performance than Eq10 of Putirka, 2016 (Fig. 15a vs. b). The statistics of the fit are similarly good for melt Al_2O_3 (Eq14). There is more scatter for other oxides, and reasonably large differences between the different provided equations, but it is not always clear which equation is better (e.g., Fig. 15c). It is notable that the worst performance is seen for experiments with oxide contents toward the tail end of the calibration dataset (gray histograms, Fig. 15). For example, Eq12 and Eq13 do a good job of predicting K_2O until ~3.5 wt% K_2O; very few experiments in the calibration dataset had such high K_2O contents (Fig. 15g). Similarly, the fit is better at lower MgO contents, where the calibration dataset is concentrated (Fig. 15e). In general, our tests demonstrate that amphibole compositions can be used to estimate the melt compositions from which they grew, as long as the results are carefully evaluated relative to the calibration range of the model.

We do not test predictions of H_2O and fO_2 using amphibole-only chemometers, as there are limited reliable experimental data with well constrained values for these parameters that were not used during model calibration.

Olivine-spinel aluminum-exchange thermometry

The aluminum content of olivine has been used as a thermometer in the mantle (De Hoog et al., 2010), and Al partitioning between olivine and spinel has been used in igneous rocks (Wan et al., 2008; Coogan et al., 2014). Here we focus on the olivine-spinel Al-exchange thermometer, which offers an advantage over Mg-Fe olivine-liquid thermometry because Al in spinel and olivine and Cr in spinel are more resistant to diffusional modification during crystal storage and transport than Fe-Mg in olivine (Spandler and O'Neill, 2010). Additionally, while olivine crystals are frequently out of equilibrium with erupted liquids (Sides et al., 2014b; Wieser et al., 2019), the fact spinels are trapped inside olivine crystals makes it more straightforward to identify mineral-mineral pairs which grew together.

The mechanism by which aluminum substitutes into olivine remains somewhat uncertain, despite the fact that constraining this reaction is vital to identify which chemical parameters should appear in a thermodynamically-constrained thermometry model. Of the possible substitutions, two mechanisms have been the focus of the most attention:

$$MgAl_2O_4^{spinel} \rightarrow (Mg, Al)AlO_4^{olivine} \quad \text{(mechanism 1)}$$

$$MgAl_2O_4^{spinel} + 2\ SiO_2 \rightarrow (MgAl_2V)Si_2O_8^{olivine} \quad \text{(mechanism 2)}$$

In the first mechanism there is a coupled substitution of Al into both the octahedral and tetrahedral sites in olivine, while the second has substitution only onto the octahedral site, with charge balance maintained by vacancies (V). It is also possible coupled substitutions could take place with Cr and Na in the octahedral site (and Al in the tetrahedral site), or Al in the octahedral site with Fe^{3+} in the tetrahedral site (Taura et al., 1998). Critically, if the mechanism 1 is dominant, the thermometer should depend only on the Al contents of the co-existing olivine and spinel (in addition to any chemical parameters controlling the activity coefficient for Al in either phase), but if mechanism 2 dominates, the thermometer will depend also on the activity of SiO_2 (a_{SiO_2}).

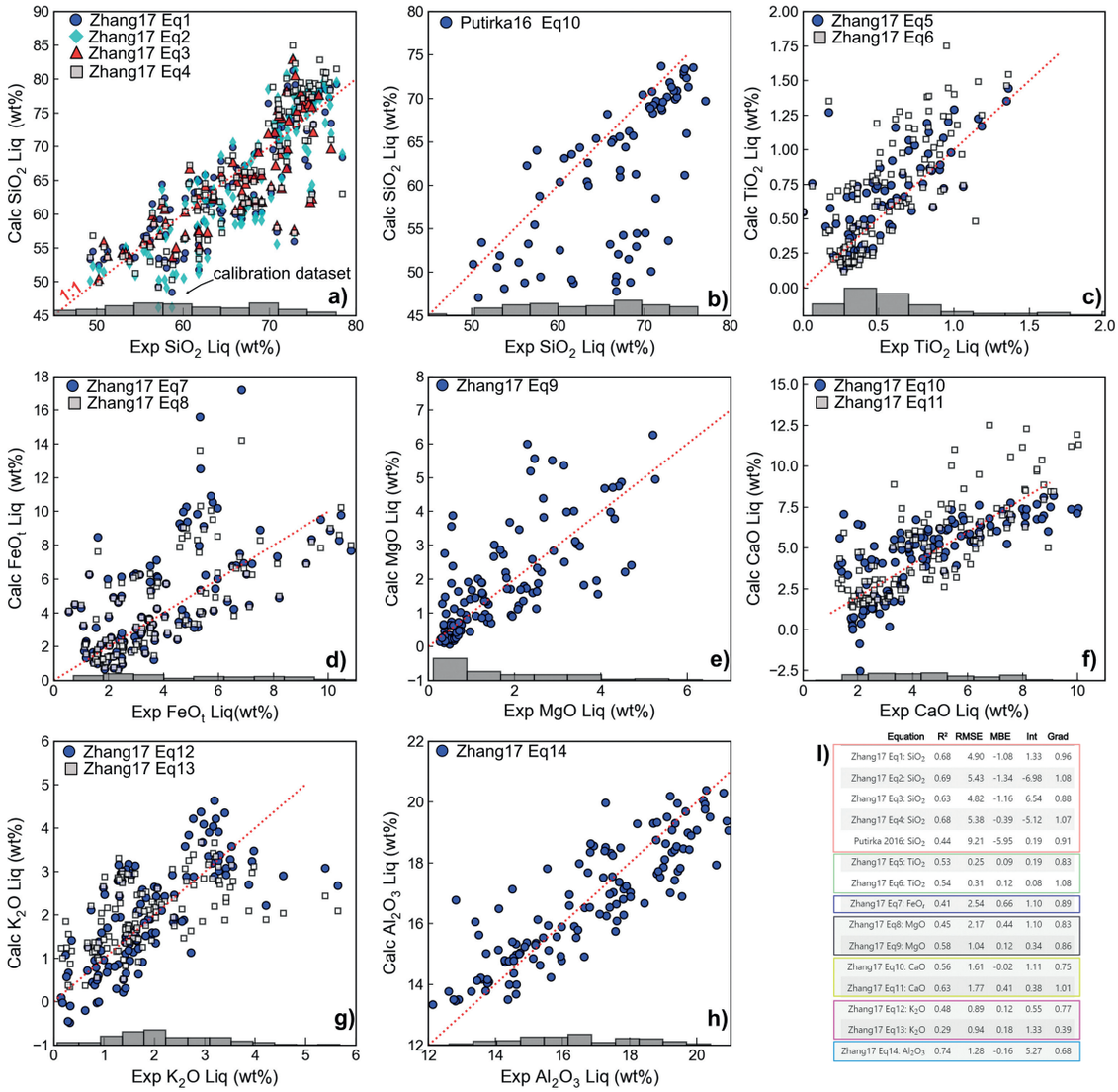

Fig. 15 Assessment of amphibole chemometers of Zhang et al. (2017, panel a, c–h) and Putirka (2016, panel b). For equation 3 and 5 of Zhang et al. (2017) and Putirka (2016) eq10, we use temperatures calculated from co-solving T from Ridolfi and Renzulli (2012) and P from Ridolfi (2021). Putirka (2016) eq10 also requires P from Ridolfi (2021) to be input. These equations have fewer datapoints on the plot, as we exclude P and Ts where Ridolfi (2021) returns an input warning. (i) Statistics for each equation. The filtered test dataset comprises 133 experimental charges not used by Zhang et al. (2017) during calibration, and 10 experimental charges from Barclay (2004), Blatter and Carmichael (2001), and Grove et al. (1997), which were used by Zhang et al. (2017) in their test, but not calibration dataset.

A thermometer based on Al-exchange was first calibrated by Wan et al. (2008) using a series of experiments at 1 bar with the bulk composition varied such that olivine and spinel co-crystallized at a range of temperatures and Cr contents. They found that the experimental data could be adequately modeled with a formula depending on the ratio of Al in olivine to Al in spinel, as well as the spinel Cr# (Cr/[Cr + Al], molar). The dependency on Cr# comes from its effect on the activity coefficient of Al in spinel. Wan et al. (2008) justified the extrapolation of the thermometer to higher P and T than the calibration dataset by comparing temperatures derived from the Al-exchange thermometer with temperatures calculated from the two-pyroxene thermometer for a suite of olivine and spinel bearing mantle xenoliths. Although there was considerable scatter around the 1:1 line ($1\sigma = 64$ °C) they found no systematic offset between the two thermometers. The behavior of the thermometer at 1 bar was tested with an additional set of experiments, for which experimental temperatures were reproduced with $1\sigma = 22$ °C.

The thermometer was further tested and recalibrated with new experimental data by Coogan et al. (2014), extending the calibration range to higher fO_2 values (ΔQFM = −0.5 to ΔQFM = +1.3). They also tested the thermometer's dependence on a_{SiO_2}. They found fO_2 had no systematic effect on the performance of the thermometer and the effect of a_{SiO_2} was within the uncertainty of the thermometer, indicating that Al incorporation into olivine by vacancy formation (mechanism 2) is unlikely to be important in

most systems. This was corroborated by an experimental diffusion study by Zhukova et al. (2017) which found that Al incorporation by vacancy formation was favored only at higher a_{SiO_2} values than is found in most igneous systems where the thermometer is applied. Further improvements have been made to both the calibration and the mathematical formulation of the thermometer by Zhang and Namur (2022).

Despite the increased calibration range of the Coogan et al. (2014) model, many natural samples possess olivine and spinel pairs with compositions which still lie outside the calibration range. This includes the study reporting the highest equilibration temperatures from this method (~1570 °C, Trela et al., 2017), which have spinels with Cr_2O_3 contents higher than any of the spinel crystals used to calibrate the thermometer. Trela et al. justified such an extrapolation based on the global correlation between the Al_2O_3 K_D and spinel Cr#, and the fact that these melts also recorded extreme mantle potential temperatures and olivine liquidus Ts. However, it has also been suggested in other locations that application of this method to spinel crystals with much higher TiO_2 contents than the calibration dataset may be invalid, because these higher TiO_2 may affect the activity coefficient of Al in spinel (e.g., Jennings et al., 2019; Wong et al., 2022). Jennings et al. (2019) suggest that in the absence of a wider calibration range, it is best to apply the thermometer only to spinel crystals that are close to the calibration range.

To robustly assess how effectively the olivine-spinel method can be extrapolated, we need a suite of experiments with compositions lying outside the current calibration range. However, because application of the thermometer relies on precise measurements of the low concentrations of Al_2O_3 in olivine, generally only experiments performed for the purpose of calibrating this thermometer can be used, and all such experiments have been used during model regression (by Wan et al., 2008, or Coogan et al., 2014). Experiments conducted for other purposes cannot be used to formulate a test dataset, because Al_2O_3 concentrations in olivine were generally not measured at all, or were measured with low precision.

The lack of an independent test dataset also makes it difficult to robustly constrain the uncertainty of this method even when applied within the calibration range. Most studies cite the quoted standard error on the fit for the calibration data as being a minimum estimate of the uncertainty (Matthews et al., 2021, 2016; Wong et al., 2022), but as discussed above, statistics calculated on calibration datasets tend not to reflect the true error when applied to data not used for calibration. A final problem with this method is that there is no independent equilibrium test to assess whether the spinel and olivine are in equilibrium. The ubiquity of Mg-Fe diffusive resetting means these elements are an unreliable test of equilibrium (c.f. Prissel et al., 2016). The slow diffusion of Al in olivine (Spandler and O'Neill, 2010) means that parts of a host olivine crystal could be out of equilibrium with their spinel inclusions if olivine crystallization occurred over a protracted time with changing temperature (or melt composition). Maps of the aluminum content of olivine crystals have revealed near-ubiquitous zoning in crystals from Iceland (Matthews et al., 2021), adding complexity to identifying equilibrium pairs. Matthews et al. (2021) and Trela et al. (2017) therefore recommended the aluminum content of olivine crystals should be mapped with a high current electron beam before identifying locations for quantitative analysis.

Plagioclase-liquid thermometry and hygrometry

Plagioclase (Plag) is a very common mineral in a wide variety of tectonic settings (e.g., MORBs, OIBs, Arcs), motivating the development of a number of thermometers, barometers and hygrometers parameterizing Plag-liquid equilibrium (Putirka, 2008, 2005; Sugawara, 2001; Waters and Lange, 2015). However, the exchange of the anorthite (An)-albite (Ab) component between liquid and Plag is sensitive to T, P, and H_2O. If none of these variables are constrained by independent methods, there is a substantial solution space to explore. There also isn't much consensus as to what equilibrium tests should be used to filter Plag-Liq pairs. Putirka (2008) note that Ab-An exchange values for experiments (K_D^{An-Ab}) are normally distributed, with experiments with $T < 1050$ °C having values of 0.1 ± 0.05, and experiments with $T > 1050$ °C having values of 0.27 ± 0.1. However, plotting our experimental data (along with the calibration dataset of Waters and Lange, 2015) shows a more continuous variation of K_D with temperature (Supporting Fig. 9) with a relatively abrupt step up to higher values between 1000 °C and 1100 °C. We tentatively suggest this step up may result from the C1–I1 structural phase transition which occurs near this temperature, and has been shown to affect plagioclase Mg partitioning behavior (Mutch et al., 2022). Using the criteria of Putirka (2008) would exclude a number of experiments close to the cut off (Supporting Fig. 9)—We instead apply an exponential fit through the experimental data, excluding experiments outside ±0.11 (Supporting Fig. 9a and b). Not applying this filter affects calculated statistics very little for the following discussion.

First, we assess sensitivity of thermometers to H_2O (Fig. 16a), thermometers to P (Fig. 16b), hygrometers to T (Fig. 16c), and hygrometers to P (Fig. 16d), by independently perturbing experimental P, T and H_2O. It is apparent from Fig. 16a that Plag-Liq temperatures are strongly sensitive to H_2O; an increase of just 1 wt% H_2O causes an average drop in T of ~31 °C for eq24a and 36 °C for eq23. This change in calculated T is comparable to the RMSE of these thermometers. Thermometers are less sensitive to P; with the calculated change in T only reaching the same magnitude as the RMSE for a change of ~10 kbar (Fig. 16b). Notably, the Plag-Liq hygrometers of Masotta and Mollo (2019), Putirka (2008) and Waters and Lange (2015) are extremely sensitive to T; calculated H_2O contents drop more than the RMSE of the hygrometer for a change in T of just 10–20 °C (Fig. 16c). These hygrometers are less sensitive to P, with P changes of 10 kbar causing variations well within the quoted RMSE (Fig. 16d).

Before assessing hygrometers using our new dataset, it is worth considering how well we really know the "true" value of H_2O in each experiment. In the compiled Plag-Liq dataset, only 33% of experiments measure H_2O by FTIR, 5% by SIMS and 4.8% by Raman spectroscopy. A further 16% report H_2O using the EPMA-based water-by-difference method, with some mention of a calibration method. 17% have calculated H_2O using a solubility model for a pure H_2O fluid or using a measured X_{H2O} value in the fluid (Andújar et al., 2015; Costa, 2004; Mandler et al., 2014), or provide enough information for us to perform these calculations using MagmaSat in VESIcal (Ghiorso and Gualda, 2015; Iacovino et al., 2021). Interestingly, Waters and Lange (2015) calibrate

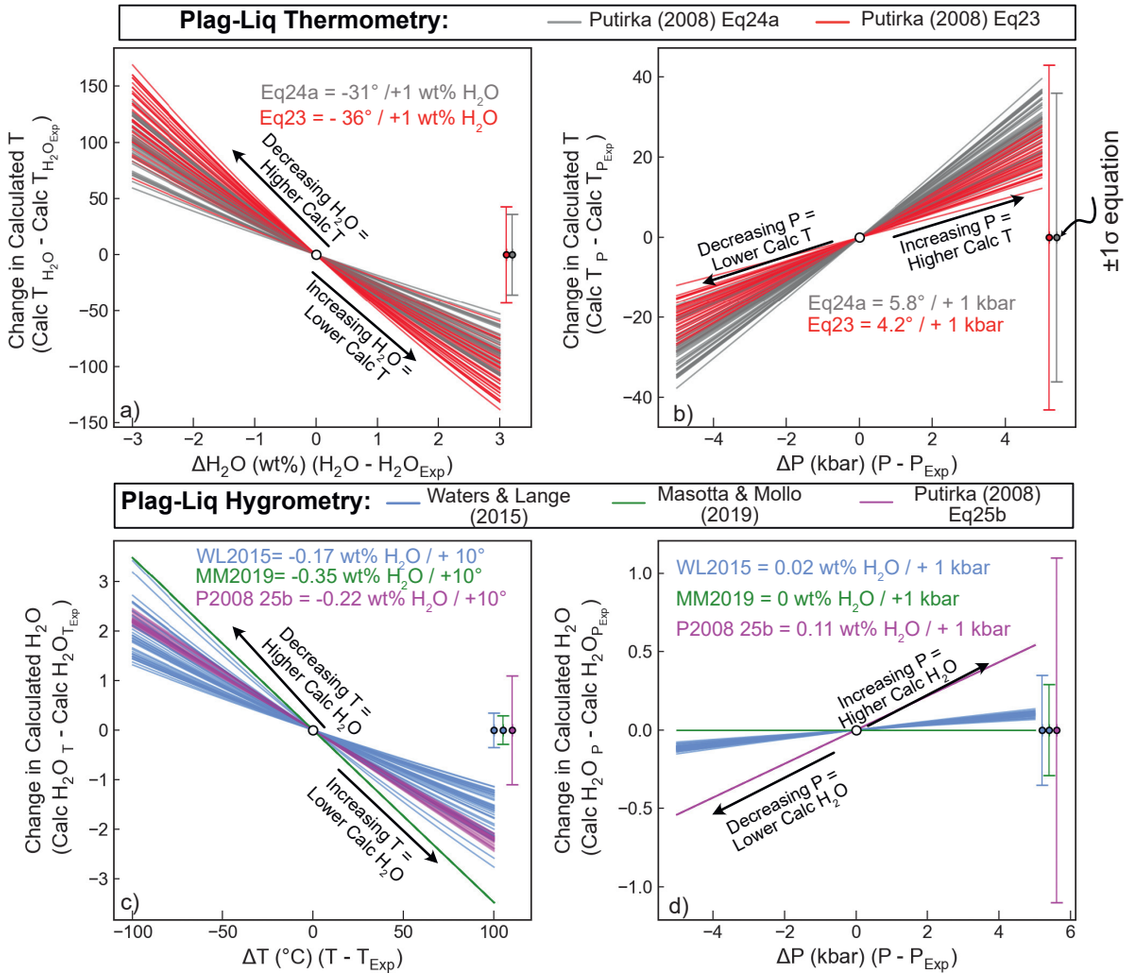

Fig. 16 Testing the sensitivity of Plag-Liq thermometers to H_2O (a) and pressure (b), and Plag-Liq hygrometers to T (c) and P (d). Forty experiments from ArcPL were randomly selected (each represented by a colored line). We perform calculations at the experimental P, T and H_2O content (0,0) on all plots, and then vary H_2O by ±3 wt% (a), P by ±5 kbar (b, d), and T by ±100 °C (c). We subtract the quantity calculated at experimental conditions from the quantity at these perturbed conditions. Error bars show the quoted RMSE on each expression. We label the average perturbation for these 40 samples on the figure.

their hygrometer by calculating H_2O in each experiment in their calibration dataset using the solubility model of Zhang et al. (2007), rather than using measured H_2O contents. If they had they used MagmaSat instead, calculated H_2O contents for these experiments would differ by an average of 0.3 wt% (the max discrepancy is 0.78 wt%), which would likely result in slightly different model parameters. Additionally, if the starting materials contained even small amounts of CO_2 (e.g., from contamination, Blatter et al., 2013), the amount of H_2O in the melt would not equal that calculated using a pure H_2O solubility model.

Ideally, we would restrict our comparison to experiments which performed FTIR, SIMS or Raman measurements of H_2O. However, this results in a much smaller dataset, with $N = 163$ vs. $N = 358$ if calibrated VBD and solubility water contents are included. This smaller dataset also has a very restricted T range (Supporting Fig. 10). It also seems unjustified to exclude experiments using solubility models, given that is how the H_2O contents were determined to calibrate the Waters and Lange (2015) model. Thus, we choose to proceed with the larger experimental dataset.

Using experimental T and P, the Waters and Lange hygrometer performs reasonably well on the ArcPL dataset, with a RMSE of ±1.2 wt% (Fig. 17a). There is a marked deviation to anomalously low calculated H_2O contents for experiments with $H_2O > 7$ wt% (cyan squares); if these are excluded, the RMSE is 0.94 wt%, although this is still far higher than the quoted RMSE of 0.35 wt%. The poor performance at high H_2O may reflect the fact that these super-hydrous compositions are poorly represented in the calibration dataset of this hygrometer (white histogram, Fig. 17a and b). When using experimental H_2O contents and pressures, the statistics reported for the thermometer of P2008 Eq24a on a global dataset are very similar to those estimated from our new dataset (we calculate RMSE = 32 °C vs. the stated RMSE = 36 °C shown as red dotted lines, Fig. 17c).

As discussed in Section "Statistics to compare models," the extreme sensitivity of Plag-Liq thermometers to H_2O content, and hygrometers to T, means that the comparisons shown in Fig. 17a and c do not accurately represent the true error when these equations are applied to natural systems, where in the vast majority of instances, neither H_2O nor T is known. To address this issue with many under constrained intensive variables, we investigate whether H_2O and T can be solved iteratively using a plagioclase hygrometer and thermometer, as an adaptation of the popular workflow of iterating barometers and thermometers. Arguably, as we perform these calculations using experimental P, we are still overestimating their performance on natural systems where H_2O, T and P are all unknown. The absence of a reliable plagioclase-liquid barometer (see Section "Plagioclase-liquid barometry") means that three-way iteration will not work. However, given these equations are far less sensitive to pressure (Fig. 16), uncertainty in pressure shouldn't hinder the model performance substantially.

Unsurprisingly, iteration of Plag-Liq hygrometers and thermometers yields worse statistics than hygrometry calculations performed using experimental T. For example, excluding experiments with $H_2O > 7$ wt%, the iterated RMSE is 1.8 wt% (vs. 0.94 wt% using experimental T; Fig. 17a vs. b). Calculated T using the iterative method are also substantially worse than those

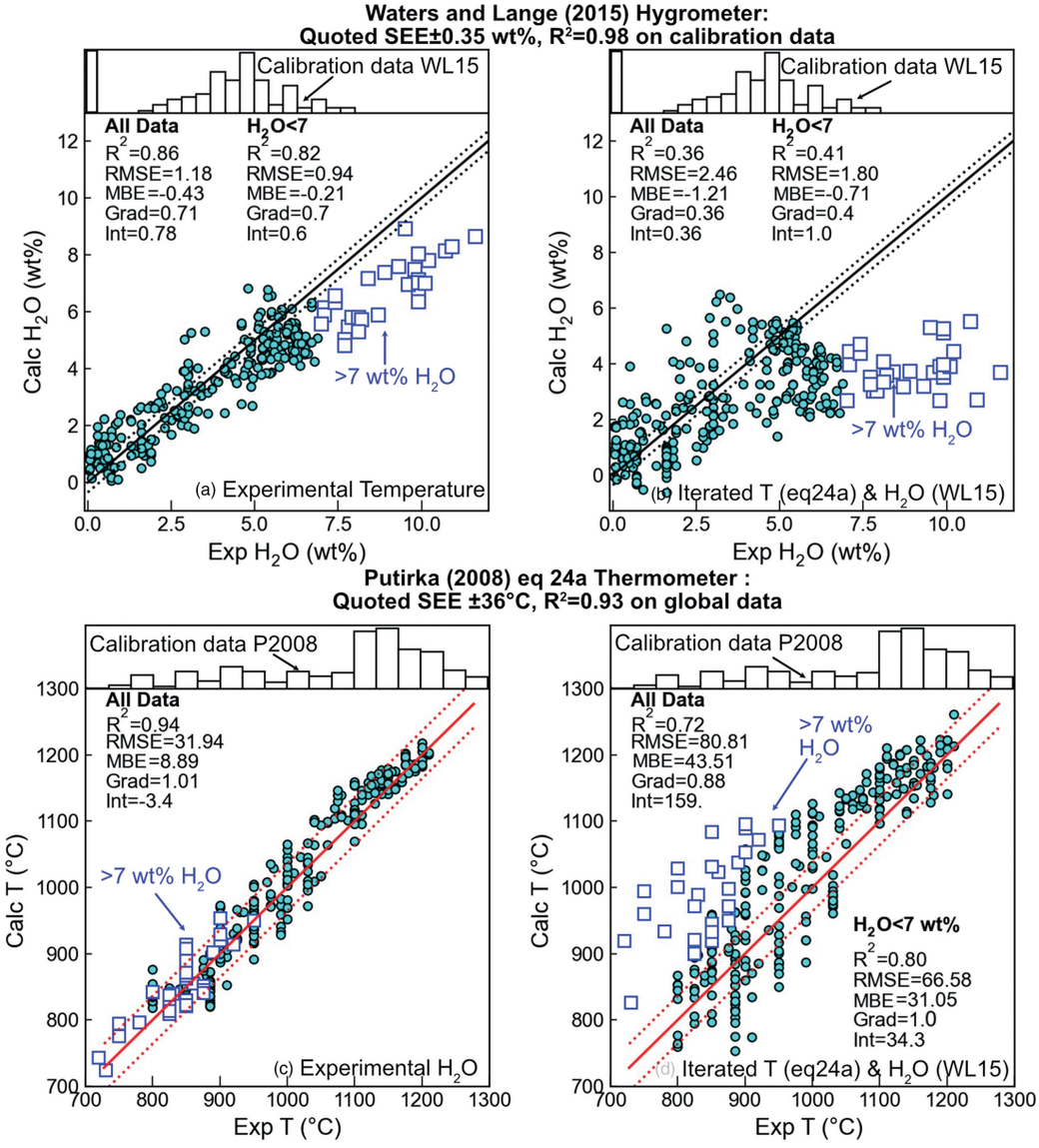

Fig. 17 Evaluating Plag-Liquid hygrometers (a and b) and thermometers (c and d). Only experiments with H_2O measured by quantitative methods are shown. (a) Calculations of H_2O using Waters and Lange (2015) with calculations performed using experimental T. The hygrometers performance drops substantially for $H_2O > 7$; the white histogram on top indicates that relatively few experiments used to calibrate this model had such high H_2O contents. (b) Iterative calculations using Waters and Lange (2015) and Putirka (2008) eq24a. (c) Calculations using Putirka (2008) eq24a and experimental H_2O contents, and (d) Iterative calculations using Putirka (2008) eq24a and Waters and Lange (2015). 1:1 line shown with ±stated RMSE for each expression.

obtained using known H_2O contents (RMSE = 81 °C vs. 32 °C for all data, RMSE = 67 °C for $H_2O < 7$ wt%, Fig. 17c and d). The statistics for these iterative calculations are more indicative of the sort of precision these methods can achieve in natural system. We also suggest that such an iterative approach may be more accurate than calculating H_2O using temperatures derived from other phases which may not have formed at the same temperature as Plag (e.g., Fe-Ti thermometry, Black and Andrews, 2020; Crabtree and Lange, 2011; Pineda et al., 2021).

The iterated thermometer has reasonable performance compared to other mineral-melt thermometers discussed, and the hygrometer has sufficient precision to distinguish dry (0–2 wt%), moderately wet (2–4 wt%) and wet (>4 wt%) lavas. However, when applied outside of the experimental products used for calibration, Plag-Liq hygrometry cannot achieve anything like the ≪0.5 wt% error often quoted for this method. Future improvements would be possible with a larger dataset of experiments where H_2O contents are known, as combining different methods for estimating H_2O contents in experiments undoubtedly adds uncertainty, particularly when using volatile contents calculated using volatile solubility models (see Wieser et al., 2022b). Additionally, given the relative success of iterating two different expressions with different underlying datasets, we suggest that recalibrating Plag-Liq hygrometers without a T term, but with compositional terms like those in Plag-Liq thermometers to incorporate the effect of temperature, may be more successful than having to iterate two independently calibrated expressions.

Plagioclase-liquid barometry

Putirka (2005) proposed a Plag-Liq barometer calibrated on 187 Plag-Liq pairs, which yields a RMSE of 1.8 kbar for the calibration dataset and RMSE of 2.2 kbar on $N = 292$ test data. However, Putirka (2008) re-evaluated this barometer using new experimental data, and found that it performed very badly, with a RMSE of 3.8 kbar excluding 1 atm data, and even worse statistics when this 1 atm data was included. They tried to find a global model to adequately predict pressures in their new dataset and found that while some regression worked on some subsets of the database, no regression could fit all data. They suggested that new experiments with the specific purpose of developing a Plag-Liq barometer are required to move forward. Despite their warning that "*the status of plagioclase-liquid as a barometer is firmly in doubt*," a concerningly large number of studies have performed Plag-Liq barometry after 2008 (e.g., Budd et al., 2016; Cheng et al., 2014; Dahren et al., 2012; Geiger et al., 2018, 2016a,b; Guo et al., 2018; Jamshidi et al., 2015; Siegburg et al., 2018). The majority of these studies quote a 2.47 kbar RMSE to justify this approach. However, this value from Putirka (2008) was only the fit to under half the data; the full dataset gave RMSE = 3.6–3.8 kbar.

As Plag-Liq barometers are still being widely used in the community, we briefly assess their sensitivity to T and H_2O, and then evaluate their performance on our new dataset. The T term in P2008 Eq25b means that the barometer is reasonably sensitive to T, with an average increase in +0.25 kbar per +10 °C increase in T (Fig. 18a). The barometer alone isn't sensitive to H_2O. However, this barometer is normally used in natural systems through iteration with the Eq24a thermometer, which is H_2O-sensitive (Fig. 18b). When iterated, the barometer is very H_2O-sensitive; an increase in H_2O by 1 wt% causes the P to drop by an average of ∼1 kbar (Fig. 18c).

We assess barometry performance on our test dataset, filtering out pairs which fail the anorthite-albite (An-Ab) equilibrium test provided in the supporting spreadsheet of Putirka (2008). Using experimental T, the barometer performs extremely poorly (RMSE = 3.8 kbar, R^2 of 0, Fig. 18d). When P and T are iterated, the performance is even worse (RMSE = 4.1 kbar, Fig. 18e). It is notable that in Fig. 18d and e that experiments performed at 2 kbar yield $P > 10$ kbar, while experiments performed at 10 kbar yield $P < 0$ kbar. To put the performance of these barometers into perspective, we compare each experimental P to a random experimental pressure drawn without replacement from the experimental dataset (Fig. 18f). The RMSE for this randomly selected pressure is almost identical to that of the iterated barometer. Thus, until new experiments are done to specifically investigate the Plag-Liq barometer, this method is only as reliable as researchers using a random number generator spanning the crustal thickness in their location of interest than Plag-Liq barometry to estimate magma storage pressures!

Fe-Ti oxides

The partitioning of Fe and Ti between cubic and rhombohedral oxides has been employed in estimating magmatic T and fO_2. The Fe-Ti geothermobarometer is advantageous owing to the rapid cation exchange between the oxides, allowing equilibrium to be restored quickly following a change in magmatic conditions (days to weeks; Venezky and Rutherford, 1999). The exchange is described by the following reaction:

$$\underset{\text{Ilmenite (rhombohed.)}}{FeTiO_3} + \underset{\text{Magnetite (cubic)}}{Fe_3O_4} = \underset{\text{Ulvospinel (cubic)}}{Fe_2TiO_4} + \underset{\text{Hematite (rhombohed.)}}{Fe_2O_3}$$

where there is solid solution between ilmenite and hematite (the rhombohedral oxides), and magnetite and ulvöspinel (spinel series, or cubic oxides). The dependence on oxygen fugacity is described by the redox reaction:

$$\underset{\text{Magnetite (cubic)}}{4Fe_3O_4} + O_2 = \underset{\text{Hematite (rhombohed.)}}{6Fe_2O_3}$$

The main challenge in calibrating a Fe-Ti geothermobarometer comes from the extremely complex solid solution of the rhombohedral oxides. Naturally occurring rhombohedral oxides incorporate significant quantities of MgO, MnO, and Al_2O_3 in addition to

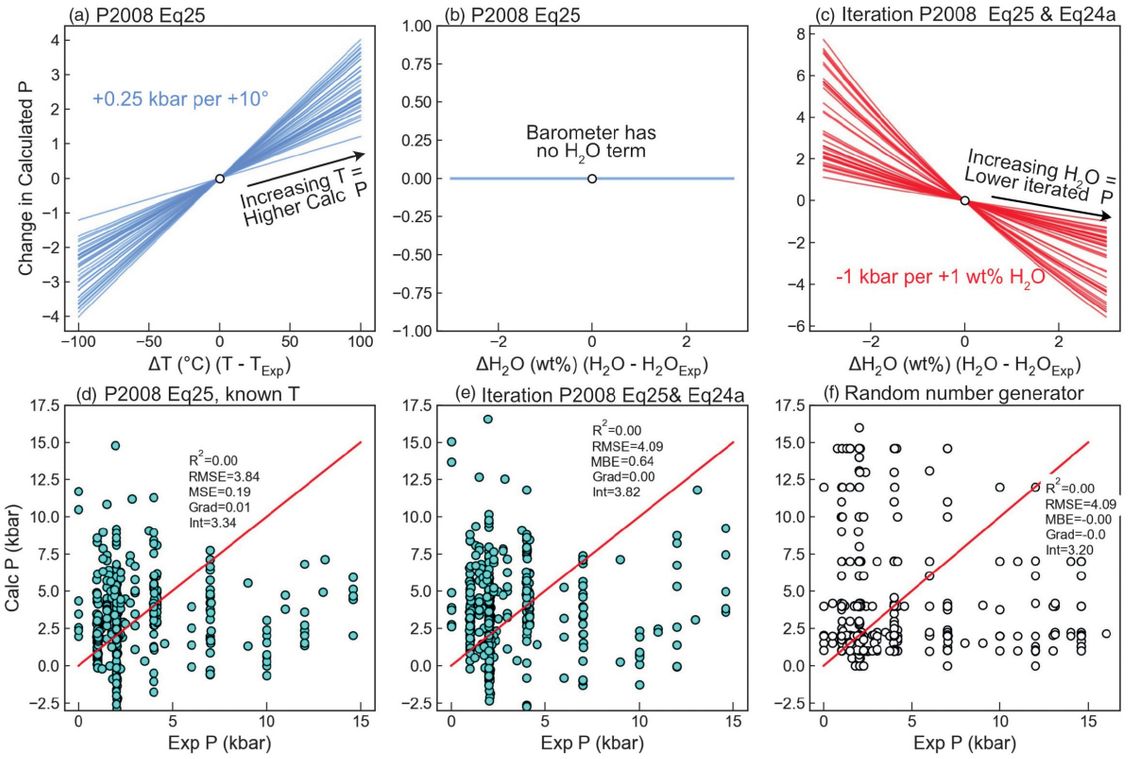

Fig. 18 Assessment of the Plag-Liq barometer of Putirka (2008) Eq25. (a and b) Assessing sensitivity of the barometer to T and H_2O for a subset of experiments. (c) Testing the sensitivity of the iterative combination of Eq25 (P) and Eq24a (T) to melt H_2O content. (d) Testing the barometer using experimental T and H_2O contents, after applying the An-Ab equilibrium test of Putirka (2008). (e) As for panel (d), but solving P and T iteratively (still using experimental H_2O). (f) Assessing the performance of a randomly selecting a Pressure value from the experimental dataset.

FeO, Fe_2O_3, and TiO_2, and have complex cation ordering transitions at lower temperatures, in addition to magnetic ordering. This means that complex expressions are required to accurately represent endmember activities, and multiple miscibility gaps exist. This complexity prevents simple empirical calibration of a geothermobarometer expression, and instead requires numerical solutions of expressions derived from a complex thermodynamic formulation.

The first implementation of the Fe-Ti geothermobarometer was made by Buddington and Lindsley (1964), and improved by a number of subsequent studies as further experimental data became available (e.g., Andersen and Lindsley, 1988; Spencer and Lindsley, 1981). The most recent update of the geothermobarometer was provided by Ghiorso and Evans (2008), building on the older thermodynamic model of Ghiorso and Sack (1991) using a calibration dataset with 5× more experiments ($N = 267$ vs. $N = 57$) and much better experimental constraints on cation ordering. In particular, the calibration range was extended to significantly higher T and fO_2, where previous versions of the model extrapolated poorly. Each update to the thermometer represents increasing sophistication of the underlying thermodynamic model, and therefore more complex numerical techniques to apply it as a geothermobarometer. An open-source implementation of the geothermobarometer was provided by Ghiorso and Prissel (2020) through the ENKI portal.

Blundy and Cashman (2008) use an independent set of experiments to estimate the uncertainty on several versions of the Fe-Ti geothermobarometer, finding one sigma uncertainties of 44 °C and 0.2 5–0.34 log units fO_2 for the Ghiorso and Evans (2008) model, but no systematic deviation at high or low temperatures. However, it is likely that the uncertainty will vary across composition space, as the sensitivity of the thermometer depends on the composition of the Fe-Ti oxides, with the best sensitivity below the NNO buffer and away from the miscibility gap (Ghiorso and Evans, 2008).

Ti in Quartz (TitaniQ) thermometer

In silicic rocks where quartz is a dominant phenocryst (e.g., granites and rhyolites), there are far fewer available thermometers relative to more mafic systems. To address this, Wark and Watson (2006) perform experiments containing quartz and rutile at 600–1000 °C at 10 kbar to produce an empirical relationship relating the Ti content of quartz to the temperature. They state that this thermometer has an uncertainty of ±2 °C at >500 °C using SIMS measurements of Ti, and that this thermometer can also be applied to systems without rutile if an independent estimate of Ti activity is obtained (e.g., from Fe-Ti equilibrium). This

thermometer was recalibrated by Kawasaki and Osanai (2008) using natural metamorphic rocks in ultrahigh temperature granulites and Thomas et al. (2010), who perform additional experiments between 5 and 20 kbar, generating an expression incorporating a term for pressure.

Importantly, Thomas et al. (2010) implied that if temperature was known independently, the equation could be inverted to solve for pressure. Wilson et al. (2012) test this inversion method on samples from the Oruanui eruption, which has been well studied, so has independent estimates of pressure from melt inclusions, and temperature and Ti activity from Fe-Ti oxides. They show that using Fe-Ti oxides to constrain temperature and Ti activity, calculated pressures from Ti in quartz are 3–10× higher than those inferred from melt inclusion saturation pressures. Similarly, if melt inclusion and Fe-Ti oxide temperatures and pressures are used, inferred Ti activity is far too low. If pressures are used from melt inclusions and Ti activity from Fe-Ti oxide, temperatures are well below the H_2O-saturated solidus for granite. Similar discrepancies with previously published pressures, temperatures and Ti activities are present for calculations on the Bishop tuff. Wilson et al. (2012) suggest that Ti activities may be highly variable in igneous systems, so activities from Fe-Ti oxides cannot be reliably used with TitaniQ. Additionally, they suggest Ti in Qtz records complex histories that cannot be simply related to changes in pressure and temperature over other variables, such as quartz growth conditions or melt composition. For example, Huang and Audétat (2012) show that Ti concentrations in Qtz depend on the crystal growth rate, so this thermobarometer should not be applied to hydrothermal fluids where growth rates are fast and highly variable.

Thomas and Bruce Watson (2012) partially rebut Wilson et al. (2012), in particular critiquing the validity of their calculations of Ti activity (e.g., Ghiorso and Gualda, 2013) and temperature estimates from Fe-Ti oxides (Ghiorso and Evans, 2008) in the Oruanui rhyolites that underly a lot of the arguments of Wilson et al. (2012). Instead, Thomas and Bruce Watson (2012) use MELTS to estimate temperature and Ti activity based on the affinity for rutile saturation from inputted melt compositions, yielding pressures similar to melt inclusions. They admit that the approach of Wark and Watson (2006) was oversimplified in its suggestion of using a fixed value of Ti activity. An excellent discussion of this Bishop Tuff controversy can be found in Putirka (2017), along with adjustment of published coefficients for different equations.

Acosta et al. (2020) identified that Ti in Qtz temperatures are offset ~100–150 °C to lower temperatures than other thermometers in silicic systems, particularly at <4 kbar. To address the source of these offsets, they perform hydrothermal quartz growth experiments at 800 °C and 1 kbar with different fluid compositions. They find that Ti in Qtz is sensitive to the Ti/Si ratio of the fluid, rather than the concentration or activity of Ti. Clearly, significantly more experimental work is needed to determine magma storage conditions precisely and accurately in silicic systems from Ti in Qtz.

Melt inclusion barometry

Melt inclusions (MIs) are small pockets of melt trapped during crystal growth, which become isolated from the external melt as the surrounding crystalline host fully encloses them. MI which were trapped from a volatile-saturated magma can be used to deduce magma storage depths because the solubility of CO_2 and H_2O in silicate melts is a strong function of pressure (Dixon, 1997; Goranson, 1931). If MI were trapped from a volatile-undersaturated magma, calculated pressures are minimum estimates (Hauri et al., 2018; Matthews et al., 2016).

Numerous recent reviews have detailed the theory, methods and advances relating to melt inclusion analysis and interpretation, as well as several specific problems associated with determination of magma storage depths from these archives. For example, Wallace et al. (2021) present a comprehensive review of olivine-hosted melt inclusions, describing melt inclusion formation, post-entrapment crystallization, and the wealth of information recorded by melt inclusions (e.g., trace element contents for tracking magma batches, processes controlling magmatic H_2O contents, calculating ascent rates, and storage pressures). Rose-Koga et al. (2021) provide a number of guidelines involving sample preparation, analysis and data reporting associated with melt inclusion analysis and interpretation. Barth and Plank (2021) discuss the processes which can alter H_2O contents in melt inclusions after their entrapment, producing regime diagrams describing how melt inclusions can act as hygrometers and barometers (revealing pre-eruptive H_2O contents and pressures), or speedometers (revealing ascent rates using H^+ diffusion). Wieser et al. (2023b) provide a detailed history of different solubility models used to calculate melt inclusion saturation pressures and highlight the large discrepancies between different models. In particular, they emphasize the importance of carefully examining the calibration range of each solubility model compared to the P-T-X range of melt inclusions from a given volcano. To avoid repetition, we refer readers to these papers, and focus our discussion on issues involving magma storage depth determinations from melt inclusions which have not yet been reviewed in detail. We specifically focus on the growing realization that the CO_2 contents of melt inclusions have been significantly underestimated, because of the presence of a substantial CO_2 within vapor bubbles that were not measured in most published studies.

Vapor bubble growth systematics

The vast majority of studies have used olivine-hosted melt inclusions to determine magma storage depths (e.g., Aster et al., 2016; Moore et al., 2015; Ruscitto et al., 2010; Sides et al., 2014a,b; Wallace et al., 2021). This focus on olivine may reflect the fact that it is one of the first crystallizing phases (important for studies focusing on the most primitive melt compositions), it is relatively abundant, and tends to have melt inclusions which are larger than those in other phases from the same sample suite (Bennett et al., 2019; Wieser et al., 2022a). It has also been suggested that the absence of cleavage in olivine makes leakage less likely than in say

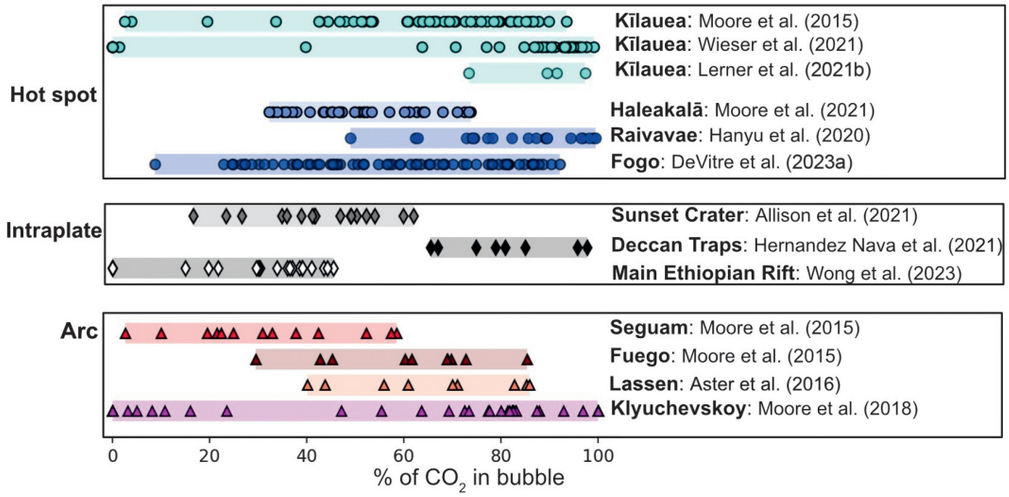

Fig. 19 Percent of the total melt inclusion CO_2 content held in the vapor bubbles from studies which have used an instrument-specific Raman calibration. Refs: Allison et al., 2021; Aster et al., 2016; DeVitre et al., 2023a; Hanyu et al., 2020; Hernandez Nava et al., 2021; Lerner et al., 2021a,b; Moore et al., 2021, 2018, 2015; Wong et al., 2023.

pyroxene, which has a strong cleavage (Kress and Ghiorso, 2004). Finally, the simple chemistry of olivine, well-constrained $K_{D,Fe\text{-}Mg}$, and low partition coefficients for REE and other trace elements of geological interest means that correcting melt inclusions for post-entrapment crystallization (PEC) is more straightforward than in plagioclase and pyroxene (Danyushevsky and Plechov, 2011; Kress and Ghiorso, 2004; Neave et al., 2017; Wieser et al., 2022a).

However, as a consequence of the fact that olivine is often the first phase to crystallize in a magma, there is significant potential for melt inclusions to experience substantial amounts of cooling prior to eruption (e.g., ~150–170 °C in high forsterite olivines from the 2018 eruption of Kīlauea, Lerner et al., 2021b; Wieser et al., 2021). This results in a large amount of PEC. Crystallization of denser olivine from less dense silicate melt, combined with differential contraction of the melt and host during cooling, causes the pressure in the inclusion to drop, driving the growth of a vapor bubble (often termed a shrinkage bubble, Kress and Ghiorso, 2004; Steele-Macinnis et al., 2011; Wallace et al., 2015). Because the solubility of CO_2 is strongly dependent on pressure, if there is sufficient time between bubble growth and *syn*-eruptive quenching, a significant proportion of the total CO_2 content of the melt inclusion will diffuse into the vapor bubble (Fig. 19, Maclennan, 2017; Wieser et al., 2021). Rapid diffusive re-equilibration of H_2O between the melt inclusion and a more H_2O-poor carrier melt as a result of fast H+ diffusion rates in olivine can also drive the growth of a vapor bubble (Aster et al., 2016; Gaetani et al., 2012).

Raman measurements of vapor bubbles

In situ Raman spectroscopic measurements of vapor bubbles over the last decade have demonstrated that a large and often dominant proportion of the total CO_2 content of melt inclusions is held within the bubble (Fig. 19, Allison et al., 2021; Aster et al., 2016; DeVitre et al., 2023a; Hanyu et al., 2020; Hernandez et al., 2018; Lerner et al., 2021a,b; Moore et al., 2015, 2018, 2021; Wieser et al., 2021). To our knowledge, the first Raman analyses of a melt inclusion vapor bubble were made by Steele-Macinnis et al. (2011) in samples from Solchiaro Volcano, Italy. They found a distinctive signal consisting of two strong peaks which is indicative of the presence of a CO_2 fluid (Fig. 20b). This contradicted the dominant hypothesis at the time that these bubbles were vacuums or voids. In two almost concurrent papers, Hartley et al. (2014) and Moore et al. (2015) present Raman measurements in melt inclusion vapor bubbles from Laki, Kīlauea, Fuego and Seguam, demonstrating that 40% to >90% of the total CO_2 is held within the bubble. More recent work has found similar proportions spanning a range of tectonic settings (Fig. 19).

Raman spectroscopic analyses of CO_2 in vapor bubbles rely on the strong correlation between the density of CO_2 and the distance between the two strong CO_2 spectral peaks collectively termed the Fermi diad. This distance is commonly called the splitting, diad splitting, or diad separation (Δ, Fig. 20b). However, the relationship between density and diad splitting has been shown to vary as a function of instrument hardware and acquisition parameters (Lamadrid et al., 2017, Fig. 20c). This means that the relationship between the diad splitting and CO_2 density must be determined for the acquisition parameters and specific data processing strategy used by each Raman laboratory. The absolute differences in CO_2 densities for a measured diad splitting on different Raman instruments are very large (Fig. 20c). For example, the vapor bubble shown in Fig. 20a yields a splitting of 103.10 cm^{-1} on the Cornell WITEC Alpha300R (Fig. 20b). Following the protocol of Lamadrid et al. (2017), this splitting is corrected based on the measured distance between two peaks from the atomic spectra of Ne to give a splitting of 103.197 cm^{-1}. Using the densimeter calibrated for this exact instrument, acquisition parameters and Ne correction regime (DeVitre et al., 2021),

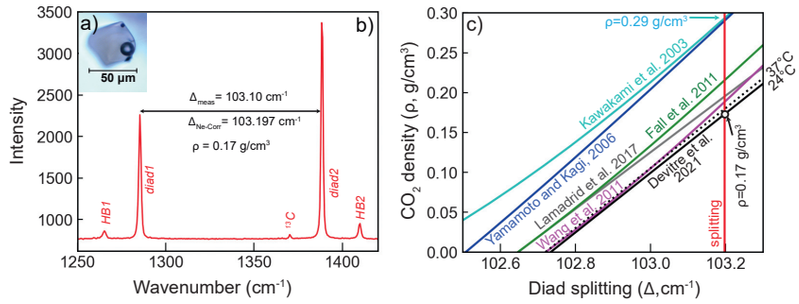

Fig. 20 Determining CO_2 density using Raman Spectroscopy. (a) Image of an olivine-hosted melt inclusion from the Twin Lakes Crater, OR. (b) Raman spectra showing the strong Fermi diad (peaks at ~1285 and 1388 cm^{-1}) with hot bands on either side. The distance between the peaks is the diad splitting ($\Delta = 103.10\ cm^{-1}$). After correction for the measured splitting of the Ne emission spectra (Lamadrid et al., 2017), this corresponds to a CO_2 density of 0.17 g/cm^3 using the splitting-density relationship developed on this specific instrument (DeVitre et al., 2021). (c) The relationship between density and splitting is different on each Raman instrument (e.g., DeVitre et al., 2021; Fall et al., 2011; Kawakami et al., 2003; Lamadrid et al., 2017; Rosso and Bodnar, 1995; Wang et al., 2019; Yamamoto and Kagi, 2006). If the Kawakami et al. (2003) densimeter was used for this vapor bubble, it would give a density of 0.29 g/cm^3.

the density of this bubble is $\rho = 0.17\ g/cm^3$ (Fig. 20c). However, if the densimeter of Kawakami et al. (2003) was instead used, the calculated density would be nearly twice as high ($\rho = 0.29\ g/cm^3$, Fig. 20c).

The most robust way to calibrate a specific Raman spectrometer for a given analytical protocol is to measure the diad splitting for ultra-pure CO_2 gas held at a variety of pressures (and thus densities). This is often achieved using a high-pressure optical cell (HPOC) or a fluid density calibration apparatus (FDCA). These apparatus feed pure CO_2 gas into a chamber where the pressure is tightly controlled, and ideally the temperature too. The measured P and T in the cell can be converted into a CO_2 density using the CO_2 equation of state (EOS, e.g., Span and Wagner, 1996). Raman measurements are made on this trapped fluid with known density, and the splitting is determined for each Raman acquisition. The relationship between the measured diad splitting and CO_2 density is often parameterized over a number of discrete density windows, as the shape of the curves vary (DeVitre et al., 2021).

Despite the importance of measuring CO_2 in melt inclusion vapor bubbles being recognized in ~2014–2015 (Hartley et al., 2014; Moore et al., 2015) and the calibration issue being highlighted in 2017 (Lamadrid et al., 2017), only published melt inclusion vapor bubble measurements made at Virginia Tech and Cornell have calibrated their Raman using a gas cell apparatus.

Only the Cornell calibration method closely controls and measures the T of the CO_2 gas directly, which is required to convert pressure into density with high accuracy and precision (DeVitre et al., 2021). A few more laboratories have used lower cost methods, where samples with known CO_2 densities are measured in a manner analogous to primary standards for EPMA calibration. For example, Allison et al. (2021) developed a calibration line by measuring the splitting-density relationships for fused silica capillary capsules (FSCC). The CO_2 density in these capsules was calculated from the mass of loaded CO_2 and the volume of the capillary, which allowed the calibration line to be determined at low densities (FSCCs ranged from 0.008–0.133 g/cm^3). Wieser et al. (2021) calibrated their Raman at low densities using 19 synthetic fluid inclusions (SFI) in quartz standards (0.04–0.14 g/cm^3) which were measured on the calibrated Virginia Tech Raman, with the same calibration being used by Wong et al. (2023). Mironov et al. (2020) use a similar approach, developing a calibration line using 8 melt inclusion vapor bubbles measured at Virginia Tech with densities between 0.013 and 0.22 g/cm^3.

Other studies measuring vapor bubbles have converted measured diad splitting into densities using a published calibration line developed in a different laboratory (e.g., Bali et al., 2018; Hartley et al., 2014; Taracsák et al., 2019; Venugopal et al., 2020b). The choice of densimeter in these studies varies widely, and the choice of one densimeter instead of another is not justified. For example, Hartley et al. (2014) using Kawakami et al. (2003, cyan line on Fig. 19c), Taracsák et al. (2019) and Venugopal et al. (2020b) using Wang et al. (2011, salmon line), and Bali et al. (2018) and Robidoux et al. (2018) using Fall et al. (2011, green line). The offset of their densities from the true value is not known, although a correction could be applied retrospectively if samples with known CO_2 densities were analyzed on the same instrument with the same analytical conditions.

Thermodynamic limits on the density of CO_2 at ambient conditions reveal a possible issue with the selection of a densimeter from the literature developed for a different instrument. The CO_2 phase diagram shows that at room T (20–26 °C), the maximum possible density of CO_2 gas is ~0.2–0.26 g/cm^3 (Fig. 21a). A vapor bubble with a higher bulk density will consist of an inner sphere of CO_2 gas of $\rho = 0.2–0.26\ g/cm^3$, and a coexisting outer shell of CO_2 liquid with $\rho > 0.7\ g/cm^3$ (see inclusion at Temp b on Fig. 21a). It is not thermodynamically possible for densities between these values to be measured by Raman spectroscopy unless the sample is heated above the critical point of CO_2 at 31 °C (supercritical CO_2 can have any density, Span and Wagner (1996)). While laser heating may occur during Raman analysis, which can account for the small number of measurements above 0.26 g/cm^3 on calibrated instruments (Fig. 21b, DeVitre et al., 2023b; Dubessy et al., 2012; Hagiwara et al., 2021), the presence of a significantly larger number of measurements with $\rho > 0.26\ g/cm^3$ in studies which did not perform an instrument specific calibration may indicate that these densities have been overestimated through selection of an inappropriate literature calibration. For example, had the Kawakami et al. (2003) densimeter been used for the melt inclusion in Fig. 20a, an impossible density of 0.29 g/cm^3 would have been obtained.

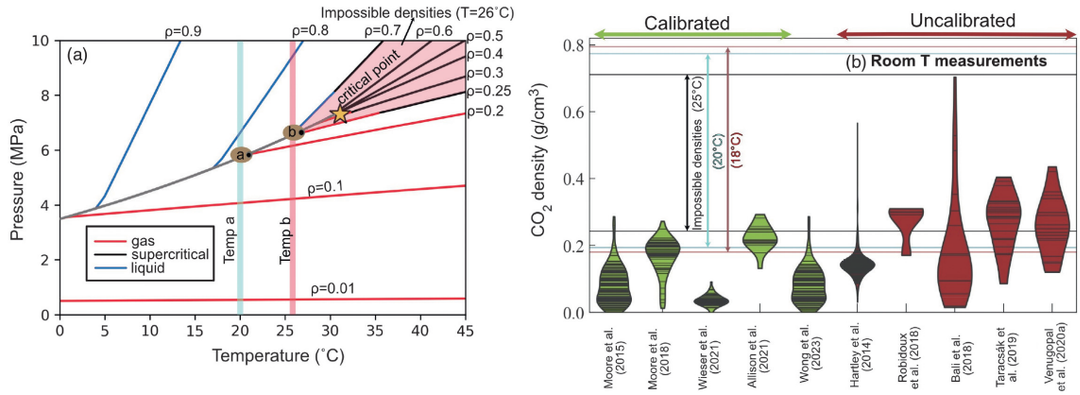

Fig. 21 (a) Phase diagram of CO_2, drawn using the NIST webbook with the Span and Wagner (1996) EOS. At Temp a (20 °C), a CO_2 fluid with $\rho > 0.2$ g/cm^3 will comprise of a vapor with $\rho = 0.194$ g/cm^3, and a liquid with $\rho = 0.773$ g/cm^3. At Temp B (26 °C), the liquid will have a density of $\sim\rho = 0.7$ g/cm^3 and the vapor will have a density of $\sim\rho = 0.25$ g/cm^3. Impossible densities at 26 °C are shown in pink. (b) Compilation of the densities of room T melt inclusion vapor bubble measurements by Raman spectroscopy, shown as a Violin plot where each horizontal line represents 1 measurement. Studies which used an instrument-specific calibration are colored green, those which did not are colored dark red. The maximum density of a vapor and minimum density of a liquid phase at three temperatures are shown with horizontal lines. Even with a room T of 25 °C, many of the densities reported by uncalibrated studies are thermodynamically impossible. Refs: (Allison et al., 2021; Bali et al., 2018; Hartley et al., 2014; Moore et al., 2018, 2015; Robidoux et al., 2018; Taracsák et al., 2019; Venugopal et al., 2020b; Wieser et al., 2021; Wong et al., 2023).

Once the relationship between measured splitting and CO_2 density for a specific Raman instrument and data reduction strategy is determined, the amount of the CO_2 held in the vapor bubble (in ppm equivalent in the glass) is calculated using mass balance:

$$CO_2(ppm\ equivalent) = 10^4\ \frac{Vol\%_{VB} \times \rho_{CO_2}}{\rho_{melt}} \tag{9}$$

Where $Vol\%_{VB}$ is the vol% of the vapor bubble, ρ_{CO_2} is the density of the CO_2 fluid (determined by Raman spectroscopy), and ρ_{melt} is the density of the silicate melt (e.g., using DensityX, Iacovino and Till, 2019). Eq. (9) can be used to demonstrate the large effect of the choice of densimeter on calculated CO_2 contents. For example, using a typical melt density (e.g., 2.7 g/cm^3) and a bubble volume percent of ~3.9% (as in the melt inclusion shown in Fig. 20a), the true vapor bubble CO_2 density of 0.17 g/cm^3 means the bubble contributes 2455 ppm CO_2 to the melt inclusion. However, if a literature densimeter was randomly chosen from those in Fig. 20b, the amount of CO_2 contributed by the bubble could be as high as 4188 ppm. These discrepancies in Raman calibration propagate to large uncertainties in magma storage depths. Wieser et al. (2021) show that for a typical Kīlauea melt inclusion with a vapor bubble occupying 5% of the inclusion volume, different densimeters could yield storage depths ranging from 4 to 18 km. Thus, constraining the splitting to CO_2 density relationship for each individual instrument is vital to avoid the introduction of very large systematic errors on calculated storage pressures (and therefore depths). Arbitrary choice of a Raman calibration is likely the largest source of systematic error in many published melt inclusion studies, only overshadowed by studies which didn't measure the bubble at all. Raman analyses without an instrument-specific calibration must be considered as qualitative. They are useful to determine whether CO_2 is present in vapor bubbles in any given system, but systematic errors spanning a factor of 3 are unacceptable when it comes to determining magma storage depths.

Other uncertainties reconstructing vapor bubbles

Once a Raman instrument is calibrated, there are four additional major sources of uncertainty affecting estimates of the amount of CO_2 held in vapor bubbles, discussed below.

Volume proportions of vapor bubbles

The first source of uncertainty is associated with determining the relative volume of the vapor bubble and melt inclusion (i.e., the $Vol\%_{VB}$ term in Eq. 9). The vast majority of melt inclusion studies estimate volumes using transmitted light images, where a best fit ellipse is fitted to the 2D outline of the melt inclusion and vapor bubble, and the third (z) dimension is estimated either as an average of the two visible axes, or the minimum of the two measured axes. Tucker et al. (2019) simulate the uncertainty associated with 2D sectioning of 3D ellipsoids, concluding that the best estimate is obtained when the third (z) dimension is calculated from the average of the 2 visible dimensions. They quantify the 1σ uncertainty of this method (−48% to +37%), which translates into a significant error in calculated storage depths where vapor bubbles contain a large proportion of the total inclusion CO_2.

However, in many cases, melt inclusions have faceted faces (e.g., Fig. 20a) or more complex shapes that can deviate significantly from a perfect ellipsoid. A best fit ellipse fitted to a more cubic shape may result in the melt volume being overestimated, and by extension, the bubble volume and CO_2 content underestimated (Hanyu et al., 2020). Mironov et al. (2020) compared methods of

calculating the z dimension using the measured x-y dimensions to direct measurements of the 3rd dimension using a vertically calibrated microscope, or by polishing two orthogonal faces so the x-y and z direction can be measured. They find that volume measurements using z from a vertically calibrated microscope or by polishing an orthogonal plane were in good agreement (±10%) with each other, but differed from the z-axis assumption methods by up to 45%. Thus, they suggest that researchers should measure the 3rd axis using these relatively low-cost methods, rather than inferring it from 2D measurements.

The most accurate and precise approach to determine the volume of each phase is to use x-ray tomography (i.e., nanoCT, μCT; Richard et al., 2019; Hanyu et al., 2020; Jorgenson et al., 2021). Hanyu et al. (2020) showed that optical methods fitting ellipses tend to overestimate MI volumes by ∼20%, even in fairly ellipsoidal shaped MI (they did not assess the effect of strong asymmetry or faceting). Attempts have been made to constrain volumes using 3D confocal Raman imaging (Aradi et al., 2021; Schiavi et al., 2020). There are additional compilations associated with the fact that the different refractive indices of the host crystal, melt and vapor can caused vertical distortion (Everall, 2010). This can make vapor bubbles appear oblate (Schiavi et al., 2020). Raman mapping is also significantly slower than μCT, and data segmentation is even more time consuming.

Most recently, DeVitre et al. (2023a) perform nano-CT scans for MI of various shapes and sizes, including those with extreme asymmetry or faceting. They compare these CT volumes to those calculated using 2D methods (measuring x-y, calculating z), and 3D methods where z is measured using a vertically calibrated microscope, or by polishing an orthogonal surface. For relatively ellipsoidal shapes, the median offset between 2D and CT methods is ∼ ±15–20% (similar to Hanyu et al., 2020), although for certain melt inclusion morphologies, the offsets can be ∼ ±50%, similar to the uncertainty estimated from slicing simulations of Tucker et al. (2019) of 1σ = −48% to +37%. For the most faceted or complex shaped MI, the systematic mis-prediction of volume inflicted using averaging or minimum axis 2D methods results in an uncertainty of up to ∼3 kbar in the saturation pressures (meaning MI could have been trapped at crustal levels at 7–8 km, or in the mantle, 17–18 km). For regular, relatively ellipsoidal-shaped melt inclusions, DeVitre et al. (2023a) find that measuring the 3rd dimension (via microscope, motorized-z stage Raman, or on an orthogonal plane) returns values within ±10% of CT measurements, although the offsets are larger for more irregularly shaped inclusions. Overall, DeVitre et al. (2023a) suggest that the orthogonal plane method is the most time-and cost-effective way to reduce uncertainty associated with bubble volumes. While CT is clearly superior, these measurements cost ∼100$/h, with high quality scans needing several hours. Data reduction is also time consuming and computationally expensive. At the moment, the CT method has not been scaled up to datasets typical of melt inclusion studies (e.g., $N > 100$) in a time or cost-effective manner. We suggest that it makes more sense to constrain the volumes of all melt inclusion to within ±20% of the true value for all melt inclusions (and published studies), compared with a smaller number of super precise CT scans, and a larger number of imprecise 2D optical methods. Developing fast, inexpensive, and precise ways to estimate vapor bubble volumes is an important frontier to address to optimize melt inclusion barometry.

Secondary phases in vapor bubbles

Another major source of uncertainty when reconstructing bubble CO_2 comes from the presence of carbon-bearing phases on the wall of the bubble, which may sequester 100s–1000s of ppm of CO_2 (Aster et al., 2016; Moore et al., 2015; Schiavi et al., 2020; Tucker et al., 2019; Venugopal et al., 2020b). These phases can be identified optically in larger bubbles (Tucker et al., 2019), and produce distinctive peaks in Raman spectra, particularly if the laser is focused near the bubble wall or line scans and 3D maps are used (Fig. 22d, Moore et al., 2015; Robidoux et al., 2018; Schiavi et al., 2020). Secondary phases can also be identified using BSE, SE and/or EDS imaging on an SEM of an exposed bubble walls (Robidoux et al., 2018; Schiavi et al., 2020; Tucker et al., 2019; Wieser et al., 2020).

The occurrence of secondary phases on bubble walls is highly variable, and their genesis is poorly understood. It appears that carbonate species (e.g., Mg, Fe, Na, and Ca carbonates) are more common in more H_2O-rich melt inclusions (e.g., arcs and wet OIBS like Fogo vs. dry OIBs such as Kīlauea). Even at a given volcano, some eruptions may have a reasonable proportion of melt inclusions containing carbonate (e.g., 1960 Kīlauea, Moore et al., 2015; Tucker et al., 2019), while other eruptions have no carbonate (e.g., 2018 Kīlauea, Lerner et al., 2021a,b; Wieser et al., 2021). Most intriguingly, even within a single crystal, some melt inclusions contain carbonate while others do not, and within a single melt inclusion with multiple bubbles, some have all their CO_2 as carbonate, some have a mix of carbonate and CO_2 fluid, and some consist only of CO_2 fluid (Fig. 23). It has been suggested that solid phases form on bubble walls as a paragenetic sequence during cooling of the vapor bubble, with sulfide precipitation at 500–700 °C, magnesite ($MgCO_3$) precipitation at <350 °C, and carbonate and sulfide precipitation from liquid H_2O at <150 °C (Robidoux et al., 2018).

Clearly, the carbonate forming reactions are subject to very local variations in condition (on the scale of the diameter of a single melt inclusion), and significant further work is required to understand them further. Quantifying the amount of CO_2 present within carbonate is non-trivial. Tucker et al. (2019) do preliminary mass balance calculations based on optical observations, considering different thicknesses of carbon on the wall of bubbles. Schiavi et al. (2020) determine the volume of carbonate and S-bearing species using 3D Raman mapping, showing that solid phases can account for 21–50% and 16–60% of the C and S budget respectively. Several experimental approaches have been developed for bubble-bearing melt inclusions as an alternative to mass balance reconstructions using Raman Spectroscopy (see below), some of which can help to resolve the carbonate problem.

Experimental homogenization approaches

Two broad homogenization strategies have been used to account for carbonate in vapor bubbles. The "in situ" strategy involves reheating of individual crystals in a Linkam or Vernadsky heating stage in an fO_2 controlled Ar or He atmosphere, making

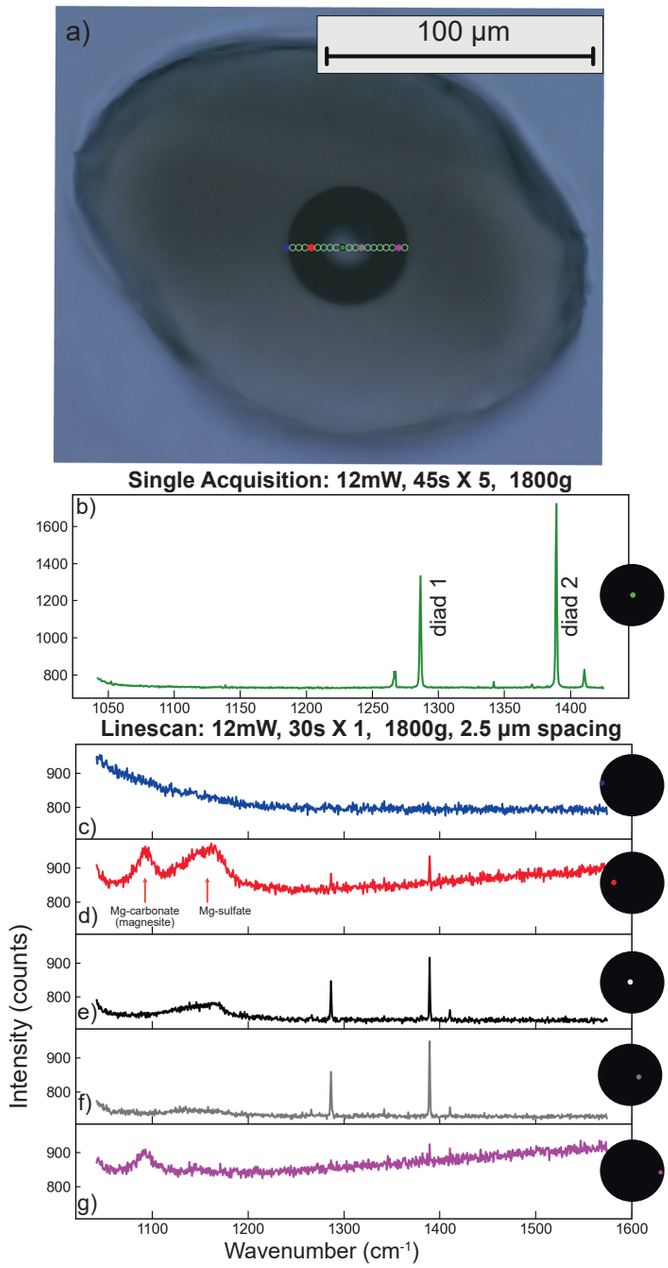

Fig. 22 Comparing a Raman acquisition in the center of a melt inclusion vapor bubble (a–b) with individual acquisitions taken as part of a line scan across the bubble (c–g). A carbonate and sulfate peak is only apparent at specific locations near the bubble edge, and does not appear in the central acquisition.

observations of the melt inclusion of interest under an optical microscope. The crystal is heated until the vapor bubble and other secondary phases redissolve, and then the sample is rapidly quenched to obtain a single-phase melt inclusion. This method allows the operator to view the melt inclusion and quench the crystal just after the inclusion homogenizes, meaning a different reheating path can be used for each crystal. The "bulk" strategy heats crystals in batches in various experimental apparatus. This had the advantage of being far less time consuming and allowing reheating at different pressures (see below), but has the disadvantage of meaning crystals trapped at a range of temperatures may be overheated or underheated (e.g., Esposito et al., 2012).

Regardless of whether a single crystal or a batch of crystals is heated, performing reheating experiments at atmospheric pressure often results in vapor bubbles not fully redissolving, or dissolving at spuriously high temperatures. For example, Le Voyer et al. (2017) individually heat melt inclusions from Mt. Shasta in the Cascade Arc using a Vernadsky-type heating stage, noting that many inclusions retain their vapor bubble, even if they are heated up to 1500 °C. The failure to homogenize the bubble has been

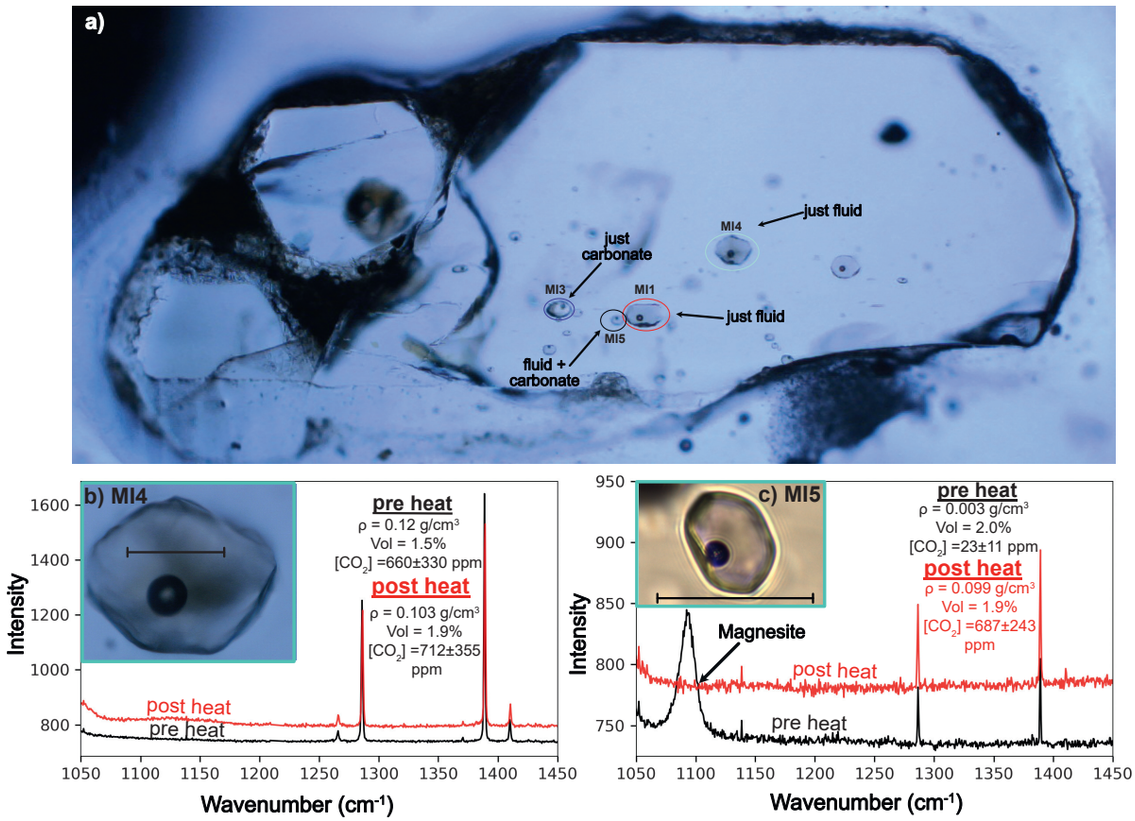

Fig. 23 Olivine crystal with numerous melt inclusions with variable carbon partitioning between fluid and solid phases. (a) Two melt inclusions (MI1, MI4) have a strong fermi diad but no carbonate phases (see spectra b), while MI5 has a fermi diad and a carbonate peak (see spectra c), and MI3 has just a carbonate peak (b) pre and post-heating, MI4 has very similar densities and volumes, with CO_2 contents within error. (c) pre-heating, MI5 has a prominent magnesite peak, and weak Fermi diads. After reheating using the method of DeVitre et al., 2023a, the amount of CO_2 increases substantially, and the magnesite peak disappears. Scalebar on b and c shows 50 μm.

attributed to the fact that at atmospheric pressure, the host exerts less pressure on the melt inclusion than was present during natural cooling at depth in the magmatic system (Danyushevsky et al., 2002; Esposito et al., 2012). Additionally, if sufficient time passes between post-entrapment crystallization and *syn*-eruptive quenching, chemical changes such as FeO or H^+-loss through the host crystal may occur, which cannot be reversed by reheating (Aster et al., 2016; Bucholz et al., 2013). To account for the excess bubble growth resulting from these irreversible changes, the inclusion must be overheated, often by many hundreds of degrees to get the bubble to disappear. This is concerning because overheating results in extensive re-equilibration between the melt inclusion and host, erasing important chemical information (e.g., major element zoning preserving quench rate, Newcombe et al., 2014; information on the amount of PEC; Wieser et al., 2021). An additional problem with heating at 1 atm is that melt inclusions may rupture/decrepitate (Tuohy et al., 2016). This is a particular problem for melt inclusions closer to the polished surface or the crystal, inclusions trapped at high pressures in the plumbing system, and more H_2O-rich inclusions.

Crystals can also be heated at elevated pressures (e.g., Piston cylinder apparatus, Rasmussen et al., 2020, internally heated gas pressure vessel, Skirius et al., 1990). The confining pressure exerted around the crystal can help prevent decrepitation, and aid bubble dissolution. To get around the issue that reheating can cause H^+ loss, and to try to resorb bubbles which have experienced extensive H^+ loss, Mironov et al. (2015) homogenize melt inclusions in an internally heated pressure vessel (IHPV) under high H_2O pressure achieved through a hydrated silicate melt matrix. Specifically, they heat melt inclusions from lavas which are thought to have diffusively lost 3 wt% H_2O (such extensive H_2O-loss favors near-ubiquitous bubble growth). They find that bubbles remain after heating under dry and damp conditions, but a large proportion disappear at higher H_2O pressures where the melt inclusion is rehydrated. This indicates that, at least in wet arc magmas, diffusive H_2O loss can account for the observation that bubbles often persists after heating (rather than resulting from a lack of confining pressure during heating in 1 atm apparatus). This reheating method was adapted by Rasmussen et al. (2020), who use a piston cylinder apparatus with KBr and $Mg(OH)_2$ as the hydrated matrix to achieve a higher experimental success rate.

One limitation of these hydrated experimental reheating methods is that the initial H_2O content of the system must be estimated when preparing the hydrated matrix, and melt inclusion water contents will be reset to this value (Buso et al., 2022). While

Rasmussen et al. (2020) use unheated melt inclusion volatile contents as a guide, there is still the risk of adding too much or too little H_2O to any specific inclusion. Additionally, while an approximate estimate of initial H_2O contents can be placed based on the point of homogenization, there is a non-unique play off between the confining P, T and H_2O content required to obtain complete homogenization of the bubble (e.g., Mironov et al., 2015 show that instead of increasing H_2O by 1 wt%, P could be increased by 6.5 kbar, or T increased by 100 °C). Thus, as with the techniques mentioned above, a limitation of these bulk homogenization experiments is that the same experimental conditions must be applied to all melt inclusions in a given capsule, even if they formed on different P-T-H_2O paths in nature (Rasmussen et al., 2020).

All the methods discussed so far attempt to redissolve the bubble entirely so measurements only need to be performed on a single homogenous glass phase. In contrast, DeVitre et al. (2023a) developed a reheating method where the aim is to redissolve the carbonate phase back into the vapor bubble where it can be measured using Raman spectroscopy, rather than redissolve the bubble into the melt. They heat individual crystals using a Linkam TS1400XY stage which heats from room T to 1400 °C at up to 200 °C/min, and quenches using Ar or He flow combined with a water-cooled plate at 240 °C/min. Heating is conducted in an fO_2 controlled Ar atmosphere to prevent oxidation of the melt and olivine. As the melt inclusion is progressively heated under a microscope, it passes through the glass transition temperature, and goes dark. The liquidus temperature is estimated from the major element composition of the system and a reasonable range of volatile contents (using MELTS, or a liq thermometer). As the liquidus temperature is reached, the melt inclusion clears to a brown glass, and the inclusion is held at this temperature for ~8–10 min to redissolve any carbonate on the walls of the bubble back into the fluid phase, before being rapidly quenched on a water-cooled plate. The advantage of this method is that it only heats the inclusion to the liquidus, rather than to higher temperatures needed to resorb the bubble. Heating to the lowest possible T is advantageous, because it limits olivine dissolution, and preserves the major element information held in the melt and olivine. Raman acquisitions collected before and after heating indicate that the carbonate is effectively redissolved back into the bubble (Fig. 23e), and that inclusions experience minimal H_2O-loss. After heating, bubbles without carbonate return CO_2 contents within uncertainty of pre-heating estimates (CO_2 contents change by $\sim \pm 10\%$ accounting for slight changes in bubble volume, or errors associated with measuring the same bubble proportion using 2D images). The main disadvantage of this method over complete homogenization methods is the fact that Raman is used to measure the bubble after heating, which requires accurate estimates of the relative volume of the vapor and melt phase.

Theoretical vapor bubble reconstruction methods

A variety of methods have also been developed to reconstruct the CO_2 content of a vapor bubble theoretically, although there have been relatively few comparisons with Raman approaches on a single inclusion basis (e.g., Aster et al., 2016; Wieser et al., 2021). Anderson and Brown (1993) investigated bubble CO_2 in a suite of melt inclusions from Kīlauea Iki, calculating an internal pressure for each melt inclusion using the measured CO_2 content in the melt phase of the inclusion. This internal pressure was then used to calculate the density of CO_2 in the coexisting vapor phase using an equation of state. To convert these densities into CO_2 amounts, they assumed all bubbles occupied 0.5 vol% prior to *syn*-eruptive quenching, which is the point at which CO_2 diffusion and bubble growth become decoupled. Riker (2005) adapt this method for Mauna Loa, calculating the pre-quench bubble volume as a factor of the T drop experienced by each melt inclusion, accounting for the fact different melt inclusions experience different amounts of cooling, and therefore PEC. Aster et al. (2016) further adapt this method for melt inclusions from Lassen, tracking the volume of a growing vapor bubble using phase volume and density information from Rhyolite-MELTS, and vapor compositions to partition elements into the bubble using the volatile solubility model of Iacono-Marziano et al. (2012). This method was also used in the Cascade Arc by Johnson and Cashman (2020) and Walowski et al. (2016).

In contrast to these methods reconstructing the vapor bubble volume *pri*or to bubble expansion accompanying syn-eruptive quenching, Tucker et al. (2019) use the equation of state method outlined by Anderson and Brown (1993) to calculate CO_2 density in the bubble, and then calculate the amount of CO_2 in the bubble using the measured bubble volume. This measured-volume method generates extremely high CO_2 estimates (and therefore storage pressures) for Hawaiian melt inclusions (4000–10,000 ppm). Using the measured bubble volume assumes that the vapor bubble and melt continue to exchange CO_2 until the glass transition T (~725 °C), such that the bubble is always in equilibrium with the measured melt composition. However, vapor bubbles experience two distinct phases of growth. The first phase of bubble growth at high T accompanies PEC or diffusive H_2O loss. High temperatures mean that CO_2 can easily diffuse from the melt into the growing vapor bubble. The second phase of bubble growth occurs during syn-eruptive quenching, where CO_2 becomes diffusion-limited as the temperature drops, but the bubble volume continues to grow until the temperature cools below the glass transition temperature (Maclennan, 2017). Wieser et al. (2021) compare the Tucker EOS method to Raman measurements at Kīlauea and show that the Tucker method overestimates bubble CO_2 by a factor of 10–20× compared to the Raman method for melt inclusions which grew most their bubble during syn-eruptive quenching (where CO_2 migration was diffusion limited). The measured-volume EOS method results in calculated magma storage depth of 5–20 km for melt inclusions while the Raman method and geophysics indicate magmas were stored at 1–5 km. For melt inclusions contained in high-Fo olivines, where most of the bubble grew during PEC at high temperatures, EOS methods are still 1.5–2× too high, as the method still neglects a non-negligible increase in bubble volume upon quench which is not accompanied by CO_2 diffusion.

To unravel the relative importance of these two stages of vapor bubble growth, good estimates are required of the amount of PEC, the amount of H_2O-loss, the quench rate, and the glass transition T (Maclennan, 2017; Rasmussen et al., 2020). Rasmussen et al. (2020) produce a Python3 tool, MIMiC (Melt Inclusion Modification Corrections), which calculates bubble CO_2 using empirical parametrizations of volume and density changes, with uncertainties quantified by Monte-Carlo techniques. However,

while constraining the amount of PEC is relatively straightforward in systems with a well-defined liquid line of descent where the initial FeO content can be easily estimated (e.g., Kīlauea *V*olcano, Wieser et al., 2021), it can be very challenging in systems with a large amount of scatter in FeO at a given MgO number (e.g., Rasmussen et al., 2017; Walowski et al., 2019), perhaps because of variability in primary FeO contents and/or mixing of diverse melt compositions (e.g., Maclennan, 2008). Reconstructing initial H_2O contents can also be very challenging, and the arc magmas at the point of melt inclusion entrapment may be substantially more hydrous than the amount of H_2O measured in the melt inclusion (Gavrilenko et al., 2019; Goltz et al., 2020).

Co-entrapped vapor bubbles

All the methods discussed thus far rely on the assumption that vapor bubbles formed after melt inclusion entrapment, through some combination of PEC, cooling, and H^+ loss. However, it has been long recognized that bubbles may also become trapped at the point of melt inclusion formation. These are termed co-entrapped bubbles. In fact, determining magma storage depths from melt inclusion saturation pressures *requires* that the melt was volatile saturated at the point of melt inclusion formation. This assumption necessitates that there is an excess volatile phase available to co-entrap. If co-entrapped bubbles are added back in by mass balance techniques, the amount of CO_2 could be drastically overestimated.

Many studies simply select a threshold volume above which bubbles are assumed to be co-entrapped (e.g., 5%, Lowenstern, 2003; Robidoux et al., 2018, or 10%, Buso et al., 2022; Moore et al., 2015). However, the maximum vapor bubble volume that can form without requiring co-entrapment is a function of the amount of cooling, and the amount of diffusion H^+ loss, and varies greatly between different eruptions (Tucker et al., 2019). Tucker et al. (2019) examine distributions of bubble volumes from Hawai'i, discarding outliers which clearly lie outside the main distribution of bubble volumes (8 vol%+). In suites where there are abundant fluid inclusions in crystals addition to bubble-bearing melt inclusions, co-entrapped bubbles can often be identified based on the fact that they have larger relative volumes than other vapor bubbles, and CO_2 densities more similar to fluid inclusions (Lerner et al., 2021a,b; Hanyu et al., 2020).

Alternatively, co-entrapped bubbles can be identified by comparing bubble volumes to variables such as H_2O content, CO_2 density, and PEC amount. For example, Wieser et al. (2021) show that there is a correlation between the amount of PEC (0–33%) and the volume of the vapor bubble (1–6 vol%) at Kīlauea. The 5% cut off from Lowenstern (2003) would clearly be inappropriate in this instance. Instead, samples lying significantly off the observed PEC-volume trend were classified as co-entrapped. While bubble growth in H_2O-poor, CO_2-rich systems can be relatively well predicted based on the amount of PEC and the cooling path alone (Maclennan, 2017; Riker, 2005; Wieser et al., 2021), it is more complicated in H_2O-rich magmas which have the potential to experience large amounts of bubble growth following H_2O loss. In these situations it is more robust to use bubble growth models, and any bubbles which exceed even the most extreme model scenarios were likely co-entrapped (e.g., Allison et al., 2021; DeVitre et al., 2023a). For example, Ruscitto et al. (2011) report the presence of bubbles with volumes spanning 6–16 vol% in high Mg olivines from Mt. Shasta, and conclude that these bubble volumes can grow through a combination of PEC and H_2O loss, so were not necessarily co-entrapped. In contrast, Allison et al. (2021) found that melt inclusions at Sunset Crater, Arizona, with bubbles >3.5 vol% were likely to be co-entrapped. These two contrasting scenarios highlight the issue with selecting a single universal volume % as a cut off.

Even using bubble growth models, the details of specific inclusion must be considered, else co-entrapped bubbles with volumes within the range predicted by the growth model could be overlooked, even if they were co-entrapped. Allison et al. (2021) subdivide bubble-bearing melt inclusions from Sunset Crater, AZ, into two groups based on bubble volumes, densities, offset trends in Olivine forsterite-total CO_2 space, and bubble growth models. They conclude that group 2 inclusions (>3.5 vol%) co-entrapped a bubble. To more robustly identify co-entrapped vapor bubbles, we are in desperate need of a forward-model of bubble growth, accounting for both H_2O and CO_2. The model of Maclennan (2017) only considered CO_2 (not H_2O) and is not publicly available. While the Monte-Carlo methods incorporated in the model of Rasmussen et al. (2020) have huge potential for this problem, currently this code only works to correct melt inclusions, and cannot be run forward for a hypothetical P-T-X path post-entrapment.

In summary, it is becoming very clear that the CO_2 contents of bubbles must be accounted for to obtain reliable storage depths, and that as a community, substantially more work is required to (a) develop reliable bubble growth models to ensure co-entrapped bubbles don't result in spurious CO_2 contents (b) calibrate individual Raman instruments and perform measurements at >33 °C and appropriate laser powers (c) improve methods to determine relative volumes (d) further investigate experimental homogenization methods.

Decrepitation

Another issue affecting melt inclusion saturation pressures is the process of decrepitation, where the internal pressure of the inclusion exceeds the strength of the host mineral, and cracks open. Decrepitation may occur during magma ascent, as the confining pressure exerted by the surrounding liquid drops. Maclennan (2017) compile a global dataset of melt inclusions from MORBs, OIBS, and continental settings, noting that 95% yield saturation pressures <2 kbar, which is close to the experimental decrepitation threshold of Wanamaker et al. (1990). Using their model of P-T-V-X evolution of melt inclusions, they conclude that pressure difference between olivine and melt causes decrepitation in most tectonic settings, so preserved CO_2 contents are minimum estimates. They note that decrepitation can be partially mitigated if the melt inclusion is trapped from a significantly undersaturated melt, so the melt inclusion experiences significant cooling and PEC at depth before it ascends, which reduces the internal pressure of their inclusion. Similarly, sequestration of CO_2 in a vapor bubble can also help to reduce the pressure of the liquid below the

decrepitation threshold. The modeling of Maclennan (2017) suggested that entrapment pressures up to twice the decrepitation threshold (i.e., ~4 kbar) could be preserved if ascent conditions allow for bubble formation while remaining below the decrepitation threshold.

However, a number of studies have been published since 2017 yielding significantly higher saturation pressures than those in compilation of Maclennan (2017), and their inferred decrepitation thresholds. In the glass phase alone, DeVitre et al. (2023a) measure CO_2 contents of up to 1.3 wt%, corresponding to glass-only saturation pressures of ~1–7 kbar. Once CO_2 in the vapor bubble is included, pressures rise to ~2–11 kbar. Similarly high glass CO_2 contents (1.2 wt%) are reported by Buso et al. (2022) in melt inclusions from the French Massive Central, corresponding to entrapment depths of >10 kbar. Once bubbles are dissolved using homogenization methods, saturation pressures in these samples extend to 15–25 kbar. Glass-only measurements from Haleakala, HI, cluster at 1–3 kbar, and extend to 2–6 kbar after accounting for the bubble (Moore et al., 2021). Melt inclusions from Isla Floreana in the Galápagos record glass-only saturation pressures of 1.1–7.2 kbar (median = 5 kbar, bubbles not measured). Critically, one bubble-free inclusion records a glass-only saturation pressure of ~7 kbar, where the lack of decrepitation cannot be attributed to the presence of a bubble keeping the inclusion below the decrepitation threshold. This high pressure overlaps with the pressure calculated from Cpx-based barometry, indicating it is not an outlier (Gleeson et al., 2022, 2021). There are numerous other examples of saturation pressures calculated from the glass phase only yield saturation pressures >2 kbar from many alkali ocean island settings (e.g., French Polynesia, Hanyu et al., 2020; El Hierro, Taracsák et al., 2019), and studies accounting for the glass and bubble exceeding 2–4 kbar (e.g., Sunset Crater, AZ, at 3–5 kbar, Allison et al., 2021, Deccan Traps between ~2–7 kbar, Hernandez Nava et al., 2021, Pico Volcano, Azores, 2.2–6.6 kbar, van Gerve et al., 2023).

The existence of melt inclusions with glass-only saturation pressures of ≫2–4 kbar and glass + bubble saturation pressures of >4 kbar suggests decrepitation may not be as important a process as Maclennan (2017) suggests. While Maclennan (2017) conclude that "*decrepitation of melt inclusions, where the inclusion ruptures and loses CO_2 to the external melt, is the dominant process that controls the observed distribution of CO_2 in the compiled data set.*" We suggest that the paucity of alkaline, CO_2-rich settings in the compilation of Maclennan (2017), combined with the fact most studies did not measure the vapor bubble, led to a dataset that was skewed to anomalously low pressures.

Melt inclusions hosted in other mineral species

While storage depths calculated from olivine-hosted melt inclusions have dominated the literature in the past few decades, there are a growing number of studies measuring melt inclusions in plagioclase (Bennett et al., 2019; Blundy et al., 2010; Drignon et al., 2019; Koleszar et al., 2012; Neave et al., 2017; Wieser et al., 2022c), ortho- and clinopyroxene (Araya et al., 2019; Koleszar et al., 2012; Wieser et al., 2022c), and amphibole (Koleszar et al., 2012). In more silicic systems, quartz-hosted inclusions are also commonly examined (Bégué et al., 2015; Quinn, 2014; Wallace et al., 1999), along with plagioclase (Bacon et al., 1992; Wright et al., 2012) and ortho- and clinopyroxene (Wright et al., 2012). Saturation pressures from inclusions in other phases often overlap with those determined from olivine (Wieser et al., 2022c), and may even yield deeper saturation pressures (Bennett et al., 2019) indicating that the "rupture resistance" of these other mineral phases may have been underestimated. Also, to gain an unbiased understanding of storage condition in a wide range of melt compositions, it is certainly advantageous to consider more than one mineral phase, particularly as in many systems olivine is restricted to a relatively narrow range of melt compositions (if present at all). However, significantly more work is needed to understand bubble growth and post-entrapment processes in these other phases, as well as rates of diffusive H_2O-loss.

Fluid inclusion barometry

Crystals growing in a fluid-saturated magma trap pockets of melt and exsolved fluid in varying proportions, varying from pure melt (melt inclusions), melt and varying proportions of fluids (melt inclusions with co-entrapped vapor bubbles), and pockets of pure fluid with little or no attached melts (fluid inclusions—FIs, Steele-Macinnis et al., 2011). In CO_2-rich, H_2O-poor volcanic systems such as mid-oceanic ridge basalts and ocean island basalts (OIBs), the exsolved vapor phase (and thus the fluid being trapped) is almost pure CO_2 at pressures >200 bars (Gerlach and Graeber, 1985). In these relatively anhydrous systems, the density of FIs can act as an excellent barometer, because the density of a CO_2-rich fluid, along with an estimate of its entrapment temperature, can be converted into a pressure using the pure CO_2 equation of state (e.g., Span and Wagner, 1996).

Despite clear potential, a Web of Science search demonstrates that FI barometry has been underutilized by igneous petrologists in the last 10 years relative to other petrological barometers (Fig. 24c and d). Conventionally, the densities of FIs have been determined by observing phase changes during heating and cooling in a temperature-controlled microscope stage (microthermometry, Sorby, 1858). The relatively small number of existing studies in volcanic settings that use microthermometric techniques to obtain FI densities show the enormous potential of this method (Fig. 24c, see Hansteen and Klugel, 2008 and refs within). For example, FI in quartz-rich xenoliths from volcanoes in the Aeolian arc yield densities corresponding to both deep (~16–20 km) and shallow (4–6 km) pressures within the crust (Frezzotti et al., 2003). The deeper regions likely represent the primary magma storage zones, with shallower depths recording re-equilibration of FI during temporary residence at shallower levels during ascent toward the surface. The ability of FI to identify multi-stage ascent has also been demonstrated at Mt. Etna (Frezzotti et al., 1991), Cabo Verde (Klügel et al., 2020), Azores (Zanon and Frezzotti, 2013) and the Canary Islands (Hansteen et al., 1998).

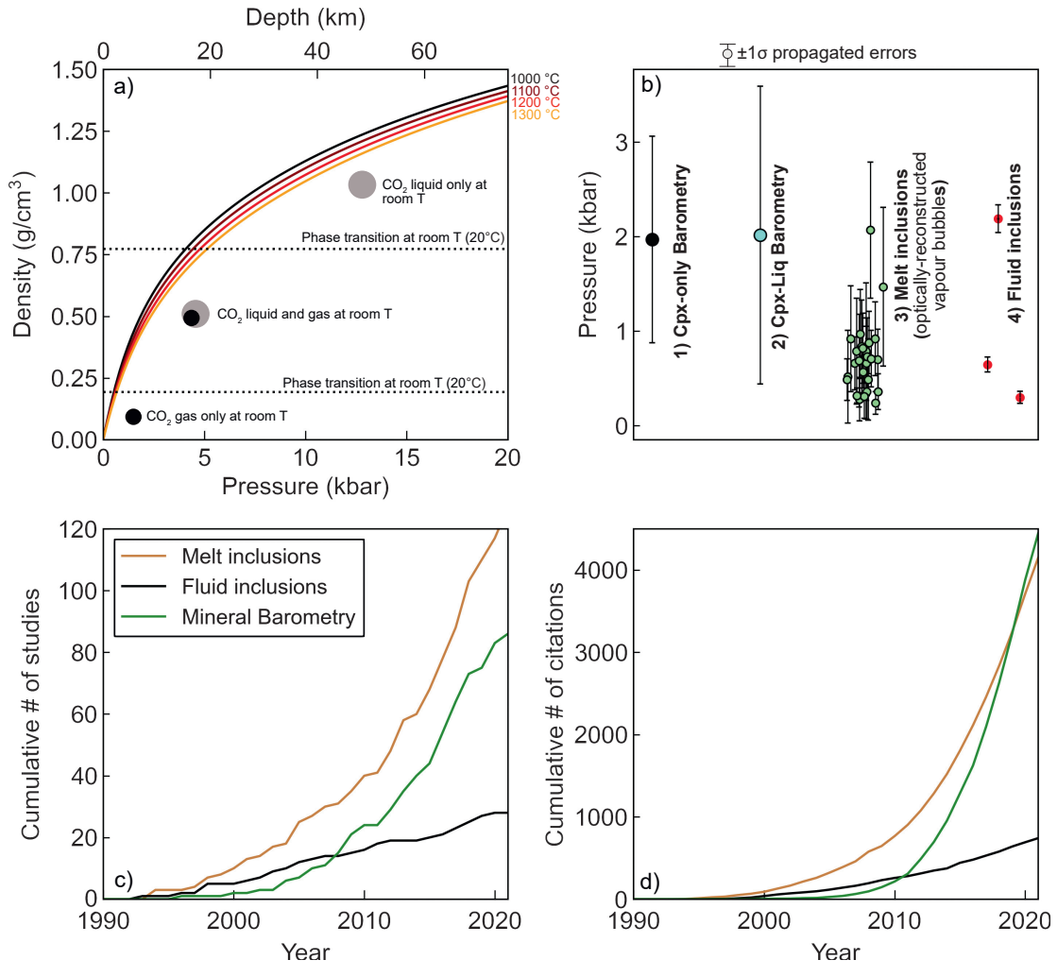

Fig. 24 (a) Relationship between pressure and the density of a pure CO_2 fluid (calculated using the Span and Wagner, 1996 EOS implemented in CoolProp, Bell et al., 2014). Storage depths shown for crustal densities of $\rho = 2700$ kg/m^3. (b) Monte Carlo simulations of 1σ errors for different barometry methods resulting from analytical uncertainty. Uncertainties calculated using Putirka (2008) eq32d-32b for Cpx-only and eq33–30 for Cpx-Liq arising from the EPMA analytical uncertainties of Neave et al. (2019, see Wieser et al., 2023d). Melt inclusion uncertainties show propagated errors in melt inclusions with vapor bubbles from the 2018 eruption of Kīlauea (Wieser et al., 2021). The large relative errors (±50%, ~1–1.5 kbar) arise from reconstructions of the volume of the vapor bubble using optical microscopy (with $z = (x + y)/2$, after Tucker et al., 2019). 1σ errors on calculated pressures for 3 FIs trapped at different crustal pressures, assuming an analytical error of 0.02 g/cm^3 from Raman spectroscopy, and uncertainty in T of 1$\sigma = 30$ K. (c) Web of Science search for number of articles using each method for barometry (full search criteria given in supporting information), and (d) number of citations to these articles.

It has long been established that Raman spectroscopy can be used as an alternative to Microthermometry to determine the density of CO_2-rich fluid inclusions (e.g., Rosso and Bodnar, 1995). This has been aided by recent improvements in the spatial and spectral resolution of confocal Raman spectroscopy, along with the development of precise ways to calibrate the relationship between peak parameters and CO_2 density for different Raman instruments (e.g., DeVitre et al., 2021; Kawakami et al., 2003; Lamadrid et al., 2017, see Section "Raman measurements of vapor bubbles"). Raman analyses have many advantages over Microthermometry:

1. Confocal Raman spectrometers are becoming increasingly common in universities, where they are widely used in other subdisciplines within the Earth Sciences (e.g., paleontology, mineral physics). In contrast, the use of heating-cooling stages and the associated expertise to conduct microthermometry is more restricted.
2. Raman analyses are faster (2–4 min) compared with the 10s of minutes required to conduct a single heating and cooling experiment. This is particularly true if there is only one fluid inclusion within the field of view at high magnification. Microthermometry can be faster on fluid inclusion trails as phase changes in multiple inclusions can be observed in one experiment.
3. Raman analyses only require only one surface to be ground down to within ~50–100 μm of the fluid inclusion, with a good enough polish to be able to visualize the FI on the Raman microscope. Normally, all that is required to achieve this is relatively

fine wet-and-dry paper (2000–7000 grade) and a quick 30 s polish on 1 μm aluminum polishing paper, while microthermometry requires a crystal to be prepared into a double polished wafer, with a good enough polish to see phase changes in detail.

4. Raman spectroscopy can be used on CO_2 fluid inclusions with a very wide range of densities (e.g., 0.01–1.24 g/cm^3) densities. In contrast, it is very hard to observe the homogenization temperature of a fluid inclusion with a density less than the critical density of CO_2 (<~0.47 g/cm^3) or a density higher than the triple point (>1.18 g/cm^3) using microthermometry (Kobayashi et al., 2012).
5. Raman spectroscopy can be applied on FIs with diameters down to ~1 μm (Frezzotti et al., 2012; Dayton et al., 2023), while it is very difficult to observe phase changes during microthermometry for inclusions <3–10 μm (Kobayashi et al., 2012). This is of particular importance in volcanic systems, where many fluid inclusions are very small (and smaller inclusions are more resistant to decrepitation during ascent, Bodnar et al., 1989; Campione et al., 2015; Wanamaker et al., 1990).

Relative to melt inclusion analyses and mineral barometry, estimates of magma storage depths from FIs are significantly more precise than both melt inclusions and mineral barometry, and have fewer sources of systematic uncertainty relating to the conversion of the measured quantity (density) to pressure (Dayton et al., 2023). Unlike the complex relationships relating pressure to mineral components (e.g., Jd in Cpx) or dissolved H_2O and CO_2 concentrations in silicate melts, the CO_2 equation of state is extremely well constrained, with very little offset between different parameterizations (~3% at 7 kbar, Span and Wagner, 1996; Sterner and Pitzer, 1994; Wieser and DeVitre, 2023). Additionally, unlike many mineral-melt barometers which are very temperature sensitive, the P calculated from the EOS is not all that sensitive to T (Fig. 24a). To demonstrate the high precision of this method, we propagate a ±30 K uncertainty in trapping temperature and a Raman analytical error of ±0.02 g/cm^3 using Monte-Carlo methods (Wieser and DeVitre, 2023), which yields a 12% 1σ error at 0.6 kbar, a 7% error at 2.2 kbar, and a 5% error at 5.2 kbar. These errors are significantly smaller than other petrological barometers (Fig. 24b). A final advantage over melt inclusion barometry is that far fewer analytical steps are required to obtain FI pressures (no μCT, FTIR/SIMS, EPMA measurements of glass), increasing the speed at which depths can be determined, at a far lower cost (Dayton et al., 2023). This speed is aided by open-source Python3 packages which help to automate data processing and CO_2 equation of state calculations, which were previously one of the slower aspects of the Raman fluid inclusion method (Wieser and DeVitre, 2023).

Several sources of systematic error must still be considered when interpreting fluid inclusion pressures. First, like melt inclusions, fluid inclusions may experience decrepitation upon ascent, which can yield anomalously low pressures, sometimes without clear textural evidence for CO_2-loss because of rehealing during ascent. Second, the pressure the inclusion exerts on the host can also cause deformation of the host through the movement of dislocations, increasing the inclusion volume. These two processes are captured by two of the three "Roedders" rules of fluid inclusions: "*2) the inclusion represents a constant volume, 3) nothing has been added or lost from the inclusion*" (Bodnar, 2017; Roedder, 1984). Namely, fluid inclusion barometry assumes that a fixed mass of CO_2 was trapped in the crystal, with a fixed volume, meaning the inclusion has a fixed density, so the density measured in the laboratory is the same as the density at which the inclusion was trapped. Decrepitation violates rule 3 (and possibly 2 as well), while re-equilibration violates rule 2. While numerous work has been conducted to determine re-equilibration processes for quartz-hosted fluid inclusions (Bakker, 2017; Boullier et al., 1989; Pecher, 1981; Qin et al., 1992), to our knowledge, the only experimental constraint on olivine re-equilibration was performed by Wanamaker and Evans (1989). They re-equilibrate fluid inclusions within San Carlos olivine by holding crystals at 1400 °C and atmospheric pressure for several days. We suggest that further experimental work is required to confidently predict how fluid inclusions in common mafic phases (olivine, pyroxenes, etc.) re-equilibrate when subject to a specific P-T-t path, with the aim of having models which can be run for a given set of samples similar to those used to model H^+ loss from melt inclusions (e.g., Barth and Plank, 2021). However, it is clear from analyses of fluid inclusions in mantle xenoliths that return crustal pressures that re-equilibration should be taken as the norm rather than the exception, and fluid inclusions will normally be reset to the pressures corresponding to the final region of prolonged magma stalling prior to eruption (Hansteen and Klugel, 2008).

The presence of fluid species other than CO_2 can also influence the accuracy of calculated pressures by both Raman spectroscopy and microthermometry. In igneous systems, the most common species present in the exsolved fluid phase are H_2O, SO_2, Cl, F, H_2S, N_2, CH_4, CO, and He. While some species are strongly Raman active with peaks that are close enough to the CO_2 diad to be visible in most high resolution acquisitions (e.g., SO_2 produces a sharp peak at 1151 cm^{-1}, Frezzotti et al., 2012), others will require a separate acquisition centered at higher wavenumbers (e.g., the N_2 peak is at 2331 cm^{-1}, CH_4 at 2917, Frezzotti et al., 2012). Monoatomic gases (e.g., He) are Raman inactive, meaning they don't produce any peaks. Molar proportions can be estimated from peak area ratios and knowledge of scattering cross sections for different Raman-active gases (Burke, 2001). Some fluid species can also be identified (and sometimes quantified) from changes in the freezing and melting temperatures by microthermometry (Hansteen and Klugel, 2008; Van Den Kerkhof, 1990), although in many cases, the relevant phase diagram has not be constrained (e.g., C O_2-He).

Mixed H_2O-CO_2 fluids present a particularly challenging problem to fluid inclusion barometry, because H_2O is so ubiquitous in the exsolved vapor phase in igneous systems. In arc magmas H_2O will be present in non negligible molar proportions at all crustal levels (≫10 mol%, Wieser et al., 2023a). Even in CO_2 dominated systems (e.g., Hawai'i), exsolved fluids will have substantial proportions of H_2O in the vapor phase as magmas ascend toward the surface. H_2O can be identified by microthermometry through the appearance of clathrate phases, and by Raman spectroscopy through identification of the O-H stretching band at ~3600 cm^{-1} (Azbej et al., 2007). However, at room temperatures, H_2O will be present as a thin film of liquid along the edge of the inclusion. This can make it very hard to identify by Raman spectroscopy, unless the inclusion is heated to the point at which H_2O dissolves

into CO_2 (e.g., to ~150 °C, Esposito et al., 2016). It has also been demonstrated that this H_2O film can react with the host crystal or a thin film of melt around the fluid inclusion (Andersen et al., 1984; Esposito et al., 2016; Frezzotti et al., 2002), and/or be lost through diffusive re-equilibration with a degassed carrier melt (Mackwell and Kohlstedt, 1990). This means it can be hard to identify which FIs had non negligible quantities of H_2O at the time of entrapment.

The presence of other fluid species complicates calculations of pressure in two ways. First, CO_2 densities obtained by Raman spectroscopy or microthermometry are normally converted into an entrapment pressure using a pure CO_2 EOS, which will deviate from the true relationship defined by a mix of the species of interest. Secondly, the reaction of secondary phases to form precipitates on walls, solid crystals, or liquid films means that the measured density represents only that of the residual CO_2 fluid phase, not the initial trapped fluid. Hansteen and Klugel (2008) discuss possible corrections for mixed fluids in their review (e.g., estimating H_2O/CO_2 ratios, then using a mixed H_2O-CO_2 EOS). Looking forward, we suggest that further work into mixed fluid equation of states and phase diagrams, as well as experiments of fluid inclusion re-equilibration, is required to continue pushing fluid inclusion barometry forward.

Can hygrometers be used as barometers?

At the opposite end of the spectra to determining pressures from fluid inclusions in CO_2-rich systems, many studies have used melt H_2O contents to estimate pressures using a pure H_2O solubility model, either through direct measurements in melt inclusions (e.g., Blundy and Cashman, 2005; Rutherford et al., 1985) or calculated H_2O contents from mineral and mineral-melt hygrometers (e.g., Amphibole Mg#s—Goltz et al., 2020; Krawczynski et al., 2012, Plag-Liq hygrometry—Ruth and Costa, 2021). For example, Ruth and Costa (2021) conduct thermobarometry on a variety of phases from Mayon Volcano (Fig. 25a). Cpx-Liq, Opx-Liq and Cpx-Opx pressures show a prominent peak at ~4 kbar (16 km). In contrast, they invoke a shallower magma reservoir at ~5 km based on H_2O-only saturation pressures calculated from plagioclase-liquid hygrometry (and H^+ in pyroxene). However, it is well established that primary arc magmas have non-negligible quantities of CO_2 (Shinohara, 2013; Wallace, 2005). While such H_2O-only calculations are often quoted as minimum estimates, it is worth questioning whether such minimum estimates are even useful, given that uncertainty in the amount of CO_2 can change the calculated pressure by up to an order of magnitude (Black and Andrews, 2020; Wieser et al., 2022b).

Using a global compilation of CO_2 in the glass phase of melt inclusions (Rasmussen et al., 2020, white histogram, Fig. 25b) and a compilation from the Cascade arc (red histogram, Fig. 25b, Wieser et al., 2023a), it is apparent that arc magmas have, at the very minimum, several hundred to a thousand ppm of CO_2. These quantities of CO_2 have a very significant effect on the calculated saturation pressures. For example, the H_2O-only saturation pressure of a typical arc basalt with 4 wt% H_2O (Plank et al., 2013) is ~1.3 kbar (white star on gray line), while the saturation pressure accounting for 1000 ppm CO_2 is 2.8 kbar (red star on gray line, Fig. 25b). In reality, it is becoming apparent that glass-only melt inclusion measurements have substantially underestimated the CO_2 content of arc magmas. A compilation of CO_2 contents in arc magmas for studies accounting the vapor bubble extends up to ~5000 ppm (Fig. 25c). For a magma with 4 wt% H_2O, a H_2O-only saturation pressure underestimates the true storage depth of a magma with 5000 ppm CO_2 by a factor of 5.7× (Fig. 25c, white star vs. green star). While H_2O depths are indeed minimum estimates, this example shows just how misleading they can be. Even for a magma with 10 wt% H_2O, H_2O-only pressures

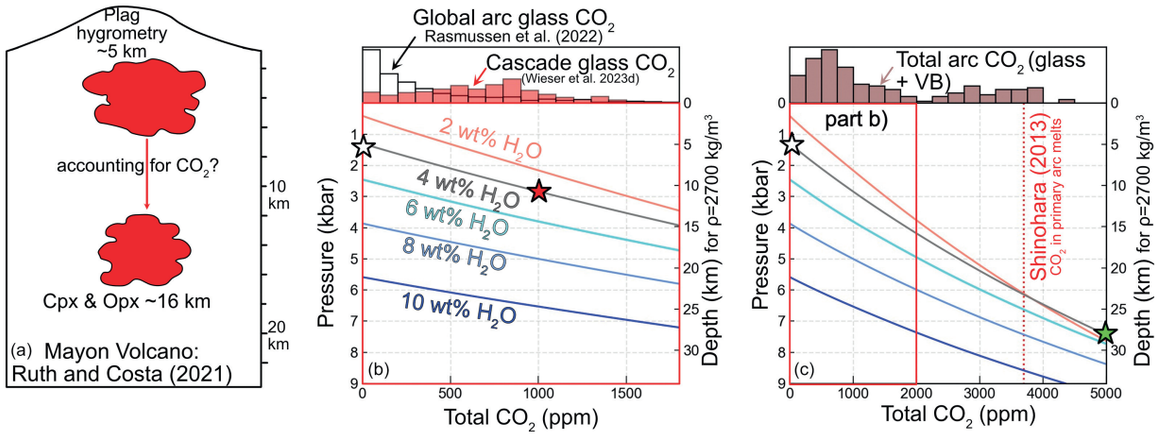

Fig. 25 Sensitivity of saturation pressures to CO_2 contents. (a) Schematic model of plumbing system at Mahon Volcano, adapted from Ruth and Costa (2021), where the depth of the shallower reservoir was calculated using a H_2O-only solubility model. (b and c) Increase in saturation pressure with CO_2 for five different H_2O contents (for a typical mafic Cascade melt composition, SiO_2 = 54.3 wt%), calculated using the solubility MagmaSat (Ghiorso and Gualda, 2015) in VESIcal (Iacovino et al., 2021). The histograms in (b) show a global compilation of arc CO_2 contents just measuring the glass phase (Rasmussen et al., 2022), and a compilation from the Cascade Arc. (c) Expanded x scale (region shown in b indicated in the red box) up to 5000 ppm CO_2. Histogram shows a compilation of total arc CO_2 accounting for the vapor bubble through Raman spectroscopy or homogenization (from Mironov et al., 2015; Moore et al., 2018; Rasmussen et al., 2020). We overlay the estimate of primitive arc magma CO_2 contents from Shinohara (2013).

underestimate by factors of ~1.8× if there is 5000 ppm CO_2. Thus, we speculate in many cases that accounting for CO_2 will push hygrometry estimated magma storage pressure substantially deeper, perhaps more in line with other barometry estimates (Fig. 25a).

Therefore, we suggest that H_2O-only saturation pressures should only be used in systems where it has been demonstrated that very little CO_2 is present, in the same way that fluid inclusion barometry using the pure CO_2 EOS can only be used in systems where there is very little H_2O in the fluid phase. In more CO_2-rich systems, saturation pressure calculations should be performed for a wide range of possible CO_2 contents to determine the uncertainty associated with H_2O-only saturation pressures.

Elastic thermobarometry/Thermoba-Raman-try

While the thermobarometric methods discussed thus far rely on changes in the chemistry of minerals, melts and fluids as a function of P and T, elastic thermobarometry relies on relative changes in the physical properties of inclusions and their hosts with changing P and T (Cisneros and Befus, 2020; Kohn et al., 2023). At the time of mineral growth and entrapment, a host-inclusion pair have the same pressure. However, after entrapment, the different thermal expansivity (change in volume with T) and compressibility (change in volume with P) of the two mineral species means that a residual pressure develops as the pair cool and/or ascend to the surface. This pressure can be deduced from the band/peak positions of minerals measured by Raman Spectroscopy, leading Kohn (2014) to term this method Thermoba-*Raman*-try! The magnitude of the residual pressure depends on the P-T path taken (see fig. 1 from Kohn, 2014). The final stress measured in the inclusion is a function of the initial P-T conditions of trapping, and the influence of differential contraction of the host-mineral pair, which can be modeled using knowledge of expansivity and incompressibility of different mineral species (Kohn, 2014). The most effective barometers are those where there is a large difference between the isothermal compressibility of the inclusion and host (i.e., changes in volume during pressure changes), while the most effective thermometers are those with the largest variations in isobaric expansivity (i.e., changes in volume during temperature changes, Kohn, 2014; Kohn et al., 2023).

Useful host-inclusion pairs are those with large differences in physical properties. There are a number of barometers using quartz inclusions, because these have a very different compressibility to host minerals such as zircon, lawsonite, and garnet (Kohn et al., 2023). Diamond is another popular host (e.g., Smith et al., 2022), along with garnet and zircon (e.g., olivine inclusions in diamond, coesite-in-diamond, and coesite-in-zircon, Kohn, 2014; Cisneros and Befus, 2020). Thermoba-Raman-try techniques can be remarkably precise; the estimated error in calculated pressure for Qtz-in-Gt inclusions is 0.3–0.5 kbar, resulting from uncertainty in the Raman band position of 0.5 cm^{-1} (Kohn, 2014). However, in general this technique has been applied to higher pressure systems where these phases are common (e.g., mantle rocks, metamorphic systems, xenoliths). Less attention has been paid to typical phases present in samples used to address volcanological research questions.

Befus et al. (2018) state that they provide the first elastic thermobarometer in igneous systems, by performing diamond anvil experiments on feldspar from atmospheric pressure to 3.6 GPa, and characterizing the Raman bands. They find a clear shift of peak positions to higher pressure from 0 to ~30 kbar, with Raman shifts of the band at ~485 cm^{-1} of ~0.42 cm^{-1} per kbar for Albite, 0.45 cm^{-1} per kbar for Andesine, and 0.31 cm^{-1} per kbar for Anorthite. Their calibrations relating peak position with pressure could be applied to natural feldspars inclusions (e.g., Fspar in Cpx, Fspar in Ol, etc.) with known compositions, using models of host-inclusion relaxation along a PT path. They suggest that the main limitation of applying this technique to shallower crustal systems where peak shifts are rather small is the limited spectral resolution of many Raman spectrometers (~1 cm^{-1}). However, the most recent generation of Raman spectrometers have improved spectral resolution (0.1–0.4 cm^{-1}). Additionally, Befus et al. (2018) fit a cubic spline to the Raman data. If the peak shape can be determined (i.e., a Lorentzian or Gaussian), peak fitting can achieve significantly higher precision than the spectra resolution (Yuan and Mayanovic, 2017). This increase in spectral resolution accompanying peak fitting is vital for precise quantification of CO_2 densities using Raman spectroscopy. For example, a WITEC alpha300R Raman has a spectra resolution of 0.57 cm^{-1} using the 1800 grating. However, the peak fitting error on the CO_2 peaks using a pseudovoigt ranges from 0.002–0.05 cm^{-1} (error quantified in lmfit implemented in DiadFit, Newville et al., 2016; Wieser and DeVitre, 2023).

To further broader the scope of the elastic thermobarometry method, Kohn (2014) evaluate 48 inclusion-host pairs and Cisneros and Befus (2020) consider >5000 pairs, highlighting a number of inclusion-host pairs that may provide useful P-T constraints in volcanic systems (e.g., magnetite in olivine, feldspar in pyroxene, feldspar in olivine). Despite the potential of this method, in the 5 years that have passed since Befus et al. (2018) and 3 years since Cisneros and Befus (2020), none of the citing studies are focused on volcanic rocks. Kohn et al. (2023) and Cisneros and Befus (2020) identify numerous roads for improvement of the elastic thermobarometry method, including the need for measurements of elastic properties at the P-T conditions of interest, calibrations of Raman shift as a function of pressure, models for inelastic processes, nonideal geometries, and a better understanding of the effects of anisotropy on ideal P. Hopefully, the recent proliferation of interest and expertise in Raman spectroscopy by igneous petrologists measuring vapor bubbles should aid such efforts.

Experimental petrology

Experimental petrology fundamentally underpins the melt inclusion, mineral-melt thermobarometry, and elastic thermobarometry methods discussed so far, because the composition of experimental products are used to calibrate models. However, experimental petrology can also be used directly to investigate magma storage conditions in a specific system. Generally, such experiments use a

starting composition characteristic of a specific volcanic system or eruption, and perform experiments at a range of P, T, H_2O and fO_2 conditions. Then, experimental phase compositions are compared to natural samples to determine the most probable storage conditions (e.g., comparing glass compositions, mineral core and rim conditions, occurrence of breakdown reactions; First et al., 2021; Rutherford et al., 1985; Weber and Castro, 2017).

For example, Voigt et al. (2022) perform experiments at 0.25–5 kbar and 850–1100 °C on natural starting materials from the 1257 Samalas eruption. They observe plagioclase and amphibole in their samples, so the PT space between amphibole breakdown and plagioclase instability places constraints on pre-eruptive PT conditions. Generally, once a stability region is identified, additional constraints are required to narrow down the magma storage conditions (Bohrson and Clague, 1988). For example, Cadoux et al. (2014) overlay Fe-Ti oxide temperatures on their phase diagrams, concluding that the observed phase assemblage at Santorini is recreated experimentally at 2 and 4 kbar.

When performing experiments to deduce magma storage conditions at a given volcano, there can be a very large solution space to explore, in terms of pressure, temperature, and fO_2. If a natural Cpx has compositional similarities to one grown experimentally at a specific set of conditions, it is difficult to quantify the true uncertainty on storage conditions without exploring a very large number of experimental conditions to determine whether a very similar composition may appear at a different set of P-T-fo_2 conditions, particularly from a slightly different bulk composition). It is also vital to consider the composition of the fluid phase, which is often described in terms of the partial pressure of H_2O (P_{H2O}) or the mole fraction of H_2O in the starting composition or exsolved fluid phase (X_{H2O}). In order to reduce the size of one variable of the solution space, many experiments are performed at water-saturated conditions (i.e., $X_{H2O} = 1$, Blatter and Carmichael, 2001; First et al., 2021; Grove et al., 1997; Nakatani et al., 2022; Sisson and Grove, 1993; Voigt et al., 2022). However, many natural systems do not contain enough H_2O to be volatile saturated at high pressures, or contain relatively large amounts of CO_2, meaning the system is volatile saturated in a fluid with $X_{H2O} < 1$. Experiments with H_2O below the quantity required for H_2O saturation have been performed (e.g., Kawamoto, 1996), as have experiments in equilibrium with mixed CO_2-H_2O fluids (e.g., Alonso-Perez et al., 2009; Cadoux et al., 2014), although these are less common than pure H_2O-saturated experiments.

Interestingly, there are many reports in the literature of phase stability being affected by X_{H2O}. For example, Keppler (1989) investigate solidus temperatures in the haplogranite system, and find that the solidus position varies as a function of the fluid phase composition for a mix of H_2O and CO_2. However, if X_{H2O} is reduced by the addition of N_2 instead, this relationship is not seen. These experiments indicate that at X_{H2O} close to 1, CO_2 is not just acting as an inert component to reduce the activity of H_2O, but is modifying the structure of the melt. Rutherford et al. (1985) also note that experiments where X_{H2O} is reduced by the presence of H_2 are not directly comparable to those where it is reduced by CO_2 (e.g., the addition of CO_2 is not following simple Henry's law behavior).

It has been shown numerous times that CO_2 affects the stability of amphibole. Given the increasing amounts of CO_2 being measured in melt inclusions in mafic arc magmas (Fig. 19), it is highly likely that the majority of high Mg# amphiboles examined at the surface formed in a system with $X_{H2O} \ll 1$. For example, Ridolfi et al. (2010) suggested that large amounts of CO_2 in a relatively high T igneous system with relatively low amounts of H_2O stabilizes OH-bearing phases such as amphibole, perhaps because of increasing solubility of OH- in the melt with increasing CO_2 dissolution.King and Holloway, 2002 suggest a reaction where molecular CO_2 and H_2O react to form carbonate and OH^-. Ridolfi et al. (2010) note that the paucity of magnesiohastingsite amphiboles in predominantly H_2O-saturated experiments may indicate that these compositions only form at high T and high CO_2 contents. Krawczynski et al. (2012) also find amphibole is stabilized at a higher temperature, with a lower Mg# in experiments with more CO_2.

In their experiments on the Mt. St. Helens dacite, Rutherford et al. (1985) find radical changes in the crystallization temperature of different phases as CO_2 is added to reduce P_{H2O} at a constant $P_{total.}$ Plag, Px and Fe-Ti oxides crystallization temperatures increase, while the amphibole liquidus temperatures decrease. Additional experiments by Rutherford and Devine (1988) confirm that the observed phase stability is produced experimentally at 920 °C, $P = 220$ Mpa, and $X_{H2O} = 0.67$. Rader and Larsen (2013) perform experiments on low MgO basaltic-andesites to constrain the impact of small amounts of CO_2, running experiments at a range of T (900–1200 °C) and P (0.001–1.8 kbar) at both $X_{H2O} = 1$ and $X_{H2O} = \sim 0.7$. They found that the plagioclase stability curve was the most sensitive to X_{H2O}, shifting 25 °C for $X_{H2O} = \sim 0.7$. They also observe shifts in amphibole stability. Finally, Cadoux et al. (2014) found that at 850 °C and 4 kbar, there was more amphibole present when X_{H2O} was 0.9 than 1. They also find changes in the stability of both pyroxenes, ilmenite and plagioclase, as well as the liquidus phase with changes in X_{H2O}. In particular, they find that orthopyroxene is generally not stable where $X_{H2O} = 1$ at 2–4 kbar.

In summary, these experiments show that X_{H2O} is clearly a very important variable to investigate in further experiments, and comparison of natural samples to experimental products should factor in possible differences in X_{H2O}. We believe this avenue of investigation is particularly important given the last decade of Raman work has clearly demonstrated that mafic arc magmas, have substantially more CO_2 (and thus a lower X_{H2O} ratio) than previously thought.

Thermobarometers based on thermodynamic modeling

Thermodynamic modeling is a powerful tool in igneous systems for exploring various hypotheses, such as whether the chemical variations in a suite of lavas can be produced by equilibrium or fractional crystallization alone, or whether processes such as crustal melting/assimilation are required (Heywood et al., 2020). However, in addition to general hypothesis testing, thermodynamical models have also been used to place quantitative constraints on magma storage conditions. Two main methods are discussed below, matching liquid lines of descent to erupted lava compositions, and multiphase saturation methods.

Liquid lines of descent

One method to determine magma storage conditions compares fractional crystallization models conducted at different conditions (e.g., P, fO_2, melt H_2O content) to observed liquid lines of descent (LLD, e.g., whole-rock XRF, glass EPMA data). Magma storage conditions in the system of interest are inferred from the model conditions providing the best fit to natural data. Comparison of observed and modeled LLDs have been used to deduce storage conditions in explosive silicic eruptions (e.g., the Campanian Ignimbrite, Campi Flegrei, Fowler et al., 2007, Bishop Tuff, Long Valley Caldera, Fowler and Spera, 2010), in trachytes and pantellerites of the East African Rift (Gleeson et al., 2017; Hutchison et al., 2018; Peccerillo, 2003; Ronga et al., 2010), in peralkaline rhyolites from Atlantic Ocean hotspots (Jeffery et al., 2017, 2016), and in basalts from Mauna Loa (Gaffney, 2002).

LLD methods are typically performed in MELTS, but other tools for modeling fractional crystallization such as COMAGMAT or Petrolog3 could also be used (Ariskin et al., 1993; Danyushevsky and Plechov, 2011). The best fit model is often determined by visual comparison between model outputs and natural samples (e.g., Fowler et al., 2007), although statistical methods have also been used. For example, Gleeson et al. (2017) develop an algorithm to calculate the smallest misfit between each datapoint and the modeled LLD path for each MELTS model. They combine the residuals of all individual samples into a weighted RMSE offset to select the best model.

In certain situations, storage conditions determined by LLD methods are supported by independent lines of evidence. For example, there is a close agreement between the depths of magma storage estimated by MELTS modeling and geophysical estimates of magma storage depths in the East African Rift (Hutchison et al., 2016; Gleeson et al., 2017), and between MELTS-estimated fO_2 and experiment constraints (Scaillet and Macdonald, 2001). However, as pointed out by Gleeson et al. (2017), even the best-fit models display systematic offsets between the MELTS predictions and the natural data for certain oxides (e.g., CaO, P_2O_5). These offsets indicate that there are several inaccuracies or omissions in the MELTS thermodynamic models (e.g., absence of amphibole or biotite in calc-alkaline magmas) that present severe limitations for the use of MELTS-based crystallization models as a thermobarometric technique.

In most cases, MELTS correctly identifies the directionality caused by changing P, H_2O, or fO_2 on the LLD (e.g., enhancing or suppressing the stability of a given mineral). However, it is unclear whether the predicted mineral stability for a given value of P, H_2O, or fO_2 is correct, or whether it is just relative differences that are trustworthy. While MELTS modeling may be able to distinguish a dry vs. wet LLD, it may not be able to determine whether a magma has 0.1 or 0.6 wt% H_2O. It has been shown that MELTS models at Kīlauea Volcano must be run with melt H_2O contents below those measured in melt inclusions to recreate the MgO content at which plagioclase and Fe-Ti oxide crystallize (Garcia, 2003; Wieser et al., 2022a). If H_2O contents were not independently constrained, such models could lead to incorrect inferences about the hydration state of magmas at this volcano.

To assess LLD methods further, we compare the measured composition of experimental liquids produced during fractional crystallization experiments on a H_2O-poor tholeiitic basalts to MELTS crystallization models (Villiger et al., 2007). We run a fractional crystallization model at the specified experimental pressure (7 kbar), fO_2 and H_2O content. We also run models at pressures of 1, 4 and 10 kbar (Fig. 26a, Supporting Figs. 11–13). We use three different MELTS versions (pMELTS, rhyolite-MELTS v1.0.2 and v1.2.0), Petrolog3 (Danyushevsky and Plechov, 2011), and MAGEMin (Riel et al., 2022) using the Holland et al. (2018) thermodynamic database. Both versions of rhyolite-MELTS show a very poor fit at the experimental pressure, predicting extensive Al_2O_3 enrichment and SiO_2 depletion relative to the experimental products (dark blue and green lines, Fig. 26e and f). The fit is far better at 1 kbar, which could lead to anomalous inferences of magma storage pressures using LLD methods. pMELTS (which has a different liquid model to rhyolite-MELTS) doesn't show such extreme Al_2O_3 enrichment at the experimental pressure, but the fit for Al_2O_3 is still better at 4 kbar than 7 kbar, and the fit for SiO_2 is best at 1–4 kbar (cyan lines). Petrolog3 shows anomalous SiO_2 depletion at all pressures (magenta lines, Fig. 26), and none of the 4 pressures used here provide a satisfactory fit to most elements (Supporting Fig. 11). The Holland et al. (2018) database, implemented through MAGEMin, certainly does the best job of recreating Al_2O_3 systematics at experimental pressures (Fig. 26e), although the fit to SiO_2 is far better at 4 kbar than 7 kbar (black lines, Fig. 26, Supporting Fig. 11).

To further assess LLD methods, we use all three MELTS versions to model the fractional crystallization experiments of Nandedkar et al. (2014) conducted on a hydrous arc basalt (Fig. 27). The hydrous nature of these experiments means that the differences between rhyolite-MELTS v1.2.0 and v1.0.2 are far more noticeable compared to the H_2O-poor experiments of Villiger et al. (2007). Using v1.2.0, models run at 4 to 10 kbar provide a reasonable match to observed oxide contents. In contrast, using v1.0.2, models run at 1 kbar provide the best fit to both SiO_2 and Al_2O_3. It is concerning that these models give such different results (even before the onset of amphibole saturation), given they are used somewhat interchangeably in the literature, and obviously, any studies conducted prior to the release of v1.2.0 in 2015 must have used v1.0.2. pMELTS shows behavior between the two rhyolite-MELTS models, with trends best recreated at 1–4 kbar.

These comparisons to two experimental studies demonstrate that the pressure you would deduce from LLD methods is very sensitive to both the choice of model and the oxides used for assessment of the "best model." In neither comparison does the experimental pressure stand out as the best model fit. It is also worth noting that for both these examples, we are using known fO_2 and H_2O contents. In reality, it is likely that P, fO_2, and H_2O are all uncertain, which leaves a very large solution space to explore (and can result in even greater ambiguity regarding the best fit model). Overall, we conclude that in their current state, LLD methods are not a reliable way to deduce storage pressure, particularly given the result can differ so much simply based on the choice of fractional crystallization model. It may be that an update of the MELTS liquid model, and/or tweaks to the thermodynamic data controlling the mineral stability fields (as performed for the qtz-2 fspar ternary) can revive this method. These comparisons also demonstrate that it is absolutely vital for papers to state the version of MELTS that they used.

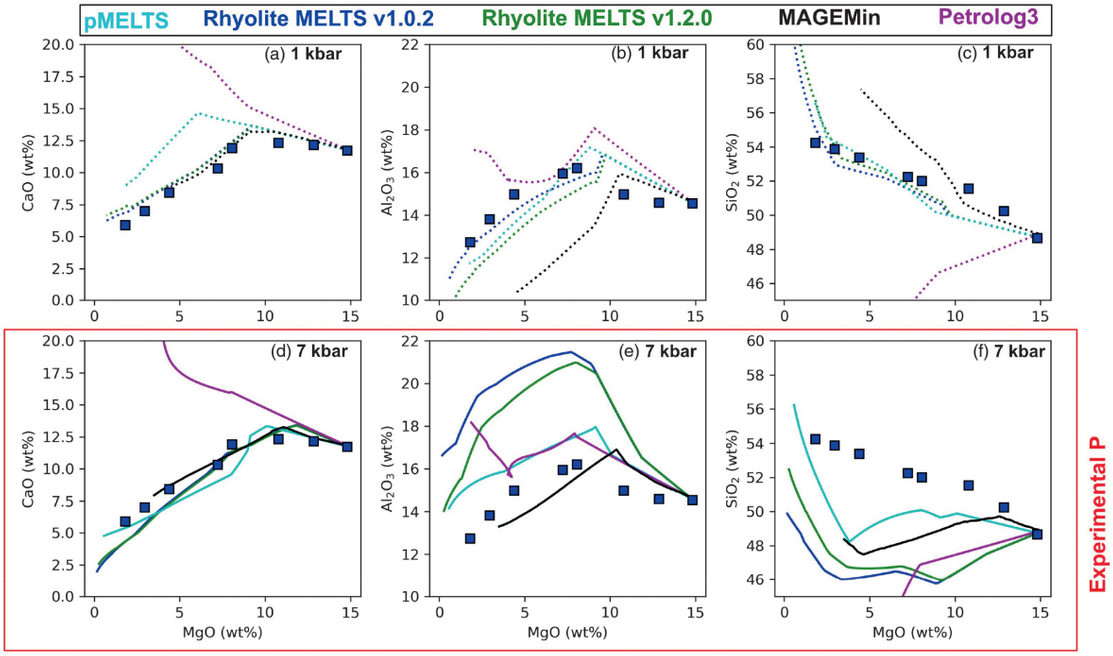

Fig. 26 Comparison of fractional crystallization models run in rhyolite-MELTS, pMELTS, Petrolog3 and MAGEMin (using the Holland et al., 2018 database) to fractional crystallization experiments from a nominally anhydrous, tholeiitic basalt from Villiger et al. (2007). MELTS calculations are performed at the experiment fO_2 (FMQ-2). Models use $H_2O = 0.2$ wt% (i.e., nominally anhydrous). While experiments were conducted at 7 kbar (solid lines), models run at 1–4 kbar provide a much better fit (particularly regarding the onset of plagioclase). Other pressures shown in Supporting Fig. 11. Calculations run using functions from PetThermoTools using underlying code from MAGEMin (Riel et al., 2022) and alphaMELTS for Python (Antoshechkina and Ghiorso, 2018).

Multi-phase saturation

As discussed in Section "Liquid-only barometry," pressure influences the location of mineral cotectics and eutectics. Under the assumption that the measured composition of a multi-phase saturated melt might contain information about the pressure of the system, rhyolite-MELTS has been used to assess the crystallization pressure of silicic and intermediate magmas that are co-saturated in quartz + feldspar (Bégué et al., 2014a,b; Gualda et al., 2019a,b; Gualda and Ghiorso, 2014; Pamukcu et al., 2015), feldspar + orthopyroxene ± quartz (Pamukçu et al., 2021) or plagioclase + clinopyroxene + orthopyroxene (Harmon et al., 2018).

Specifically, these methods work by performing crystallization calculations at several discrete pressures (and a specified H_2O and fO_2 value). T is progressively dropped from the liquidus, and the appearance of different phases is tracked. Mineral saturation curves (see Fig. 28a and b) are determined from these individual isobaric crystallization calculations. A residual is defined as the maximum temperature difference between the saturation curves of the phases of interest (ΔT, Fig. 28c). In many cases, the mineral saturation curves never intersect exactly at a single point, so the best fit pressure is calculated from the minimum point on the residual curve. A solution is only considered valid if the residual T gets within a pre-specified threshold value of a perfect intersection (e.g., 5 °C for quartz—plagioclase—alkali-feldspar equilibrium; Gualda and Ghiorso, 2014). This threshold varies—in their application of this method to Plag-Opx-Cpx equilibrium, Harmon et al. (2018) compare experimental pressures to calculated pressures with minimum residuals from 16 °C to 121 °C.

One complication of these methods is that the position of saturation curves is not just influenced by pressure, but also fO_2 and melt H_2O content. This means for natural samples; calculations must be repeated at different melt fO_2 and H_2O contents to identify the location of the minimum T offset in multivariate P-fO_2-H_2O space. In an ideal world, this would mean that MELTS could be used simultaneously as a barometer, oxybarometer, and hygrometer. However, in reality, this very large solution space can result in substantial errors which do not reproduce the true conditions for P, fO_2, or H_2O.

Quartz—2 feldspar MELTS barometry

Interest in using the co-saturation of quartz, plagioclase and alkali feldspar to help constrain pressure (Gualda and Ghiorso, 2014) led to a substantial update to the MELTS algorithm in 2012. To recreate the near invariant behavior of pumice and melt inclusion compositions from the early erupted Bishop Tuff, which have major and trace element systematics indicating a low degree of freedom, Gualda et al. (2012) tweaked the enthalpy of formation of quartz and the potassic endmember in the alkali-feldspar solid-solution (Gualda et al., 2012, rhyolite-MELTS v1.0.2). Using this new model, and the residual T method described above, Gualda and Ghiorso (2014) estimated storage pressures for a variety of quartz-saturated rhyolites. They compared these MELTS

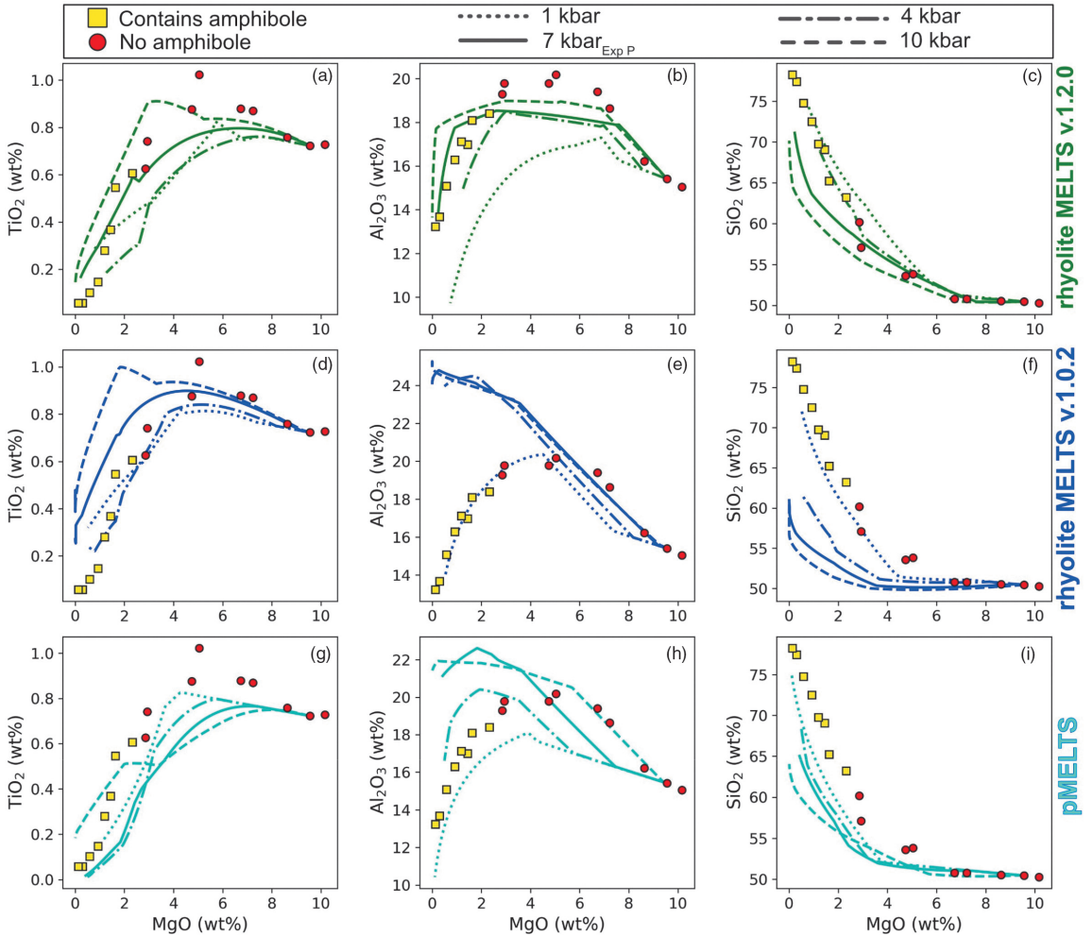

Fig. 27 Comparison of fractional crystallization experiments from Nandedkar et al. (2014) at 7 kbar to MELTS fractional crystallization models run using three different versions. Experiments containing amphibole are shown in yellow, those without amphibole in red. Models were run at the experimental fO_2 and H_2O contents, at 1, 4, 7 (Exp P) and 10 kbar. MELTS models run in PetThermoTools using alphaMELTS for Python (Antoshechkina and Ghiorso, 2018; Gleeson and Wieser, 2023).

pressures to H_2O-CO_2 saturation pressures from the Bishop Tuff (Anderson et al., 2000), the Younger Toba Tuff (Chesner and Luhr, 2010) and the Mamaku Ignimbrite (Bégué et al., 2014a,b), and found a close agreement. Based on the results in Fig. 10 of Gualda and Ghiorso (2014), we calculate an R^2 of 0.75 and RMSE of 0.38 kbar for $N = 30$ melt inclusions when performing a linear regression between the two pressure methods. However, it is worth noting $N = 17$ of these inclusions are from the Bishop Tuff, which was used as a reference point for the tweaking of the rhyolite-MELTS model calibration (Gualda et al., 2012).

The rhyolite-MELTS quartz-feldspar barometers have since been applied to several other systems worldwide. This includes the application of quartz-plagioclase barometry to rhyolitic eruptions of the Taupo Volcano Zone, New Zealand (Bégué et al., 2014a,b), and quartz—2 feldspar barometry on the matrix glass of the Peach Spring Tuff, USA (Pamukcu et al., 2015). Results indicate spatial and/or temporal variations in the magma storage conditions beneath the Taupo Volcanic Zone, as well as correlations between rhyolite SiO_2 contents and pressure (Bégué et al., 2014a,b; Pamukçu et al., 2020). However, these papers have been the subject to a number of comment-reply articles (e.g., Wilson et al., 2021; Pamukçu et al., 2021). Most notably, Wilke et al. (2017) question the accuracy of the rhyolite-MELTS barometer following comparison of the results of the rhyolite-MELTS geobarometer with their new empirical expression for the pressure of quartz + feldspar saturated liquids. The Wilke et al. (2017) method, termed DERP (Determining Rhyolitic Pressures), is based on experiments on haplogranitic compositions which examine the influence of melt H_2O content (or activity) and the normative melt An content (largely determined by the CaO content of the liquid) on the quartz + feldspar saturation surface. Comparison of rhyolite-MELTS and DERP barometric results for melt-inclusions and matrix glasses from the Taupo Volcanic Zone and Peach Spring Tuff revealed a large offset between the two methods, with DERP typically returning pressures around twice as large as those from rhyolite-MELTS. The pressure discrepancy between the two methods correlates with the normative melt An content (a key part of the new DERP parameterization). Thus, Wilke et al. (2017) suggest that rhyolite-MELTS may underestimate magma storage pressures in Ca-bearing rhyolites.

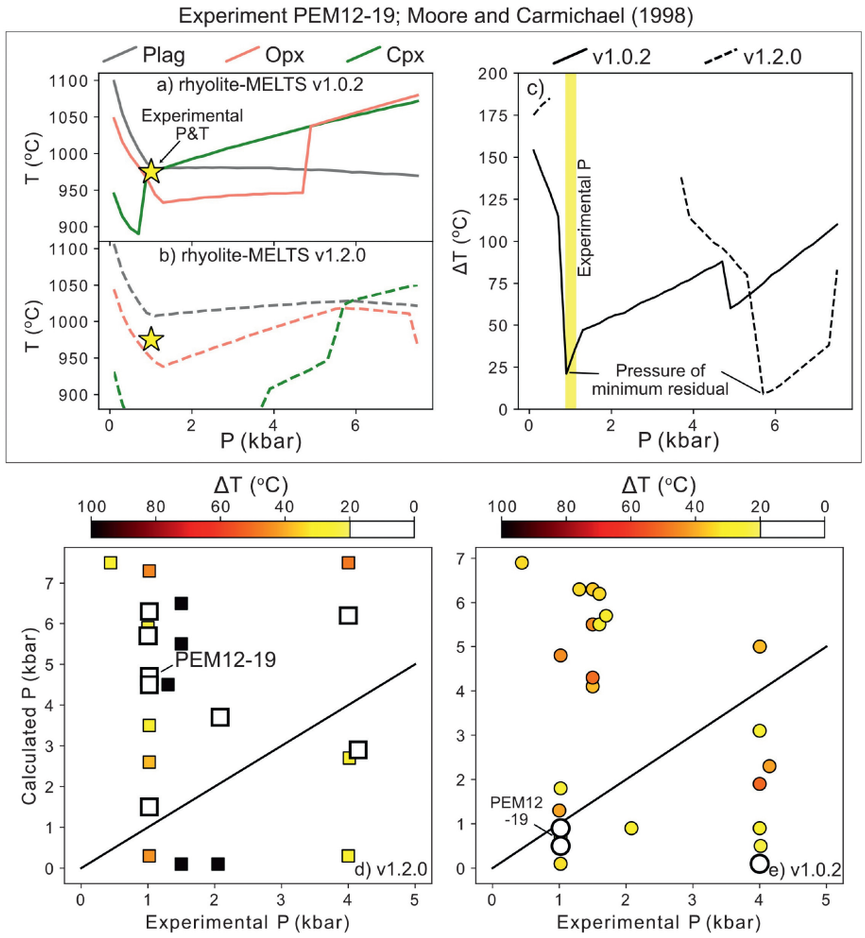

Fig. 28 Assessing the Rhyolite-MELTS Plagioclase 2–pyroxene barometer of Harmon et al. (2018). (a and b) Constructing mineral stability curves for Exp. PEM12–19 from Moore and Carmichael (1998) at the experimental fO_2 ($\triangle$NNO = +1.1) and H_2O contents (3.6 wt%) using two different versions of MELTS. (c) Calculating the maximum T offset at each pressure between the three mineral stability curves. The MELTS barometer takes the minimum residual as the best-fit T ($\sim$1 and $\sim$6 kbar depending on the MELTS version). (d and e) Testing the barometer using experiments saturated in Plag-2 pyroxene from ArcPL. Calculated pressure using the residual method plotted against the experimental pressure (calculations performed at experiment fO_2 and H_2O contents). Experiments with residuals $<$20 °C are colored white, and larger residuals are colored based on this value. Calculations performed in PetThermoTools (Gleeson and Wieser, 2023) using alphaMELTS for Python (Antoshechkina and Ghiorso, 2018).

The relative performance of DERP and rhyolite-MELTS was revisited by Gualda et al. (2019a), who argued that the use of multi-linear fits in DERP to determine the influence of P, An content and H_2O on the quartz—2 feldspar thermal minimum results in a functional form that is thermodynamically impossible. As a result, extrapolation of the DERP barometer outside the range of the calibration data could lead to systematic errors. Gualda et al. (2019a) also question the use of TitaniQ as an independent barometer to compare to the results from DERP in Wilke et al. (2017), as the Ti content in quartz is influenced by a wide number of parameters, including the mineral growth rate and melt Ti activity, which may limit its use as a geobarometer (see Section "Ti in Quartz (TitaniQ) thermometer"). To assess the performance of the 2 barometers in natural systems, Gualda et al. (2019a) compare the results derived from the DERP and rhyolite-MELTS barometers to independent pressure estimates from H_2O-CO_2 saturation in melt inclusions and amphibole thermobarometry. In general, the barometric estimates provided by rhyolite-MELTS are in good agreement with the independent pressure estimates. DERP-derived pressures are typically more scattered and often higher than those derived from melt inclusion or amphibole thermobarometry. However, given the issues with amphibole barometry discussed in Section "Amphibole thermobarometry and chemometry," such comparisons should perhaps be interpreted with caution.

Wilke et al. (2019) responded by stating that the large range of pressures estimated by the DERP barometer in Gualda et al. (2019a) result from its sensitivity to glass Na_2O and K_2O contents, which are used to calculate the normative Qz, Or, and Ab values. Wilke et al. (2019) suggest that in hydrous rhyolitic glasses where alkali migration during analysis is a serious issue, DERP should only be used to estimate storage pressures when data quality is assured. Furthermore, Wilke et al. (2019) note that the experimental compositions used to calibrate the DERP barometer all contain normative corundum, meaning that all CaO in the glass is used to calculate the normative An content. In many natural samples, such as the Peach Spring Tuff used by Pamukcu et al. (2015), lower Al_2O_3 contents

lead to the presence of Wollastonite in the normative mineral assemblage, complicating the determination of the normative melt An content. As normative Wollastonite is not accounted for in the DERP calibration, this could explain the offset between the two barometers and indicates that DERP should only be used in systems where Corundum appears in the normative mineral assemblage.

Therefore, for most natural samples where Wollastonite is present in the normative mineral assemblage, the rhyolite-MELTS geobarometer might be more reliable than empirically calibrated alternatives like DERP. Nevertheless, questions regarding the application of the rhyolite-MELTS barometer to natural systems still remain, especially as the test datasets used to evaluate this barometer have largely relied on data from the Bishop Tuff (part of the 2012 recalibration). In addition, inclusion of the updated H_2O-CO_2 model of Ghiorso and Gualda (2015) in the rhyolite-MELTS calculations (i.e., rhyolite-MELTS v1.2.0) shifts the quartz and sanidine stability field to such an extent that three-phase saturation cannot be simulated at any pressure for the compositions of Gualda and Ghiorso (2014) and Pamukcu et al. (2015). This indicates a further update to the quartz + sanidine ± plagioclase thermodynamic properties is necessary to align with the most recent volatile solubility model, so that users do not need to perform calculations for magmas that are typically saturated in a mixed H_2O-CO_2 fluid phase using an outdated H_2O solubility model. We are not of aware of any new experiments that can be used as an independent test dataset for resolving the rhyolite-MELTS vs. DERP debate, and to gain a better understanding of the errors involved. Further experiments in this compositional range will be vital to move this approach forward in a rigorous manner.

Plagioclase—2 pyroxene MELTS barometry

Since the inception of the rhyolite-MELTS geobarometer in 2014, the general method developed by Gualda and Ghiorso (2014) has been expanded to examine the cosaturation of other phases. Harmon et al. (2018) suggest that the rhyolite-MELTS thermodynamic models could be used to provide pressure estimates for intermediate magmas saturated in plagioclase, clinopyroxene and orthopyroxene.

To test the plagioclase—2 pyroxene geobarometer, Harmon et al. (2018) apply their method to $N = 8$ experiments from the LEPR database that are saturated in two or more of the phases of interest. As the plagioclase—2 pyroxene geobarometer is significantly more sensitive to fO_2 and H_2O content than the quartz—2 feldspar barometer, they use experimental fO_2 and melt H_2O contents for these comparisons to eliminate the multi-dimensional aspects of mineral stability, allowing the best fit pressure to be found. Of the eight experiments they use, only one returned a ΔT between the three mineral saturation curves of <20 °C, with the minimum T offset for other experiments stretching up to 121 °C (i.e., the saturation curves never intercept close to one another).

The single experiment with a minimum T offset <20 °C (1140mf #27 from Grove et al., 1997) returns a predicted P only 0.2 kbar offset from the true experimental P. If the threshold for the minimum T offset is relaxed, experiments with minimum T offsets between 20 °C and 40 °C also return P estimates within ~0.25 kbar of their experimental P. However, convergence of only 4 experiments conducted at 0.001–1 kbar where fO_2 and melt H_2O contents are known makes it difficult to robustly assess the accuracy of the geobarometer in natural systems when applied up to the 5 kbar limit suggested by Harmon et al. (2018).

We investigate this method further using experiments from variably hydrous magmas cosaturated in plagioclase, orthopyroxene and clinopyroxene present in the ArcPL dataset conducted within the pressure range suggested by Harmon et al. (2018, 0–5 kbar, $N = 39$). As for LLD methods, we find that the version of MELTS used for these calculations drastically affects calculations of storage conditions. Based on the flowchart provided on the MELTS-OFM website (https://melts.ofm-research.org/MELTS-decision-tree.html), rhyolite-MELTS v1.2.0 should be used for these calculations, as these experiments contain dissolved H_2O (±CO_2). However, Harmon et al. (2018) use rhyolite-MELTS v.1.0.2. For experiment PEM12–19 from Moore and Carmichael (1998), rhyolite-MELTS v.1.0.2 shows very good convergence at pressures very close to the experimental pressure (Fig. 28e). However, if rhyolite-MELTS v1.2.0 is used, the clinopyroxene stability curve intersects plagioclase and orthopyroxene at significantly higher pressures (Fig. 28a and b), leading to an estimated pressure of ~6 kbar (~5 kbar higher than the experimental pressure, Fig. 28c).

Using rhyolite-MELTS v.1.0.2 for all N = 39 experiments, $N = 10$ do not stabilize all three phases at any pressure within 200 °C of the liquidus, which is well below the experimental temperature in all cases. Of the remaining $N = 29$, models converging with a residual <20 °C are shown in white on Fig. 28e. These produce a very poor fit to experimental pressures. If we take solutions produced at larger temperature residuals (following Harmon et al., 2018), there is no improvement. Using rhyolite-MELTS v1.2.0, $N = 23$ experiments converge, and the fit is similarly bad, regardless of the temperature residual used. The poor performance of this barometer supports previous criticism of the ability of MELTS to correctly predict the stability of these phases, pointing to the need for better solution models (e.g., Klein et al., 2023; Nandedkar et al., 2014; Villiger, 2004).

Other methods

There are of course, numerous methods to determine magma storage conditions which we do not address in detail, both in the interest of relative(!) brevity, and because there is insufficient experimental data that wasn't used during calibration to independently test models.

Additional thermometers have been developed based on:

- REE exchange between Opx-Cpx (Liang et al., 2013) and Plag-Cpx (Sun and Liang, 2012).
- Mg-Fe exchange between Ol and Cpx (Loucks, 1996).
- Ni partitioning between Ol-Liq (resulting in a H_2O-independent thermometer, Pu et al., 2021, 2017).

- NaSi-CaAl exchange between Amp-Plag (e.g., Holland and Blundy, 1994; Molina et al., 2021a,b).
- Mg exchange between Plag-Cpx (Sun and Lissenberg, 2018).

Hygrometers based on:

- Cpx-Liq major element equilibrium (e.g., Armienti et al., 2013; Perinelli et al., 2016).
- Ca exchange between Ol-Liq (Gavrilenko et al., 2016).
- The highest amphibole Mg# found in a given volcanic system (Krawczynski et al., 2012).
- The H_2O contents of nominally-anhydrous minerals (NAMs) combined with models of mineral-melt partition coefficients (Demouchy et al., 2006; Towbin et al., 2023; Wade et al., 2008).
- Projections of the Ol, Cpx and Plag cotectics (Klein et al., 2023).

Barometers based on:

- Sr/Y contents of erupted lavas (Profeta et al., 2016).
- "Multi-reaction" thermodynamic models relying on underlying thermodynamic models (e.g., Holland and Powell, 2011), where predicted and observed mineral compositions and phase fractions are compared to constrain storage conditions (e.g., Ziberna et al., 2017; Nicoli and Matthews, 2019).
- Al-Si partitioning between Plag and Amp (Molina et al., 2015).
- CaO-Mg# systematics of mid-oceanic ridge magmas (Villiger et al., 2007).

Of course, these lists are by no means exhaustive.

Future developments should be guided by FAIR principles

It is becoming increasingly recognized that science needs to shift to a framework where research products are findable, accessible, interoperable, and reusable (FAIR; Wilkinson et al., 2016). In the context of this review, this applies not only to research data such as analyses of experimental products and natural samples, but also the tools and workflows used to perform calculations. There have already been a few advances in this regard, but there is still lots of work to do.

From a data availability perspective, a major advance came in 2008 when Hirschmann et al. (2008) created a web-hosted database of 6600 experiments (The Library of Experimental Phase Relationships—LEPR), which has been widely used to calibrate thermobarometers and chemometers (e.g., Jorgenson et al., 2022; Petrelli et al., 2020). This compilation represented a massive team effort, requiring most phase compositions and experimental contents to be hand typed from data tables in journal pdfs. However, the dataset is still incomplete in many aspects, stemming from missing data in the underlying publications, compounded with missing information during digitization. For example, in the LEPR download of Cpx-Liq pairs used by Petrelli et al. (2020) and Putirka (2008), 66% of experiments do not have a glass H_2O content. In some cases, this represents the fact that the experiment was anhydrous, in others the fact that authors didn't report water, and sometimes because H_2O is often presented in a different table from other chemical information, so may have been missed during digitization. Similarly, with few exceptions, the number of analyses used to produce the reported average for each phase (e.g., $N = 5$ Cpx analyses, $N = 8$ amphibole analyses) is not present in LEPR. This makes filtering out experiments with compositions defined by small numbers of analyses challenging, despite the importance of this given the influence of analytical error on minor, pressure-sensitive components (e.g., Cpx, Wieser et al., 2023d).

A bigger challenge is that there is no framework by which LEPR is regularly updated as new experimental studies are published. This has led to individual authors digitizing data from new experimental papers to use in their studies (Jorgenson et al., 2022; Wieser et al., 2023b,d). However, individual researchers who compile data may not have time to format the data for inclusion in the database, because under the current academic publishing model, they would gain no credit for this activity (citations for use of the database would go to the original author team from 2008). One option moving forwards may be for LEPR to frequently publish new short articles or Zenodo dois with an author team consisting of anyone who has contributed to the curation of the database (as for open-source tools such as Numpy; e.g., Dubois et al., 1996; Harris et al., 2020). Alternatively, the responsibility could be placed on people generating experimental data, where journal editors do not allow a paper to be accepted before the experimental data is included in these databases. However, in many cases, journals do not mandate authors submitting data to a specific repository. Community guidelines for exactly how such data should be presented and archived are clearly needed, as well as ways for data curation to be valued when hiring and promoting researchers (e.g., Klöcking et al., 2023).

From a modeling perspective, there has been a rapid increase in the number of open-source, easy to use tools to perform calculations, making workflows more user-friendly, faster, more reproducible, and less susceptible to user-error and version control issues. In general, calculation tools have evolved from GUIs, macro-enabled spreadsheets and web apps, many of which require users to type every single composition by hand, to python or R packages where calculations can be run on any number of rows in a user-supplied spreadsheet using Jupyter Notebooks or web apps built on top of underlying code (e.g., Iacovino and Till, 2019; Weber and Blundy, 2023). One major contribution toward more automated workflows for thermodynamic calculations has been the development of the ENKI ThermoEngine (Thermoengine Code Contributors, 2022), allowing the MELTS thermodynamic models and algorithms to be accessed through Python. This allows creation of modeling workflows or complete packages based on the underlying MELTS models. For example, the Fe-Ti oxybarometer and thermometer of Ghiorso and Evans (2008) was originally

released as a web-app where users had to hand-type major element values for each Fe-Ti oxide pair. Ghiorso and Prissel (2020) produced a Jupyter Notebook using Thermoengine which allows users to upload an excel file of matching Fe-Ti oxide pairs, and all calculations are performed automatically. The open-source Python3 package VESIcal (Iacovino et al., 2021) relies on Thermoengine to perform calculations using the solubility model Magmasat (which requires MELTS, Ghiorso and Gualda, 2015), as well as six other popular solubility models. Prior to VESIcal, saturation pressure calculations were performed using different tools for each solubility model in a myriad of different environments (Excel macros, web apps, web servers). The MELTS thermodynamic models can also be accessed using the alphaMELTS for MATLAB/Python packages (Antoshechkina and Ghiorso, 2018). As the alphaMELTS for MATLAB/Python packages (and Thermoengine) requires some familiarity with the underlying language, easy-to-use, wrapped functions for common workflows have been released through the PetThermoTools package (used for calculations in Section "Thermobarometers based on thermodynamic modeling," Gleeson and Wieser, 2023).

Calculations using the thermodynamic models/databases developed by Holland and Powell (e.g., Holland and Powell, 1998; Holland et al., 2018), which represent the main thermodynamic alternative to MELTS in igneous petrology, have traditionally performed using a variety of different software packages, including THERMOCALC, Perple_X (Connolly, 2009, 2005) and Theriak-Domino (de Capitani and Petrakakis, 2010). The learning curve for these tools is typically quite steep, and it remains challenging to integrate the thermodynamic calculations with common coding languages such as MATLAB, Python3 or Julia. Development of MAGEMin, a C-based Gibbs Free Energy minimization software with a Julia interface (Riel et al., 2022), improves the ease with which thermodynamic calculations can be performed using the Holland et al. (2018) thermodynamic database. This is because the Julia MAGEMin functions can also be called from Python, allowing the PetThermoTools package to perform the same calculations (e.g., LLD, phase stability, phase diagrams) using MELTS or the Holland et al. (2018) thermodynamic models. Such advances are vital for model intercomparison (e.g., Fig. 26).

For mineral-melt thermobarometry and chemometry, calculations using 100 s of different popular equations can be performed in Thermobar (Wieser et al., 2022c). This tool replaces a very large number of separate Excel spreadsheets, R and Matlab code released by individual thermobarometry papers, and allows more complex functions such as assessing all possible pairs for equilibrium, error propagation, etc. For thermobarometers applicable to determining the conditions of mantle melting, Python package meltPT (McNab and Ball, 2023) can perform calculations using 12 published expressions.

These open-source python tools allow investigation of science questions that were almost impossible to address using existing tools requiring extensive manual data input. For example, Wieser et al., 2022b use VESIcal to perform detailed comparisons of solubility models, including their sensitivities to different parameters. Wieser et al., 2023b,d use Thermobar to propagate thermobarometry uncertainties using Monte Carlo methods. Open-source packages also have the advantage that the source code is version-controlled, and hosted in a public repository (e.g., on GitHub) for anyone to inspect and adapt. This means even if the authors stop supporting the packages (perhaps moving from academia into industry), packages could be updated by a third party, although as a community we need to develop clear guidelines as to how academic credit can be allocated for such an activity. This is in stark constant with numerous tools which have simply become unusable with age. For example, the volatile solubility model of Duan (2014) is no longer available at the website listed in the paper, requiring authors to rely on archived downloads (e.g., Allison et al., 2022 rely on a version archived in 2017).

Conclusions

A wide diversity of methods have been used to determine the pressures, temperatures, H_2O contents and fO_2 conditions at which magmas are stored and staged as they transit from the mantle to the surface. Many thermometry methods work very well, indicating that silicate melt (±mineral) compositions are sensitive to temperature. However, it becomes more challenging to determine temperatures in subsolidus or highly evolved systems. In contrast, many mineral-based and thermodynamical barometers perform extremely poorly, struggling to reliably distinguish between different storage regions in the crust within ~10–20 km. This partially results from thermodynamic limitations; mineral compositions are not overly sensitive to the relatively narrow range of crustal pressures in the vast majority of tectonic settings on Earth. This fundamental limitation is not helped by sources of analytical and experimental uncertainty (e.g., poor analytical precision for minor components, missing elements such as Cr, Wieser et al., 2023d). Melt inclusion barometers have the potential to be more precise, because the dissolution of CO_2 and H_2O is highly sensitive to pressure. However, extensive work is required to resolve discrepancies between solubility models, as well as evaluate the influence of the vapor bubble CO_2 in many published datasets. Raman-based fluid inclusion barometry has great potential in specific systems, although further work is needed to investigate the influence of decrepitation, elastic relaxation, and the influence of other fluid species (e.g., SO_2, H_2O, etc.). Thermoba-Raman-try also shows great promise, but further work is required. We propose a number of key frontiers to address in the pursuit of better thermobarometers, hygrometers and oxybarometers:

(1) The development of higher quality experimental and thermodynamic datasets with precisely constrained mineral compositions, fluid compositions, dissolved volatile contents, and experimental fO_2 for recalibration of thermodynamic models (e.g., MELTS), mineral-melt thermobarometers, and volatile solubility models. Sufficient experiments need to be performed to allow full isolation of a test dataset during model calibration. More experiments using mixed H_2O-CO_2 fluids are needed to understand phase stability in relatively CO_2-rich arc magmas.

(2) Development of more robust infrastructure for compiling experimental and thermodynamic data, alongside sufficient metadata to assess data quality. This will ensure that application of new and exciting machine learning techniques will not suffer from the old computer science adage: "garbage-in, garbage-out."
(3) Further development of open-source methods to perform calculations, allowing easy intercomparison between models, error propagation, and model updates as new data becomes available.

Useful websites and open-source tools

Tools for petrologists—Contains Excel workbooks and python scripts for a number of petrographic workflows	http://www.kaylaiacovino.com/tools-for-petrologists/
VESIcal GitHub, Read The Docs and YouTube pages: Code, documentation and worked examples for calculations using VESIcal	https://github.com/kaylai/VESIcal, https://vesical.readthedocs.io/en/latest/
Thermobar read the docs page: Code, documentation and worked examples for thermobarometry and hygrometry calculations using Thermobar	https://github.com/PennyWieser/Thermobar, https://thermobar.readthedocs.io/en/latest/
Putirka spreadsheet compilation: Contains Excel spreadsheets for different thermobarometry calculations	https://sites.google.com/mail.fresnostate.edu/keithputirka/home
MELTS resources: Links to download different MELTS calculation tools (Matlab, python, GUI, excel)	https://magmasource.caltech.edu/gitlist
MAGEMin: Gibbs free energy minimization solver package	https://github.com/ComputationalThermodynamics/MAGEMin
PetThermoTools: Python package for performing common thermodynamic calculations with MAGEMin or MELTS	https://github.com/gleesonm1/PetThermoTools https://github.com/gleesonm1/pyMELTScalc
ENKI server: Allows calculations to be performed using ThermoEngine infrastructure without the need for local installation	http://enki-portal.org/
NIST Web Book: Phase data for CO_2 and calculations using the CO_2 equation of state	https://webbook.nist.gov/cgi/cbook.cgi?ID=C124389&Mask=1
DiadFit: Code, documentation and worked examples using DiadFit for fitting Raman data and performing EOS calculations for CO_2	https://github.com/PennyWieser/DiadFit, https://diadfit.readthedocs.io/en/latest/
CoolProp: A python package for calculating thermodynamic properties of gases using different equation of states	http://www.coolprop.org/

Data availability

Jupyter Notebooks and Excel workbooks used to make each figure are available on GitHub (https://github.com/PennyWieser/Thermobarometry_Review_2023).

Acknowledgments

PW thanks Iris Buisman for extracting the Cpx counting statistics shown in Fig. 8, and Keith Putirka for many useful conversations on thermobarometry and the formulation and calibration of his equations. PW acknowledges support from NSF EAR **2217371, the Rose Hills Innovator Program, and start up funds to UC Berkeley for computing resources. EG acknowledges support from NSF** EAR 2216738.

Author contributions

PW coordinated the review and was responsible for writing and coding for all sections except those listed below. MG was responsible for the OPAM, MELTS and MAGEMin sections with edits from PW, and SM for Ol-Sp and Fe-Ti oxide sections. CD contributed text to the MI and FI section, and provided edits throughout. CD and EG contributed to analytical work. All authors reviewed and edited the text.

References

Acosta MD, Watkins JM, Reed MH, Donovan JJ, and DePaolo DJ (2020) Ti-in-quartz: Evaluating the role of kinetics in high temperature crystal growth experiments. *Geochimica et Cosmochimica Acta* 281: 149–167. https://doi.org/10.1016/j.gca.2020.04.030.

Afonso JC, Ranalli G, and Fernàndez M (2007) Density structure and buoyancy of the oceanic lithosphere revisited. *Geophysical Research Letters* 34: L10302. https://doi.org/10.1029/2007GL029515.

Ague JJ (1997) Thermodynamic calculation of emplacement pressures for batholithic rocks, California: Implications for the aluminum-in-hornblende barometer. *Geology* 25: 563. https://doi.org/10.1130/0091-7613(1997)025<0563:TCOEPF>2.3.CO;2.
Allison CM, Roggensack K, and Clarke AB (2021) Highly explosive basaltic eruptions driven by CO2 exsolution. *Nature Communications* 12: 217. https://doi.org/10.1038/s41467-020-20354-2.
Allison CM, Roggensack K, and Clarke AB (2022) MafiCH: A general model for H2O–CO2 solubility in mafic magmas. *Contributions to Mineralogy and Petrology* 177: 40. https://doi.org/10.1007/s00410-022-01903-y.
Almeev RR, Bolte T, Nash BP, Holtz F, Erdmann M, and Cathey HE (2012) High-temperature, low-H2O silicic magmas of the Yellowstone hotspot: An experimental study of rhyolite from the Bruneau–Jarbidge Eruptive Center, Central Snake River Plain, USA. *Journal of Petrology* 53: 1837–1866. https://doi.org/10.1093/petrology/egs035.
Alonso-Perez R, Müntener O, and Ulmer P (2009) Igneous garnet and amphibole fractionation in the roots of island arcs: Experimental constraints on andesitic liquids. *Contributions to Mineralogy and Petrology* 157: 541–558. https://doi.org/10.1007/s00410-008-0351-8.
Andersen DJ and Lindsley DH (1988) Internally consistent solution models for Fe-Mg-Mn-Ti oxides; Fe-Ti oxides. *American Mineralogist* 73: 714–726.
Andersen T, O'Reilly SY, and Griffin WL (1984) The trapped fluid phase in upper mantle xenoliths from Victoria, Australia: Implications for mantle metasomatism. *Contributions to Mineralogy and Petrology* 88: 72–85. https://doi.org/10.1007/BF00371413.
Anderson JL (1996) Status of thermobarometry in granitic batholiths. In: *The Third Hutton Symposium on the Origin of Granites and Related Rocks*. Geological Society of America. https://doi.org/10.1130/0-8137-2315-9.125.
Anderson AT and Brown (1993) CO2 contents and formation pressures of some Kilauean melt inclusions. *American Mineralogist* 78: 794–803.
Anderson AT, Davis AM, and Lu F (2000) Evolution of Bishop Tuff rhyolitic magma based on melt and magnetite inclusions and zoned phenocrysts. *Journal of Petrology* 41: 449–473.
Anderson JL, Barth AP, Wooden JL, and Mazdab F (2008) Thermometers and thermobarometers in granitic systems. *Reviews in Mineralogy and Geochemistry* 69: 121–142. https://doi.org/10.2138/rmg.2008.69.4.
Andújar J, Scaillet B, Pichavant M, and Druitt TH (2015) Differentiation conditions of a basaltic magma from santorini, and its bearing on the production of andesite in arc settings. *Journal of Petrology* 56: 765–794. https://doi.org/10.1093/petrology/egv016.
Antoshechkina P and Ghiorso MS (2018) MELTS for MATLAB: A new educational and research tool for computational thermodynamics. In: *AGU Fall Meeting, abstract #ED44B-23*.
Aradi L, Spránitz T, Guzmics T, Szabó C, and Berkesi M (2021) 3D Raman imaging of multiphase fluid and melt inclusions: Challenges and perspectives. In: *Goldschmidt 2021 Abstracts. Presented at the Goldschmidt 2021*. European Association of Geochemistry. https://doi.org/10.7185/gold2021.8071. Virtual.
Araya N, Nakamura M, Yasuda A, Okumura S, Sato T, Iguchi M, Miki D, and Geshi N (2019) Shallow magma pre-charge during repeated Plinian eruptions at Sakurajima volcano. *Scientific Reports* 9: 1979. https://doi.org/10.1038/s41598-019-38494-x.
Ariskin AA, Frenkel MY, Barmina GS, and Nielsen RL (1993) Comagmat: A Fortran program to model magma differentiation processes. *Computers & Geosciences* 19: 1155–1170. https://doi.org/10.1016/0098-3004(93)90020-6.
Armienti P, Perinelli C, and Putirka KD (2013) A new model to estimate deep-level magma ascent rates, with applications to Mt. Etna (Sicily, Italy). *Journal of Petrology* 54: 795–813. https://doi.org/10.1093/petrology/egs085.
Aster EM, Wallace PJ, Moore LR, Watkins J, Gazel E, and Bodnar RJ (2016) Reconstructing CO2 concentrations in basaltic melt inclusions using Raman analysis of vapor bubbles. *Journal of Volcanology and Geothermal Research* 323: 148–162. https://doi.org/10.1016/j.jvolgeores.2016.04.028.
Azbej T, Severs MJ, Rusk BG, and Bodnar RJ (2007) In situ quantitative analysis of individual H2O–CO2 fluid inclusions by laser Raman spectroscopy. *Chemical Geology* 237: 255–263. https://doi.org/10.1016/j.chemgeo.2006.06.025.
Bacon CR, Newman S, and Stolper E (1992) Water, CO2, Cl, and F in melt inclusions in phenocrysts from three Holocene explosive eruptions, Crater Lake, Oregon. *American Mineralogist* 77: 1021–1030.
Bakker R (2017) Re-equilibration processes in fluid inclusion assemblages. *Minerals* 7: 117. https://doi.org/10.3390/min7070117.
Bali E, Hartley ME, Halldórsson SA, Gudfinnsson GH, and Jakobsson S (2018) Melt inclusion constraints on volatile systematics and degassing history of the 2014–2015 Holuhraun eruption, Iceland. *Contributions to Mineralogy and Petrology* 173: 9. https://doi.org/10.1007/s00410-017-1434-1.
Barclay J (2004) A hornblende basalt from Western Mexico: Water-saturated phase relations constrain a pressure-temperature window of eruptibility. *Journal of Petrology* 45: 485–506. https://doi.org/10.1093/petrology/egg091.
Barker AK, Rydeblad EM, and Silva SM (2021) Magma storage at Ocean Islands: Insights from Cape Verde. In: *Crustal Magmatic System Evolution: Anatomy, Architecture, and Physico-Chemical Processes*, pp. 45–78. https://doi.org/10.1002/9781119564485.ch3.
Barth A and Plank T (2021) The ins and outs of water in olivine-hosted melt inclusions: Hygrometer vs. speedometer. *Frontiers in Earth Science* 9: 614004. https://doi.org/10.3389/feart.2021.614004.
Barth A, Newcombe M, Plank T, Gonnermann H, Hajimirza S, Soto GJ, Saballos A, and Hauri E (2019) Magma decompression rate correlates with explosivity at basaltic volcanoes—Constraints from water diffusion in olivine. *Journal of Volcanology and Geothermal Research* 387: 106664. https://doi.org/10.1016/j.jvolgeores.2019.106664.
Baxter RJM, Maclennan J, Neave DA, and Thordarson T (2023) Depth of magma storage under Iceland controlled by magma fluxes. *Geochemistry, Geophysics, Geosystems* 24: e2022GC010811. https://doi.org/10.1029/2022GC010811.
Beattie P (1993) Olivine-melt and orthopyroxene-melt equilibria. *Contributions to Mineralogy and Petrology* 115: 103–111. https://doi.org/10.1007/BF00712982.
Befus KS, Lin J-F, Cisneros M, and Fu S (2018) Feldspar Raman shift and application as a magmatic thermobarometer. *American Mineralogist* 103: 600–609. https://doi.org/10.2138/am-2018-6157.
Bégué F, Deering CD, Gravley DM, Kennedy BM, Chambefort I, Gualda GAR, and Bachmann O (2014a) Extraction, storage and eruption of multiple isolated magma batches in the paired Mamaku and Ohakuri Eruption, Taupo Volcanic Zone, New Zealand. *Journal of Petrology* 55: 1653–1684. https://doi.org/10.1093/petrology/egu038.
Bégué F, Gualda GAR, Ghiorso MS, Pamukcu AS, Kennedy BM, Gravley DM, Deering CD, and Chambefort I (2014b) Phase-equilibrium geobarometers for silicic rocks based on rhyolite-MELTS. Part 2: Application to Taupo Volcanic Zone rhyolites. *Contributions to Mineralogy and Petrology* 168: 1082. https://doi.org/10.1007/s00410-014-1082-7.
Bégué F, Gravley DM, Chambefort I, Deering CD, and Kennedy BM (2015) Magmatic volatile distribution as recorded by rhyolitic melt inclusions in the Taupo Volcanic Zone, New Zealand. *Geological Society, London, Special Publications* 410: 71–94. https://doi.org/10.1144/SP410.4.
Bell IH, Wronski J, Quoilin S, and Lemort V (2014) Pure and pseudo-pure fluid thermophysical property evaluation and the open-source thermophysical property library CoolProp. *Industrial and Engineering Chemistry Research* 53: 2498–2508. https://doi.org/10.1021/ie4033999.
Bell AF, La Femina PC, Ruiz M, Amelung F, Bagnardi M, Bean CJ, Bernard B, Ebinger C, Gleeson M, Grannell J, Hernandez S, Higgins M, Liorzou C, Lundgren P, Meier NJ, Möllhoff M, Oliva S-J, Ruiz AG, and Stock MJ (2021) Caldera resurgence during the 2018 eruption of Sierra Negra volcano, Galápagos Islands. *Nature Communications* 12: 1397. https://doi.org/10.1038/s41467-021-21596-4.
Bennett EN, Jenner FE, Millet M-A, Cashman KV, and Lissenberg CJ (2019) Deep roots for mid-ocean-ridge volcanoes revealed by plagioclase-hosted melt inclusions. *Nature* 572: 235–239. https://doi.org/10.1038/s41586-019-1448-0.
Berman RG (1988) Internally-consistent thermodynamic data for minerals in the system Na2O-K2O-CaO-MgO-FeO-Fe2O3-Al2O3-SiO2-TiO2-H2O-CO2. *Journal of Petrology* 29: 445–522. https://doi.org/10.1093/petrology/29.2.445.
Berndt J (2004) An experimental investigation of the influence of water and oxygen fugacity on differentiation of MORB at 200 MPa. *Journal of Petrology* 46: 135–167. https://doi.org/10.1093/petrology/egh066.
Black BA and Andrews BJ (2020) Petrologic imaging of the architecture of magma reservoirs feeding caldera-forming eruptions. *Earth and Planetary Science Letters* 552: 116572. https://doi.org/10.1016/j.epsl.2020.116572.
Blatter DL and Carmichael ISE (2001) Hydrous phase equilibria of a Mexican high-silica andesite: A candidate for a mantle origin? *Geochimica et Cosmochimica Acta* 65: 4043–4065. https://doi.org/10.1016/S0016-7037(01)00708-6.

Blatter DL, Sisson TW, and Hankins WB (2013) Crystallization of oxidized, moderately hydrous arc basalt at mid- to lower-crustal pressures: Implications for andesite genesis. *Contributions to Mineralogy and Petrology* 166: 861–886. https://doi.org/10.1007/s00410-013-0920-3.

Blatter DL, Sisson TW, and Hankins WB (2023) Garnet stability in arc basalt, andesite, and dacite—An experimental study. *Contributions to Mineralogy and Petrology* 178. https://doi.org/10.1007/s00410-023-02008.

Blundy J (2022) Chemical differentiation by mineralogical buffering in crustal hot zones. *Journal of Petrology* 63: egac054. https://doi.org/10.1093/petrology/egac054.

Blundy J and Cashman K (2005) Rapid decompression-driven crystallization recorded by melt inclusions from Mount St. Helens volcano. *Geology* 33(10): 793–796.

Blundy J and Cashman K (2008) Petrologic reconstruction of magmatic system variables and processes. *Reviews in Mineralogy and Geochemistry* 69: 179–239. https://doi.org/10.2138/rmg.2008.69.6.

Blundy J, Cashman KV, Rust A, and Witham F (2010) A case for CO2-rich arc magmas. *Earth and Planetary Science Letters* 290: 289–301. https://doi.org/10.1016/j.epsl.2009.12.013.

Bodnar RJ (2017) Fluid inclusions. In: Bobrowsky P and Marker B (eds.) *Encyclopedia of Engineering Geology. Encyclopedia of Earth Sciences Series*, pp. 1–5. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-39193-9_225-1.

Bodnar RJ, Binns PR, and Hall DL (1989) Synthetic fluid inclusions—VI. Quantitative evaluation of the decrepitation behaviour of fluid inclusions in quartz at one atmosphere confining pressure. *Journal of Metamorphic Geology* 7: 229–242. https://doi.org/10.1111/j.1525-1314.1989.tb00586.x.

Bogaerts M, Scaillet B, and Auwera JV (2006) Phase equilibria of the Lyngdal Granodiorite (Norway): Implications for the origin of metaluminous ferroan granitoids. *Journal of Petrology* 47: 2405–2431. https://doi.org/10.1093/petrology/egl049.

Bohrson WA and Clague DA (1988) Origin of ultramafic xenoliths containing exsolved pyroxenes from Hualalai Volcano, Hawaii. *Contributions to Mineralogy and Petrology* 100: 139–155. https://doi.org/10.1007/BF00373581.

Bolte T, Holtz F, Almeev R, and Nash B (2015) The Blacktail Creek Tuff: An analytical and experimental study of rhyolites from the Heise volcanic field, Yellowstone hotspot system. *Contributions to Mineralogy and Petrology* 169: 15. https://doi.org/10.1007/s00410-015-1112-0.

Böttcher N, Taron J, Kolditz O, Liedl R, and Park C-H (2012) Comparison of equations of state for carbon dioxide for numerical simulations. In: *Proceedings ModelCARE2011, Leipzig, Germany. IAHS Publ. 355*, p. 9.

Boulanger M, France L, Deans JRL, Ferrando C, Lissenberg CJ, and von der Handt A (2020) Magma reservoir formation and evolution at a slow-spreading center (Atlantis Bank, Southwest Indian Ridge). *Frontiers in Earth Science* 8: 554598. https://doi.org/10.3389/feart.2020.554598.

Boullier A-M, Michot G, Pecher A, and Barres O (1989) Diffusion and/or plastic deformation around fluid inclusions in synthetic quartz: New investigations. In: Bridgwater D (ed.) *Fluid Movements—Element Transport and the Composition of the Deep Crust*, pp. 345–360. Netherlands, Dordrecht: Springer. https://doi.org/10.1007/978-94-009-0991-5_28.

Brugman KK and Till CB (2019) A low-aluminum clinopyroxene-liquid geothermometer for high-silica magmatic systems. *American Mineralogist* 104: 996–1004. https://doi.org/10.2138/am-2019-6842.

Bucholz CE, Gaetani GA, Behn MD, and Shimizu N (2013) Post-entrapment modification of volatiles and oxygen fugacity in olivine-hosted melt inclusions. *Earth and Planetary Science Letters* 374: 145–155. https://doi.org/10.1016/j.epsl.2013.05.033.

Budd DA, Troll VR, Dahren B, and Burchardt S (2016) Persistent multitiered magma plumbing beneath Katla volcano, Iceland. *Geochemistry, Geophysics, Geosystems* 17: 966–980. https://doi.org/10.1002/2015GC006118.

Buddington AF and Lindsley DH (1964) Iron-titanium oxide minerals and synthetic equivalents. *Journal of Petrology* 5: 310–357. https://doi.org/10.1093/petrology/5.2.310.

Burke EAJ (2001) Raman microspectrometry of fluid inclusions. *Lithos* 55: 139–158. https://doi.org/10.1016/S0024-4937(00)00043-8.

Buso R, Laporte D, Schiavi F, Cluzel N, and Fonquernie C (2022) High-pressure homogenization of olivine-hosted CO_2-rich melt inclusions in a piston cylinder: Insight into the volatile content of primary mantle melts. *European Journal of Mineralogy* 34: 325–349. https://doi.org/10.5194/ejm-34-325-2022.

Cadoux A, Scaillet B, Druitt TH, and Deloule E (2014) Magma storage conditions of large Plinian eruptions of Santorini volcano (Greece). *Journal of Petrology* 55: 1129–1171. https://doi.org/10.1093/petrology/egu021.

Campione M, Malaspina N, and Frezzotti ML (2015) Threshold size for fluid inclusion decrepitation: Size for fluid inclusion decrepitation. *Journal of Geophysical Research: Solid Earth* 120: 7396–7402. https://doi.org/10.1002/2015JB012086.

Caracciolo A, Bali E, Guðfinnsson GH, Kahl M, Halldórsson SA, Hartley ME, and Gunnarsson H (2020) Temporal evolution of magma and crystal mush storage conditions in the Bárðarbunga-Veiðivötn volcanic system, Iceland. *Lithos* 352–353: 105234. https://doi.org/10.1016/j.lithos.2019.105234.

Caracciolo A, Halldórsson SA, Bali E, Marshall EW, Jeon H, Whitehouse MJ, Barnes JD, Guðfinnsson GH, Kahl M, and Hartley ME (2022) Oxygen isotope evidence for progressively assimilating trans-crustal magma plumbing systems in Iceland. *Geology* 50: 796–800. https://doi.org/10.1130/G49874.1.

Chakraborty S and Dohmen R (2022) Diffusion chronometry of volcanic rocks: Looking backward and forward. *Bulletin of Volcanology* 84: 57. https://doi.org/10.1007/s00445-022-01565-5.

Chen YJ (1992) Oceanic crustal thickness versus spreading rate. *Geophysical Research Letters* 19(8): 753–756. https://doi.org/10.1029/92GL00161.

Cheng L-L, Yang Z-F, Zeng L, Wang Y, and Luo Z-H (2014) Giant plagioclase growth during storage of basaltic magma in Emeishan Large Igneous Province, SW China. *Contributions to Mineralogy and Petrology* 167: 971. https://doi.org/10.1007/s00410-014-0971-0.

Chesner CA and Luhr JF (2010) A melt inclusion study of the Toba Tuffs, Sumatra, Indonesia. *Journal of Volcanology and Geothermal Research* 197: 259–278. https://doi.org/10.1016/j.jvolgeores.2010.06.001.

Cisneros M and Befus KS (2020) Applications and limitations of elastic thermobarometry: Insights from elastic modeling of inclusion-host pairs and example case studies. *Geochemistry, Geophysics, Geosystems* 21(10): e2020GC009231. https://doi.org/10.1029/2020GC009231.

Connolly JAD (2005) Computation of phase equilibria by linear programming: A tool for geodynamic modeling and its application to subduction zone decarbonation. *Earth and Planetary Science Letters* 236: 524–541. https://doi.org/10.1016/j.epsl.2005.04.033.

Connolly JAD (2009) The geodynamic equation of state: What and how: Geodynamic equation of state-what and how. *Geochemistry, Geophysics, Geosystems* 10. https://doi.org/10.1029/2009GC002540.

Coogan LA, Saunders AD, and Wilson RN (2014) Aluminum-in-olivine thermometry of primitive basalts: Evidence of an anomalously hot mantle source for large igneous provinces. *Chemical Geology* 368: 1–10. https://doi.org/10.1016/j.chemgeo.2014.01.004.

Costa F (2004) Petrological and experimental constraints on the pre-eruption conditions of Holocene dacite from Volcan San Pedro (36 S, Chilean Andes) and the importance of sulphur in silicic subduction-related magmas. *Journal of Petrology* 45: 855–881. https://doi.org/10.1093/petrology/egg114.

Costa F, Shea T, and Ubide T (2020) Diffusion chronometry and the timescales of magmatic processes. *Nature Reviews Earth and Environment* 1: 201–214. https://doi.org/10.1038/s43017-020-0038-x.

Crabtree SM and Lange RA (2011) Complex phenocryst textures and zoning patterns in andesites and dacites: Evidence of degassing-induced rapid crystallization? *Journal of Petrology* 52: 3–38. https://doi.org/10.1093/petrology/egq067.

Dahren B, Troll VR, Andersson UB, Chadwick JP, Gardner MF, Jaxybulatov K, and Koulakov I (2012) Magma plumbing beneath Anak Krakatau volcano, Indonesia: Evidence for multiple magma storage regions. *Contributions to Mineralogy and Petrology* 163: 631–651. https://doi.org/10.1007/s00410-011-0690-8.

Dal-Negro A, Manoli S, Secca L, and Piccirillo EM (1989) Megacrystic clinopyroxenes from Victoria (Australia); crystal chemical comparisons of pyroxenes from high and low pressure regimes. *European Journal of Mineralogy* 1(1): 105–121.

Danyushevsky LV and Plechov P (2011) Petrolog3: Integrated software for modeling crystallization processes. *Geochemistry, Geophysics, Geosystems* 12. https://doi.org/10.1029/2011GC003516.

Danyushevsky LV, McNeill AW, and Sobolev AV (2002) Experimental and petrological studies of melt inclusions in phenocrysts from mantle-derived magmas: An overview of techniques, advantages and complications. *Chemical Geology* 183: 5–24. https://doi.org/10.1016/S0009-2541(01)00369-2.

Dayton K, Gazel E, Wieser P, Troll VR, Carracedo JC, La Madrid H, Roman DC, Ward J, Aulinas M, Geiger H, Deegan FM, Gisbert G, and Perez-Torrado FJ (2023) Deep magma storage during the 2021 La Palma eruption. *Science Advances* 9: eade7641. https://doi.org/10.1126/sciadv.ade7641.

de Capitani C and Petrakakis K (2010) The computation of equilibrium assemblage diagrams with Theriak/Domino software. *American Mineralogist* 95: 1006–1016. https://doi.org/10.2138/am.2010.3354.

De Hoog JC, Gall L, and Cornell DH (2010) Trace-element geochemistry of mantle olivine and application to mantle petrogenesis and geothermobarometry. *Chemical Geology* 270(1–4): 196–215.

Demouchy S, Jacobsen SD, Gaillard F, and Stern CR (2006) Rapid magma ascent recorded by water diffusion profiles in mantle olivine. *Geology* 34: 429. https://doi.org/10.1130/G22386.1.

DeVitre CL, Allison CM, and Gazel E (2021) A high-precision CO2 densimeter for Raman spectroscopy using a fluid density calibration apparatus. *Chemical Geology* 584: 120522. https://doi.org/10.1016/j.chemgeo.2021.120522.

DeVitre CL, Gazel E, Ramalho RS, Venugopal S, Steele-MacInnis M, Hua J, Allison CM, Moore LR, Carracedo JC, and Monteleone B (2023a) Oceanic intraplate explosive eruptions fed directly from the mantle. *Proceedings. National Academy of Sciences. United States of America* 120: e2302093120. https://doi.org/10.1073/pnas.2302093120.

DeVitre CL, Dayton K, Gazel E, Pamukçu A, Gaetani G, and Wieser PE (2023b) Laser heating effect on Raman analysis of CO2 co-existing as liquid and vapor in olivine-hosted melt inclusion bubbles. *Volcanica* 6: 201–219. https://doi.org/10.30909/vol.06.02.201219.

Dixon JE (1997) Degassing of alkalic basalts. *American Mineralogist* 82: 368–378. https://doi.org/10.2138/am-1997-3-415.

Dobbin KK and Simon RM (2011) Optimally splitting cases for training and testing high dimensional classifiers. *BMC Medical Genomics* 4: 31. https://doi.org/10.1186/1755-8794-4-31.

Drignon MJ, Nielsen RL, Tepley FJ, and Bodnar RJ (2019) Upper mantle origin of plagioclase megacrysts from plagioclase-ultraphyric mid-oceanic ridge basalt. *Geology* 47: 43–46. https://doi.org/10.1130/G45542.1.

Duan X (2014) A general model for predicting the solubility behavior of H2O–CO2 fluids in silicate melts over a wide range of pressure, temperature and compositions. *Geochimica et Cosmochimica Acta* 125: 582–609. https://doi.org/10.1016/j.gca.2013.10.018.

Dubessy J, Caumon M-C, Rull F, and Sharma S (2012) Instrumentation in Raman spectroscopy: Elementary theory and practice. In: Ferraris G, Dubessy J, Caumon M-C, and Rull F (eds.) *Raman Spectroscopy Applied to Earth Sciences and Cultural Heritage*, pp. 83–172. European Mineralogical Union. https://doi.org/10.1180/EMU-notes.12.3.

Dubois PF, Hinsen K, and Hugunin J (1996) Numerical python. *Computers in Physics* 10: 262. https://doi.org/10.1063/1.4822400.

Erdmann M and Koepke J (2016) Silica-rich lavas in the oceanic crust: Experimental evidence for fractional crystallization under low water activity. *Contributions to Mineralogy and Petrology* 171: 83. https://doi.org/10.1007/s00410-016-1294-0.

Erdmann S, Martel C, Pichavant M, and Kushnir A (2014) Amphibole as an archivist of magmatic crystallization conditions: Problems, potential, and implications for inferring magma storage prior to the paroxysmal 2010 eruption of Mount Merapi, Indonesia. *Contributions to Mineralogy and Petrology* 167: 1016. https://doi.org/10.1007/s00410-014-1016-4.

Erdmann S, Martel C, Pichavant M, Bourdier J-L, Champallier R, Komorowski J-C, and Cholik N (2016) Constraints from phase equilibrium experiments on pre-eruptive storage conditions in mixed magma systems: A case study on crystal-rich basaltic andesites from Mount Merapi, Indonesia. *Journal of Petrology* 57: 535–560. https://doi.org/10.1093/petrology/egw019.

Esposito R, Klebesz R, Bartoli O, Klyukin Y, Moncada D, Doherty A, and Bodnar R (2012) Application of the Linkam TS1400XY heating stage to melt inclusion studies. *Open Geosciences* 4. https://doi.org/10.2478/s13533-011-0054-y.

Esposito R, Lamadrid HM, Redi D, Steele-MacInnis M, Bodnar RJ, Manning CE, De Vivo B, Cannatelli C, and Lima A (2016) Detection of liquid H_2O in vapor bubbles in reheated melt inclusions: Implications for magmatic fluid composition and volatile budgets of magmas? *American Mineralogist* 101: 1691–1695. https://doi.org/10.2138/am-2016-5689.

Everall NJ (2010) Confocal Raman microscopy: Common errors and artefacts. *Analyst* 135: 2512. https://doi.org/10.1039/c0an00371a.

Fall A, Tattitch B, and Bodnar RJ (2011) Combined microthermometric and Raman spectroscopic technique to determine the salinity of H2O–CO2–NaCl fluid inclusions based on clathrate melting. *Geochimica et Cosmochimica Acta* 75: 951–964. https://doi.org/10.1016/j.gca.2010.11.021.

First EC, Hammer JE, Ruprecht P, and Rutherford M (2021) Experimental constraints on dacite magma storage beneath Volcán Quizapu, Chile. *Journal of Petrology* 62: egab027. https://doi.org/10.1093/petrology/egab027.

Fowler SJ and Spera FJ (2010) A metamodel for crustal magmatism: Phase equilibria of giant ignimbrites. *Journal of Petrology* 51: 1783–1830. https://doi.org/10.1093/petrology/egq039.

Fowler SJ, Spera FJ, Bohrson WA, Belkin HE, and De Vivo B (2007) Phase equilibria constraints on the chemical and physical evolution of the campanian ignimbrite. *Journal of Petrology* 48: 459–493. https://doi.org/10.1093/petrology/egl068.

Frezzotti ML, De Vivo B, and Clocchiatti R (1991) Melt-mineral-fluid interactions in ultramafic nodules from alkaline lavas of Mount Etna (Sicily, Italy): Melt and fluid inclusion evidence. *Journal of Volcanology and Geothermal Research* 47: 209–219. https://doi.org/10.1016/0377-0273(91)90001-G.

Frezzotti ML, Andersen T, Neumann E-R, and Simonsen SL (2002) Carbonatite melt–CO2 fluid inclusions in mantle xenoliths from Tenerife, Canary Islands: A story of trapping, immiscibility and fluid–rock interaction in the upper mantle. *Lithos* 64: 77–96. https://doi.org/10.1016/S0024-4937(02)00178-0.

Frezzotti ML, Peccerillo A, and Bonelli R (2003) Magma ascent rates and depths of crustal magma reservoirs beneath the Aeolian volcanic Arc (Italy): Inferences from fluid and melt inclusions in xenoliths. In: *Developments in Volcanology*, pp. 185–205. Elsevier. https://doi.org/10.1016/S1871-644X(03)80030-X.

Frezzotti ML, Tecce F, and Casagli A (2012) Raman spectroscopy for fluid inclusion analysis. *Journal of Geochemical Exploration* 112: 1–20. https://doi.org/10.1016/j.gexplo.2011.09.009.

Gaetani GA, O'Leary JA, Shimizu N, Bucholz CE, and Newville M (2012) Rapid reequilibration of H2O and oxygen fugacity in olivine-hosted melt inclusions. *Geology* 40: 915–918. https://doi.org/10.1130/G32992.1.

Gaffney AM (2002) Environments of crystallization and compositional diversity of Mauna Loa Xenoliths. *Journal of Petrology* 43: 963–981. https://doi.org/10.1093/petrology/43.6.963.

Gao R, Lassiter JC, Clague DA, and Bohrson WA (2022) Evolution of Hawaiian volcano magmatic plumbing system and implications for melt/edifice and melt/lithosphere interaction: Constraints from Hualālai Xenoliths. *Journal of Petrology* 63: egac091. https://doi.org/10.1093/petrology/egac091.

Garcia MO (2003) A petrologic perspective of Kilauea Volcano's summit magma reservoir. *Journal of Petrology* 44: 2313–2339. https://doi.org/10.1093/petrology/egg079.

Gardner JE, Befus KS, Gualda GAR, and Ghiorso MS (2014) Experimental constraints on rhyolite-MELTS and the Late Bishop Tuff magma body. *Contributions to Mineralogy and Petrology* 168: 1051. https://doi.org/10.1007/s00410-014-1051-1.

Gavrilenko M, Herzberg C, Vidito C, Carr MJ, Tenner T, and Ozerov A (2016) A calcium-in-olivine geohygrometer and its application to subduction zone magmatism. *Journal of Petrology* 57: 1811–1832. https://doi.org/10.1093/petrology/egw062.

Gavrilenko M, Krawczynski M, Ruprecht P, Li W, and Catalano JG (2019) The quench control of water estimates in convergent margin magmas. *American Mineralogist* 104: 936–948. https://doi.org/10.2138/am-2019-6735.

Geiger H, Barker AK, and Troll VR (2016a) Locating the depth of magma supply for volcanic eruptions, insights from Mt. Cameroon. *Scientific Reports* 6: 33629. https://doi.org/10.1038/srep33629.

Geiger H, Mattsson T, Deegan FM, Troll VR, Burchardt S, Gudmundsson Ó, Tryggvason A, Krumbholz M, and Harris C (2016b) Magma plumbing for the 2014–2015 Holuhraun eruption, Iceland. *Geochemistry, Geophysics, Geosystems* 17: 2953–2968. https://doi.org/10.1002/2016GC006317.

Geiger H, Troll VR, Jolis EM, Deegan FM, Harris C, Hilton DR, and Freda C (2018) Multi-level magma plumbing at Agung and Batur volcanoes increases risk of hazardous eruptions. *Scientific Reports* 8: 10547. https://doi.org/10.1038/s41598-018-28125-2.

Geist D, Naumann T, and Larson P (1998) Evolution of Galapagos magmas: Mantle and crustal fractionation without assimilation. *Journal of Petrology* 39: 953–971. https://doi.org/10.1093/petroj/39.5.953.

Gerlach TM and Graeber EJ (1985) Volatile budget of Kilauea volcano. *Nature* 313: 273–277. https://doi.org/10.1038/313273a0.

Ghiorso MS (1990) Thermodynamic properties of hematite—ilmenite—geikielite solid solutions. *Contributions to Mineralogy and Petrology* 104: 645–667. https://doi.org/10.1007/BF01167285.
Ghiorso MS and Evans BW (2008) Thermodynamics of rhombohedral oxide solid solutions and a revision of the FE-TI two-oxide geothermometer and oxygen-barometer. *American Journal of Science* 308: 957–1039. https://doi.org/10.2475/09.2008.01.
Ghiorso MS and Gualda GAR (2013) A method for estimating the activity of titania in magmatic liquids from the compositions of coexisting rhombohedral and cubic iron–titanium oxides. *Contributions to Mineralogy and Petrology* 165: 73–81. https://doi.org/10.1007/s00410-012-0792-y.
Ghiorso MS and Gualda GAR (2015) An H_2O–CO_2 mixed fluid saturation model compatible with rhyolite-MELTS. *Contributions to Mineralogy and Petrology* 169: 1–30. https://doi.org/10.1007/s00410-015-1141-8.
Ghiorso M and Prissel K (2020) *ENKI Cloud App: Implementation of the Fe-Ti Oxide Geothermooxybarometer of Ghiorso and Evans, 2008*. https://doi.org/10.5281/ZENODO.3866660.
Ghiorso MS and Sack O (1991) Fe-Ti oxide geothermometry: Thermodynamic formulation and the estimation of intensive variables in silicic magmas. *Contributions to Mineralogy and Petrology* 108: 485–510. https://doi.org/10.1007/BF00303452.
Ghiorso MS and Sack RO (1995) Chemical mass transfer in magmatic processes IV. A revised and internally consistent thermodynamic model for the interpolation and extrapolation of liquid-solid equilibria in magmatic systems at elevated temperatures and pressures. *Contributions to Mineralogy and Petrology* 119: 197–212. https://doi.org/10.1007/BF00307281.
Ghiorso MS, Hirschmann MM, Reiners PW, and Kress VC (2002) The pMELTS: A revision of MELTS for improved calculation of phase relations and major element partitioning related to partial melting of the mantle to 3GPa. *Geochemistry, Geophysics, Geosystems* 3: 1–35. https://doi.org/10.1029/2001GC000217.
Gibson SA and Geist D (2010) Geochemical and geophysical estimates of lithospheric thickness variation beneath Galápagos. *Earth and Planetary Science Letters* 300: 275–286. https://doi.org/10.1016/j.epsl.2010.10.002.
Gleeson MLM and Gibson SA (2021) Insights into the nature of plume-ridge interaction and outflux of H_2O from the Galápagos spreading center. *Geochemistry, Geophysics, Geosystems* 22. https://doi.org/10.1029/2020GC009560.
Gleeson MLM, Stock MJ, Pyle DM, Mather TA, Hutchison W, Yirgu G, and Wade J (2017) Constraining magma storage conditions at a restless volcano in the Main Ethiopian Rift using phase equilibria models. *Journal of Volcanology and Geothermal Research* 337: 44–61. https://doi.org/10.1016/j.jvolgeores.2017.02.026.
Gleeson MLM, Gibson SA, and Stock MJ (2021) Upper mantle mush zones beneath low melt flux ocean island volcanoes: Insights from Isla Floreana, Galápagos. *Journal of Petrology* 61: egaa094. https://doi.org/10.1093/petrology/egaa094.
Gleeson MLM, Gibson SA, and Stock MJ (2022) Constraints on the behaviour and content of volatiles in Galápagos magmas from melt inclusions and nominally anhydrous minerals. *Geochimica et Cosmochimica Acta* 319: 168–190. https://doi.org/10.1016/j.gca.2021.11.005.
Gleeson and Wieser (2023) PetThermoTools v0.2.0. *Zenodo*. https://doi.org/10.5281/zenodo.7758493.
Goltz AE, Krawczynski MJ, Gavrilenko M, Gorbach NV, and Ruprecht P (2020) Evidence for superhydrous primitive arc magmas from mafic enclaves at Shiveluch volcano, Kamchatka. *Contributions to Mineralogy and Petrology* 175: 115. https://doi.org/10.1007/s00410-020-01746-5.
Goranson RW (1931) The solubility of water in granite magmas. *American Journal of Science* s5-22: 481–502. https://doi.org/10.2475/ajs.s5-22.132.481.
Grove TL, Kinzler RJ, and Bryan WB (1992) Fractionation of mid-ocean ridge basalt (MORB). In: *Mantle Flow and Melt Generation at Mid-Ocean Ridges*, pp. 281–311. American Geophysical Union.
Grove TL, Donnelly-Nolan JM, and Housh T (1997) Magmatic processes that generated the rhyolite of Glass Mountain, Medicine Lake volcano, N. California. *Contributions to Mineralogy and Petrology* 127: 205–223. https://doi.org/10.1007/s004100050276.
Gualda GAR and Ghiorso MS (2014) Phase-equilibrium geobarometers for silicic rocks based on rhyolite-MELTS. Part 1: Principles, procedures, and evaluation of the method. *Contributions to Mineralogy and Petrology* 168: 1033. https://doi.org/10.1007/s00410-014-1033-3.
Gualda GAR, Ghiorso MS, Lemons RV, and Carley TL (2012) Rhyolite-MELTS: A modified calibration of MELTS optimized for silica-rich, fluid-bearing magmatic systems. *Journal of Petrology* 53: 875–890. https://doi.org/10.1093/petrology/egr080.
Gualda GAR, Bégué F, Pamukcu AS, and Ghiorso MS (2019a) Rhyolite-MELTS vs DERP—Newer does not make it better: A comment on 'the effect of anorthite content and water on Quartz–Feldspar cotectic compositions in the rhyolitic system and implications for geobarometry' by Wilke et al. (2017; Journal of Petrology, 58, 789–818). *Journal of Petrology* 60: 855–864. https://doi.org/10.1093/petrology/egz003.
Gualda GAR, Gravley DM, Deering CD, and Ghiorso MS (2019b) Magma extraction pressures and the architecture of volcanic plumbing systems. *Earth and Planetary Science Letters* 522: 118–124. https://doi.org/10.1016/j.epsl.2019.06.020.
Guo K, Zhai S-K, Wang X-Y, Yu Z-H, Lai Z-Q, Chen S, Song Z-J, Ma Y, Chen Z-X, Li X-H, and Zeng Z-G (2018) The dynamics of the southern Okinawa Trough magmatic system: New insights from the microanalysis of the an contents, trace element concentrations and Sr isotopic compositions of plagioclase hosted in basalts and silicic rocks. *Chemical Geology* 497: 146–161. https://doi.org/10.1016/j.chemgeo.2018.09.002.
Hagiwara Y, Kawano T, Takahata K, Torimoto J, and Yamamoto J (2021) Temperature dependence of a Raman CO_2 densimeter from 23°C to 200°C and 7.2 to 248.7 MPa: Evaluation of density underestimation by laser heating. *Journal of Raman Specroscopy* 52: 1744–1757. https://doi.org/10.1002/jrs.6188.
Halldórsson SA, Marshall EW, Caracciolo A, Matthews S, Bali E, Rasmussen MB, Ranta E, Robin JG, Guðfinnsson GH, Sigmarsson O, Maclennan J, Jackson MG, Whitehouse MJ, Jeon H, van der Meer QHA, Mibei GK, Kalliokoski MH, Repczynska MM, Rúnarsdóttir RH, Sigurðsson G, Pfeffer MA, Scott SW, Kjartansdóttir R, Kleine BI, Oppenheimer C, Aiuppa A, Ilyinskaya E, Bitetto M, Giudice G, and Stefánsson A (2022) Rapid shifting of a deep magmatic source at Fagradalsfjall volcano, Iceland. *Nature* 609: 529–534. https://doi.org/10.1038/s41586-022-04981-x.
Hammer J, Jacob S, Welsch B, Hellebrand E, and Sinton J (2016) Clinopyroxene in postshield Haleakala ankaramite: 1. Efficacy of thermobarometry. *Contributions to Mineralogy and Petrology* 171: 1–23. https://doi.org/10.1007/s00410-015-1212-x.
Hammestrom JE and Zen E-A (1986) Aluminum in hornblende: An empirical igneous geobarometer. *American Mineralogist* 71: 1297–1313.
Hansteen TH and Klugel A (2008) Fluid Inclusion thermobarometry as a tracer for magmatic processes. *Reviews in Mineralogy and Geochemistry* 69: 143–177. https://doi.org/10.2138/rmg.2008.69.5.
Hansteen TH, Klügel A, and Schmincke H-U (1998) Multi-stage magma ascent beneath the Canary Islands: Evidence from fluid inclusions. *Contributions to Mineralogy and Petrology* 132: 48–64. https://doi.org/10.1007/s004100050404.
Hanyu T, Yamamoto J, Kimoto K, Shimizu K, and Ushikubo T (2020) Determination of total CO2 in melt inclusions with shrinkage bubbles. *Chemical Geology* 557: 119855. https://doi.org/10.1016/j.chemgeo.2020.119855.
Harmon LJ, Cowlyn J, Gualda GAR, and Ghiorso MS (2018) Phase-equilibrium geobarometers for silicic rocks based on rhyolite-MELTS. Part 4: Plagioclase, orthopyroxene, clinopyroxene, glass geobarometer, and application to Mt. Ruapehu, New Zealand. *Contributions to Mineralogy and Petrology* 173: 7. https://doi.org/10.1007/s00410-017-1428-z.
Harris CR, Millman KJ, van der Walt SJ, Gommers R, Virtanen P, Cournapeau D, Wieser E, Taylor J, Berg S, Smith NJ, Kern R, Picus M, Hoyer S, van Kerkwijk MH, Brett M, Haldane A, del Río JF, Wiebe M, Peterson P, Gérard-Marchant P, Sheppard K, Reddy T, Weckesser W, Abbasi H, Gohlke C, and Oliphant TE (2020) Array programming with NumPy. *Nature* 585: 357–362. https://doi.org/10.1038/s41586-020-2649-2.
Hartley ME, Bali E, Maclennan J, Neave DA, and Halldórsson SA (2018) Melt inclusion constraints on petrogenesis of the 2014–2015 Holuhraun eruption, Iceland. *Contributions to Mineralogy and Petrology* 173: 1–23.
Hartley ME, Maclennan J, Edmonds M, and Thordarson T (2014) Reconstructing the deep CO2 degassing behaviour of large basaltic fissure eruptions. *Earth and Planetary Science Letters* 393: 120–131. https://doi.org/10.1016/j.epsl.2014.02.031.
Hauri EH, Maclennan J, McKenzie D, Gronvold K, Oskarsson N, and Shimizu N (2018) CO2 content beneath northern Iceland and the variability of mantle carbon. *Geology* 46: 55–58. https://doi.org/10.1130/G39413.1.
Helz RT (1982) Chapter 2, Experimental studies of amphibole stability: Phase relations and compositions of amphiboles produced in studies of the melting behavior of rocks. *Reviews in Mineralogy and Geochemistry* 9(1): 279–353.

Helz RT and Thornber CR (1987) Geothermometry of Kilauea Iki lava lake, Hawaii. *Bulletin of Volcanology* 49: 651–668. https://doi.org/10.1007/BF01080357.
Hernandez Nava A, Black BA, Gibson SA, Bodnar RJ, Renne PR, and Vanderkluysen L (2021) Reconciling early Deccan Traps CO_2 outgassing and pre-KPB global climate. *Proceedings. National Academy of Sciences. United States of America* 118: e2007797118. https://doi.org/10.1073/pnas.2007797118.
Hernandez PR, Woodcock A, Estrada M, and Schultz PW (2018) Undergraduate research experiences broaden diversity in the scientific workforce. *Bioscience* 68: 204–211. https://doi.org/10.1093/biosci/bix163.
Hernández-Uribe D, Spera FJ, Bohrson WA, and Heinonen JS (2022) A comparative study of two-phase equilibria modeling tools: MORB equilibrium states at variable pressure and H_2O concentrations. *American Mineralogist* 107(9): 1789–1806.
Herzberg C (2004) Partial crystallization of mid-ocean ridge basalts in the crust and mantle. *Journal of Petrology* 45: 2389–2405. https://doi.org/10.1093/petrology/egh040.
Heywood LJ, DeBari SM, Gill JB, Straub SM, Schindlbeck-Belo JC, Escobar-Burciaga RD, and Woodhead J (2020) Across-arc diversity in rhyolites from an intra-oceanic arc: Evidence from IODP Site U1437, Izu-Bonin rear arc, and surrounding area. *Geochemistry, Geophysics, Geosystems* 21. https://doi.org/10.1029/2019GC008353.
Higgins O, Sheldrake T, and Caricchi L (2022) Machine learning thermobarometry and chemometry using amphibole and clinopyroxene: A window into the roots of an arc volcano (Mount Liamuiga, Saint Kitts). *Contributions to Mineralogy and Petrology* 177: 10. https://doi.org/10.1007/s00410-021-01874-6.
Hirschmann MM, Ghiorso MS, Davis FA, Gordon SM, Mukherjee S, Grove TL, Krawczynski M, Medard E, and Till CB (2008) Library of experimental phase relations (LEPR): A database and web portal for experimental magmatic phase equilibria data: LIBRARY OF EXPERIMENTAL PHASE RELATIONS. *Geochemistry, Geophysics, Geosystems* 9: n/a-n/a. https://doi.org/10.1029/2007GC001894.
Holland T and Blundy J (1994) Non-ideal interactions in calcic amphiboles and their bearing on amphibole-plagioclase thermometry. *Contributions to Mineralogy and Petrology* 116: 433–447. https://doi.org/10.1007/BF00310910.
Holland TJB and Powell R (1990) An enlarged and updated internally consistent thermodynamic dataset with uncertainties and correlations: The system K2O–Na2O–CaO–MgO–MnO–FeO–Fe2O3–Al2O3–TiO2–SiO2–C–H2–O2. *Journal of Metamorphic Geology* 8: 89–124. https://doi.org/10.1111/j.1525-1314.1990.tb00458.x.
Holland TJB and Powell R (1998) An internally consistent thermodynamic data set for phases of petrological interest. *Journal of Metamorphic Geology* 16: 309–343. https://doi.org/10.1111/j.1525-1314.1998.00140.x.
Holland TJB and Powell R (2004) An internally consistent thermodynamic data set for phases of petrological interest. *Journal of Metamorphic Geology* 16: 309–343. https://doi.org/10.1111/j.1525-1314.1998.00140.x.
Holland TJB and Powell R (2011) An improved and extended internally consistent thermodynamic dataset for phases of petrological interest, involving a new equation of state for solids: Thermodynamic dataset for phases of petrological interest. *Journal of Metamorphic Geology* 29: 333–383. https://doi.org/10.1111/j.1525-1314.2010.00923.x.
Holland TJB, Green ECR, and Powell R (2018) Melting of peridotites through to granites: A simple thermodynamic model in the system KNCFMASHTOCr. *Journal of Petrology* 59: 881–900. https://doi.org/10.1093/petrology/egy048.
Hollister L, Grissom G, Peters EK, Stowell HH, and Slsson V (1987) Confirmation of the empirical correlation of Al in hornblende with pressure of solidification of calc-alkaline plutons. *American Mineralogist* 72(3–4): 231–239.
Huang R and Audétat A (2012) The titanium-in-quartz (TitaniQ) thermobarometer: A critical examination and re-calibration. *Geochimica et Cosmochimica Acta* 84: 75–89. https://doi.org/10.1016/j.gca.2012.01.009.
Humphreys MCS, Cooper GF, Zhang J, Loewen M, Kent AJR, Macpherson CG, and Davidson JP (2019) Unravelling the complexity of magma plumbing at Mount St. Helens: A new trace element partitioning scheme for amphibole. *Contributions to Mineralogy and Petrology* 174: 9. https://doi.org/10.1007/s00410-018-1543-5.
Hurai V (2010) Fluid inclusion geobarometry: Pressure corrections for immiscible H2O–CH4 and H2O–CO2 fluids. *Chemical Geology* 278: 201–211. https://doi.org/10.1016/j.chemgeo.2010.09.014.
Hutchison W, Mather TA, Pyle DM, Boyce AJ, Gleeson MLM, Yirgu G, Blundy JD, Ferguson DJ, Vye-Brown C, Millar IL, Sims KWW, and Finch AA (2018) The evolution of magma during continental rifting: New constraints from the isotopic and trace element signatures of silicic magmas from Ethiopian volcanoes. *Earth and Planetary Science Letters* 489: 203–218. https://doi.org/10.1016/j.epsl.2018.02.027.
Hutchison W, Pyle DM, Mather TA, Yirgu G, Biggs J, Cohen BE, Barfod DN, and Lewi E (2016) The eruptive history and magmatic evolution of Aluto volcano: New insights into silicic peralkaline volcanism in the Ethiopian rift. *Journal of Volcanology and Geothermal Research* 328: 9–33.
Iacono-Marziano G, Morizet Y, Le Trong E, and Gaillard F (2012) New experimental data and semi-empirical parameterization of H2O–CO2 solubility in mafic melts. *Geochimica et Cosmochimica Acta* 97: 1–23. https://doi.org/10.1016/j.gca.2012.08.035.
Iacovino K and Till C (2019) DensityX: A program for calculating the densities of hydrous magmatic liquids from 427-1,627 °C and up to 30 kbar. *Volcanica* 2: 1–10. https://doi.org/10.30909/vol.02.01.0110.
Iacovino K, Matthews S, Wieser PE, Moore G, and Bégué F (2021) VESIcal part I: An open-source thermodynamic model engine for mixed volatile (H2O-CO2) solubility in silicate melt. *Earth and Space Science* 8(11): e2020EA001584. https://doi.org/10.1029/2020EA001584.
Jamshidi K, Ghasemi H, Troll VR, Sadeghian M, and Dahren B (2015) Magma storage and plumbing of adakite-type post-ophiolite intrusions in the Sabzevar ophiolitic zone, northeast Iran. *Solid Earth* 6: 49–72. https://doi.org/10.5194/se-6-49-2015.
Jeffery AJ, Gertisser R, O'Driscoll B, Pacheco JM, Whitley S, Pimentel A, and Self S (2016) Temporal evolution of a post-caldera, mildly peralkaline magmatic system: Furnas volcano, São Miguel, Azores. *Contributions to Mineralogy and Petrology* 171: 42. https://doi.org/10.1007/s00410-016-1235-y.
Jeffery AJ, Gertisser R, Self S, Pimentel A, O'Driscoll B, and Pacheco JM (2017) Petrogenesis of the peralkaline ignimbrites of Terceira, Azores. *Journal of Petrology* 58: 2365–2402. https://doi.org/10.1093/petrology/egy012.
Jennings ES and Holland TJB (2015) A simple thermodynamic model for melting of peridotite in the system NCFMASOCr. *Journal of Petrology* 56: 869–892. https://doi.org/10.1093/petrology/egv020.
Jennings ES, Gibson SA, and Maclennan J (2019) Hot primary melts and mantle source for the Paraná-Etendeka flood basalt province: New constraints from Al-in-olivine thermometry. *Chemical Geology* 529: 119287. https://doi.org/10.1016/j.chemgeo.2019.119287.
Johnson ER and Cashman KV (2020) Understanding the storage conditions and fluctuating eruption style of a young monogenetic volcano: Blue Lake crater (<3 ka), High Cascades, Oregon. *Journal of Volcanology and Geothermal Research* 408: 107103. https://doi.org/10.1016/j.jvolgeores.2020.107103.
Johnson MC and Rutherford MJ (1989) Experimental calibration of the aluminum-in-hornblende geobarometer with application to Long Valley caldera (California) volcanic rocks. *Geology* 17: 837. https://doi.org/10.1130/0091-7613(1989)017<0837:ECOTAI>2.3.CO;2.
Jorgenson C, Caricchi L, Stueckelberger M, Fevola G, and Weber G (2021) *A Myriad of Melt Inclusions: A Synchrotron Microtomography Study of Melt Inclusions and Vapour Bubbles from Colli Albani (Italy) (other).* pico. https://doi.org/10.5194/egusphere-egu21-13287.
Jorgenson C, Higgins O, Petrelli M, Bégué F, and Caricchi L (2022) A machine learning-based approach to clinopyroxene thermobarometry: Model optimization and distribution for use in earth sciences. *JGR Solid Earth* 127. https://doi.org/10.1029/2021JB022904.
Kawakami Y, Yamamoto J, and Kagi H (2003) Micro-raman densimeter for CO_2 inclusions in mantle-derived minerals. *Applied Spectroscopy* 57: 1333–1339. https://doi.org/10.1366/000370203322554473.
Kawamoto T (1996) Experimental constraints on differentiation and H2O abundance of calc-alkaline magmas. *Earth and Planetary Science Letters* 144: 577–589. https://doi.org/10.1016/S0012-821X(96)00182-3.
Kawasaki T and Osanai Y (2008) Empirical thermometer of TiO_2 in quartz for ultrahigh-temperature granulites of East Antarctica. *Geological Society, London, Special Publications* 308: 419–430. https://doi.org/10.1144/SP308.21.
Kelley DF and Barton M (2008) Pressures of crystallization of Icelandic magmas. *Journal of Petrology* 49(3): 465–492. https://doi.org/10.1093/petrology/egm089.
Keppler H (1989) The influence of the fluid phase composition on the solidus temperatures in the haplogranite system NaAlSi3O8-KAlSi3O8-SiO2-H2O-CO2. *Contributions to Mineralogy and Petrology* 102: 321–327. https://doi.org/10.1007/BF00373725.

King PL and Holloway JR (2002) CO2 solubility and speciation in intermediate (andesitic) melts: The role of H2O and composition. *Geochimica et Cosmochimica Acta* 66: 1627–1640. https://doi.org/10.1016/S0016-7037(01)00872-9.

Klein BZ, Jagoutz O, Schmidt MW, and Kueter N (2023) A global assessment of the controls on the fractionation of arc magmas. *Geochemistry, Geophysics, Geosystems* 24: e2023GC010888. https://doi.org/10.1029/2023GC010888.

Klöcking M, Wyborn L, Lehnert KA, Ware B, Prent AM, Profeta L, Kohlmann F, Noble W, Bruno I, Lambart S, Ananuer H, Barber ND, Becker H, Brodbeck M, Deng H, Deng K, Elger K, De Souza Franco G, Gao Y, Ghasera KM, Hezel DC, Huang J, Kerswell B, Koch H, Lanati AW, Ter Maat G, Martínez-Villegas N, Nana Yobo L, Redaa A, Schäfer W, Swing MR, Taylor RJM, Traun MK, Whelan J, and Zhou T (2023) Community recommendations for geochemical data, services and analytical capabilities in the 21st century. *Geochimica et Cosmochimica Acta* 351: 192–205. https://doi.org/10.1016/j.gca.2023.04.024.

Klügel A (1998) Reactions between mantle xenoliths and host magma beneath La Palma (Canary Islands): Constraints on magma ascent rates and crustal reservoirs. *Contributions to Mineralogy and Petrology* 131: 237–257.

Klügel A, Day S, Schmid M, and Faria B (2020) Magma plumbing during the 2014–2015 eruption of fogo (Cape Verde Islands). *Frontiers in Earth Science* 8: 157. https://doi.org/10.3389/feart.2020.00157.

Kobayashi T, Yamamoto J, Hirajima T, Ishibashi H, Hirano N, Lai Y, Prikhod'ko VS, and Arai S (2012) Conformity and precision of CO_2 densimetry in CO_2 inclusions: Microthermometry *versus* Raman microspectroscopic densimetry: Conformity and precision of CO_2 densimetry in CO_2 inclusions. *Journal of Raman Specroscopy* 43: 1126–1133. https://doi.org/10.1002/jrs.3134.

Kohn MJ (2014) "Thermoba-Raman-try": Calibration of spectroscopic barometers and thermometers for mineral inclusions. *Earth and Planetary Science Letters* 388: 187–196. https://doi.org/10.1016/j.epsl.2013.11.054.

Kohn MJ and Spear FS (1991) Error propagation for barometers: 2. Application to rocks. *American Mineralogist* 76(1–2): 138–147.

Kohn MJ, Mazzucchelli ML, and Alvaro M (2023) Elastic thermobarometry. *Annual Review of Earth and Planetary Sciences* 51: 331–366.

Koleszar AM, Kent AJR, Wallace PJ, and Scott WE (2012) Controls on long-term low explosivity at andesitic arc volcanoes: Insights from Mount Hood, Oregon. *Journal of Volcanology and Geothermal Research* 219–220: 1–14. https://doi.org/10.1016/j.jvolgeores.2012.01.003.

Krawczynski MJ, Grove TL, and Behrens H (2012) Amphibole stability in primitive arc magmas: Effects of temperature, H2O content, and oxygen fugacity. *Contributions to Mineralogy and Petrology* 164: 317–339. https://doi.org/10.1007/s00410-012-0740-x.

Kress VC and Ghiorso MS (2004) Thermodynamic modeling of post-entrapment crystallization in igneous phases. *Journal of Volcanology and Geothermal Research* 137: 247–260. https://doi.org/10.1016/j.jvolgeores.2004.05.012.

Lamadrid HM, Moore LR, Moncada D, Rimstidt JD, Burruss RC, and Bodnar RJ (2017) Reassessment of the Raman CO2 densimeter. *Chemical Geology* 450: 210–222. https://doi.org/10.1016/j.chemgeo.2016.12.034.

Larocque J and Canil D (2010) The role of amphibole in the evolution of arc magmas and crust: The case from the Jurassic Bonanza arc section, Vancouver Island, Canada. *Contributions to Mineralogy and Petrology* 159: 475–492. https://doi.org/10.1007/s00410-009-0436-z.

Le Voyer M, Kelley KA, Cottrell E, and Hauri EH (2017) Heterogeneity in mantle carbon content from CO2-undersaturated basalts. *Nature Communications* 8: 14062. https://doi.org/10.1038/ncomms14062.

Leahy GM, Collins JA, Wolfe CJ, Laske G, and Solomon SC (2010) Underplating of the Hawaiian Swell: Evidence from teleseismic receiver functions: Underplating of the Hawaiian Swell. *Geophysical Journal International* 183: 313–329. https://doi.org/10.1111/j.1365-246X.2010.04720.x.

Lee C-TA and Anderson DL (2015) Continental crust formation at arcs, the arclogite "delamination" cycle, and one origin for fertile melting anomalies in the mantle. *Science Bulletin* 60: 1141–1156. https://doi.org/10.1007/s11434-015-0828-6.

Lerner AH, Sublett DM, Cauley C, Wallace P, and Bodnar RJ (2021a) *Magma Storage Depths and Excess CO2 Fluids From the Explosive Keanakakoi tephra (Kilauea Volcano, Hawaii) Based on Measurements of Melt and Fluid Inclusions*. AGU Fall Meeting 2021, Abstract ID: DI25B-0044.

Lerner AH, Wallace P, and Shea T (2021b) The petrologic and degassing behavior of sulfur and other magmatic volatiles from the 2018 eruption of Kīlauea, Hawai'i: Melt concentrations, magma storage depths, and magma recycling. *Bulletin of Volcanology* 83(43): 1–32.

Li X, Kind R, Yuan X, Wölbern I, and Hanka W (2004) Rejuvenation of the lithosphere by the Hawaiian plume. *Nature* 427: 827–829. https://doi.org/10.1038/nature02349.

Liang Y, Sun C, and Yao L (2013) A REE-in-two-pyroxene thermometer for mafic and ultramafic rocks. *Geochimica et Cosmochimica Acta* 102: 246–260. https://doi.org/10.1016/j.gca.2012.10.035.

Lissenberg CJ and MacLeod CJ (2016) A reactive porous flow control on mid-ocean ridge magmatic evolution. *Journal of Petrology* 57: 2195–2220. https://doi.org/10.1093/petrology/egw074.

Liu Y, Zhang Y, and Behrens H (2005) Solubility of H2O in rhyolitic melts at low pressures and a new empirical model for mixed H2O–CO2 solubility in rhyolitic melts. *Journal of Volcanology and Geothermal Research* 143: 219–235. https://doi.org/10.1016/j.jvolgeores.2004.09.019.

Lones MA (2021) How to avoid machine learning pitfalls: A guide for academic researchers. *arXiv*. 2108.02497 [cs].

Loucks RR (1996) A precise olivine-augite Mg-Fe-exchange geothermometer. *Contributions to Mineralogy and Petrology* 125: 140–150. https://doi.org/10.1007/s004100050211.

Lowenstern JB (2003) Melt inclusions come of age: Volatiles, volcanoes, and sorby's legacy. In: *Developments in Volcanology*, pp. 1–21. Elsevier. https://doi.org/10.1016/S1871-644X(03)80021-9.

Ludden JN (1978) Magmatic evolution of the basaltic shield volcanoes of Reunion Island. *Journal of Volcanology and Geothermal Research* 4: 171–198. https://doi.org/10.1016/0377-0273(78)90035-5.

Luo B, Wang Z, Song J, Qian Y, He Q, Li Y, Head JW, Moynier F, Xiao L, and Becker H (2023) The magmatic architecture and evolution of the Chang'e-5 lunar basalts. *Nature Geoscience*: 1–8.

Mackwell SJ and Kohlstedt DL (1990) Diffusion of hydrogen in olivine: Implications for water in the mantle. *Journal of Geophysical Research* 95: 5079. https://doi.org/10.1029/JB095iB04p05079.

Maclennan J (2008) Concurrent mixing and cooling of melts under Iceland. *Journal of Petrology* 49: 1931–1953. https://doi.org/10.1093/petrology/egn052.

Maclennan J (2017) Bubble formation and decrepitation control the CO2 content of olivine-hosted melt inclusions. *Geochemistry, Geophysics, Geosystems* 18: 597–616. https://doi.org/10.1002/2016GC006633.

Maclennan J, McKenzie D, Gronvöld K, and Slater L (2001) Crustal accretion under northern Iceland. *Earth and Planetary Science Letters* 191: 295–310. https://doi.org/10.1016/S0012-821X(01)00420-4.

Mandler BE, Donnelly-Nolan JM, and Grove TL (2014) Straddling the tholeiitic/calc-alkaline transition: The effects of modest amounts of water on magmatic differentiation at Newberry Volcano, Oregon. *Contributions to Mineralogy and Petrology* 168: 1066. https://doi.org/10.1007/s00410-014-1066-7.

Marxer F, Ulmer P, and Müntener O (2022) Polybaric fractional crystallisation of arc magmas: An experimental study simulating trans-crustal magmatic systems. *Contributions to Mineralogy and Petrology* 177: 3. https://doi.org/10.1007/s00410-021-01856-8.

Masotta M and Mollo S (2019) A new plagioclase-liquid hygrometer specific to trachytic systems. *Minerals* 9: 375. https://doi.org/10.3390/min9060375.

Masotta M, Mollo S, Freda C, Gaeta M, and Moore G (2013) Clinopyroxene–liquid thermometers and barometers specific to alkaline differentiated magmas. *Contributions to Mineralogy and Petrology* 166: 1545–1561. https://doi.org/10.1007/s00410-013-0927-9.

Matthews S, Shorttle O, and Maclennan J (2016) The temperature of the Icelandic mantle from olivine-spinel aluminum exchange thermometry. *Geochemistry, Geophysics, Geosystems* 17: 4725–4752. https://doi.org/10.1002/2016GC006497.

Matthews S, Wong K, Shorttle O, Edmonds M, and Maclennan J (2021) Do olivine crystallization temperatures faithfully record mantle temperature variability? *Geochemistry, Geophysics, Geosystems* 22. https://doi.org/10.1029/2020GC009157.

McNab F and Ball P (2023) meltPT: A Python package for basaltic whole-rock thermobarometric analysis with application to Hawai'i. *Volcanica* 6: 63–76. https://doi.org/10.30909/vol.06.01.6376.

Médard E and Le Pennec J-L (2022) Petrologic imaging of the magma reservoirs that feed large silicic eruptions. *Lithos* 428–429: 106812. https://doi.org/10.1016/j.lithos.2022.106812.

Melekhova E, Blundy J, Robertson R, and Humphreys MCS (2015) Experimental evidence for Polybaric differentiation of primitive arc basalt beneath St. Vincent, Lesser Antilles. *Journal of Petrology* 56: 161–192. https://doi.org/10.1093/petrology/egu074.

Menke W (1999) Crustal isostasy indicates anomalous densities beneath Iceland. *Geophysical Research Letters* 26(9): 1215–1218.

Mironov N, Portnyagin M, Botcharnikov R, Gurenko A, Hoernle K, and Holtz F (2015) Quantification of the CO2 budget and H2O–CO2 systematics in subduction-zone magmas through the experimental hydration of melt inclusions in olivine at high H2O pressure. *Earth and Planetary Science Letters* 425: 1–11. https://doi.org/10.1016/j.epsl.2015.05.043.

Molina JF, Moreno JA, Castro A, Rodríguez C, and Fershtater GB (2015) Calcic amphibole thermobarometry in metamorphic and igneous rocks: New calibrations based on plagioclase/amphibole Al-Si partitioning and amphibole/liquid Mg partitioning. *Lithos* 232: 286–305. https://doi.org/10.1016/j.lithos.2015.06.027.

Mironov NL, Tobelko DP, Smirnov SZ, Portnyagin MV, and Krasheninnikov SP (2020) Estimation of CO_2 content in the gas phase of melt inclusions using Raman spectroscopy: Case study of inclusions in olivine from the Karymsky volcano (Kamchatka). *Russian Geology and Geophysics* 61(5–6): 600–610.

Molina JF, Cambeses A, Moreno JA, Morales I, Lázaro C, Montero P, and Bea F (2021a) A cautionary note on amphibole geobarometry. In: *The 2nd International Electronic Conference on Mineral Science. Presented at the IECMS 2021*, p. 17. MDPI. https://doi.org/10.3390/iecms2021-09346.

Molina JF, Cambeses A, Moreno JA, Morales I, Montero P, and Bea F (2021b) A reassessment of the amphibole-plagioclase NaSi-CaAl exchange thermometer with applications to igneous and high-grade metamorphic rocks. *American Mineralogist* 106: 782–800. https://doi.org/10.2138/am-2021-7400.

Mollo S, Putirka K, Misiti V, Soligo M, and Scarlato P (2013) A new test for equilibrium based on clinopyroxene–melt pairs: Clues on the solidification temperatures of Etnean alkaline melts at post-eruptive conditions. *Chemical Geology* 352: 92–100. https://doi.org/10.1016/j.chemgeo.2013.05.026.

Moore G and Carmichael ISE (1998) The hydrous phase equilibria (to 3 kbar) of an andesite and basaltic andesite from western Mexico: Constraints on water content and conditions of phenocryst growth. *Contributions to Mineralogy and Petrology* 130: 304–319. https://doi.org/10.1007/s004100050367.

Moore LR, Gazel E, Tuohy R, Lloyd AS, Esposito R, Steele-MacInnis M, Hauri EH, Wallace PJ, Plank T, and Bodnar RJ (2015) Bubbles matter: An assessment of the contribution of vapor bubbles to melt inclusion volatile budgets. *American Mineralogist* 100: 806–823. https://doi.org/10.2138/am-2015-5036.

Moore LR, Mironov N, Portnyagin M, Gazel E, and Bodnar RJ (2018) Volatile contents of primitive bubble-bearing melt inclusions from Klyuchevskoy volcano, Kamchatka: Comparison of volatile contents determined by mass-balance versus experimental homogenization. *Journal of Volcanology and Geothermal Research* 358: 124–131. https://doi.org/10.1016/j.jvolgeores.2018.03.007.

Moore LR, Gazel E, and Bodnar RJ (2021) The volatile budget of Hawaiian magmatism: Constraints from melt inclusions from Haleakala volcano, Hawaii. *Journal of Volcanology and Geothermal Research* 410: 107144. https://doi.org/10.1016/j.jvolgeores.2020.107144.

Müntener O, Kelemen PB, and Grove TL (2001) The role of H2O during crystallization of primitive arc magmas under uppermost mantle conditions and genesis of igneous pyroxenites: An experimental study. *Contributions to Mineralogy and Petrology* 141: 643–658. https://doi.org/10.1007/s004100100266.

Müntener O, Kelemen PB, and Grove TL (2001) The role of H_2O during crystallization of primitive arc magmas under uppermost mantle conditions and genesis of igneous pyroxenites: An experimental study. *Contributions to Mineralogy and Petrology* 141: 643–658.

Mutch EJF, Blundy JD, Tattitch BC, Cooper FJ, and Brooker RA (2016) An experimental study of amphibole stability in low-pressure granitic magmas and a revised Al-in-hornblende geobarometer. *Contributions to Mineralogy and Petrology* 171: 85. https://doi.org/10.1007/s00410-016-1298-9.

Mutch EJF, Maclennan J, Holland TJB, and Buisman I (2019a) Millennial storage of near-Moho magma. *Science* 365: 260–264. https://doi.org/10.1126/science.aax4092.

Mutch EJF, Maclennan J, Shorttle O, Edmonds M, and Rudge JF (2019b) Rapid transcrustal magma movement under Iceland. *Nature Geoscience* 12: 569–574. https://doi.org/10.1038/s41561-019-0376-9.

Mutch EJF, Maclennan J, and Madden-Nadeau AL (2022) The dichotomous nature of Mg partitioning between plagioclase and melt: Implications for diffusion chronometry. *Geochimica et Cosmochimica Acta* 339: 173–189. https://doi.org/10.1016/j.gca.2022.10.035.

Nakatani T, Kudo T, and Suzuki T (2022) Experimental constraints on magma storage conditions of two caldera-forming eruptions at Towada volcano, Japan. *JGR Solid Earth* 127. https://doi.org/10.1029/2021JB023665.

Nandedkar RH, Ulmer P, and Müntener O (2014) Fractional crystallization of primitive, hydrous arc magmas: An experimental study at 0.7 GPa. *Contributions to Mineralogy and Petrology* 167: 1015. https://doi.org/10.1007/s00410-014-1015-5.

Neave DA and Putirka KD (2017) A new clinopyroxene-liquid barometer, and implications for magma storage pressures under Icelandic rift zones. *American Mineralogist* 102: 777–794. https://doi.org/10.2138/am-2017-5968.

Neave DA, Hartley ME, Maclennan J, Edmonds M, and Thordarson T (2017) Volatile and light lithophile elements in high-anorthite plagioclase-hosted melt inclusions from Iceland. *Geochimica et Cosmochimica Acta* 205: 100–118. https://doi.org/10.1016/j.gca.2017.02.009.

Neave DA, Bali E, Guðfinnsson GH, Halldórsson SA, Kahl M, Schmidt A-S, and Holtz F (2019) Clinopyroxene–liquid equilibria and geothermobarometry in natural and experimental tholeiites: The 2014–2015 Holuhraun Eruption, Iceland. *Journal of Petrology* 60: 1653–1680. https://doi.org/10.1093/petrology/egz042.

Newcombe ME, Fabbrizio A, Zhang Y, Ma C, Le Voyer M, Guan Y, Eiler JM, Saal AE, and Stolper EM (2014) Chemical zonation in olivine-hosted melt inclusions. *Contributions to Mineralogy and Petrology* 168: 1030. https://doi.org/10.1007/s00410-014-1030-6.

Newville M, Stensitzki T, Allen DB, Rawlik M, Ingargiola A, and Nelson A (2016) *LMFIT: Non-Linear Least-Square Minimization and Curve-Fitting for Python*. Astrophysics Source Code Library ascl-1606.

Nicoli G and Matthews S (2019) The Hebridean Igneous Province plumbing system: A phase equilibria perspective. *Lithos* 348: 105194.

Nimis P (1995) A clinopyroxene geobarometer for basaltic systems based on crystal-structure modeling. *Contributions to Mineralogy and Petrology* 121: 115–125. https://doi.org/10.1007/s004100050093.

Onuma K and Tohara T (1983) Effect of chromium on phase relations in the join forsterite-anorthite-diopside in air at 1 atm. *Contributions to Mineralogy and Petrology* 84: 174–181.

Pamukcu AS, Gualda GAR, Ghiorso MS, Miller CF, and McCracken RG (2015) Phase-equilibrium geobarometers for silicic rocks based on rhyolite-MELTS—Part 3: Application to the Peach Spring Tuff (Arizona–California–Nevada, USA). *Contributions to Mineralogy and Petrology* 169: 33. https://doi.org/10.1007/s00410-015-1122-y.

Pamukçu AS, Wright KA, Gualda GAR, and Gravley D (2020) Magma residence and eruption at the Taupo Volcanic Center (Taupo Volcanic Zone, New Zealand): Insights from rhyolite-MELTS geobarometry, diffusion chronometry, and crystal textures. *Contributions to Mineralogy and Petrology* 175: 48. https://doi.org/10.1007/s00410-020-01684-2.

Pamukçu AS, Gualda GAR, and Gravley DM (2021) Rhyolite-MELTS and the storage and extraction of large-volume crystal-poor rhyolitic melts at the Taupō Volcanic Center: A reply to Wilson et al. (2021). *Contributions to Mineralogy and Petrology* 176: 82. https://doi.org/10.1007/s00410-021-01840-2.

Papale P, Moretti R, and Barbato D (2006) The compositional dependence of the saturation surface of H2O+CO2 fluids in silicate melts. *Chemical Geology* 229: 78–95. https://doi.org/10.1016/j.chemgeo.2006.01.013.

Parat F, Streck M, Holtz F, and Almeev RR (2014) Experimental study into the petrogenesis of crystal-rich basaltic to andesitic magmas at Arenal volcano. *Contributions to Mineralogy and Petrology* 168: 1040.

Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, Killeen T, Lin Z, Gimelshein N, and Antiga L (2019) Pytorch: An imperative style, high-performance deep learning library. *Advances in Neural Information Processing Systems* 32.

Peccerillo A (2003) Relationships between Mafic and Peralkaline Silicic magmatism in continental rift settings: A petrological, geochemical and isotopic study of the Gedemsa volcano, Central Ethiopian Rift. *Journal of Petrology* 44: 2003–2032. https://doi.org/10.1093/petrology/egg068.

Pecher A (1981) Experimental decrepitation and re-equilibration of fluid inclusions in synthetic quartz. *Tectonophysics* 78: 567–583. https://doi.org/10.1016/0040-1951(81)90029-9.

Perinelli C, Mollo S, Gaeta M, De Cristofaro SP, Palladino DM, Armienti P, Scarlato P, and Putirka KD (2016) An improved clinopyroxene-based hygrometer for Etnean magmas and implications for eruption triggering mechanisms. *American Mineralogist* 101: 2774–2777. https://doi.org/10.2138/am-2016-5916.

Petrelli M, Caricchi L, and Perugini D (2020) Machine learning thermo-barometry: Application to clinopyroxene-bearing magmas. *Journal of Geophysical Research - Solid Earth* 125. https://doi.org/10.1029/2020JB020130.

Pichavant M and Macdonald R (2007) Crystallization of primitive basaltic magmas at crustal pressures and genesis of the calc-alkaline igneous suite: Experimental evidence from St Vincent, Lesser Antilles arc. *Contributions to Mineralogy and Petrology* 154: 535–558. https://doi.org/10.1007/s00410-007-0208-6.

Pineda C, Hammer J, First E, and Morata D (2021) Storage conditions of a caldera-forming volcanic eruption: Insights from the Pudahuel rhyolitic ignimbrite in central Chile (32° 10′S). *Lithos* 400–401: 106382. https://doi.org/10.1016/j.lithos.2021.106382.

Plank T, Kelley KA, Zimmer MM, Hauri EH, and Wallace PJ (2013) Why do mafic arc magmas contain ~4wt% water on average? *Earth and Planetary Science Letters* 364: 168–179. https://doi.org/10.1016/j.epsl.2012.11.044.

Powell R, Holland T, and Worley B (1998) Calculating phase diagrams involving solid solutions via non-linear equations, with examples using THERMOCALC. *Journal of Metamorphic Geology* 16: 577–588.

Prissel TC, Parman SW, and Head JW (2016) Formation of the lunar highlands Mg-suite as told by spinel. *American Mineralogist* 101: 1624–1635. https://doi.org/10.2138/am-2016-5581.

Pritchard ME, Mather TA, McNutt SR, Delgado FJ, and Reath K (2019) Thoughts on the criteria to determine the origin of volcanic unrest as magmatic or non-magmatic. *Philosophical Transactions of the Royal Society A* 377: 20180008. https://doi.org/10.1098/rsta.2018.0008.

Profeta L, Ducea MN, Chapman JB, Paterson SR, Gonzales SMH, Kirsch M, Petrescu L, and DeCelles PG (2016) Quantifying crustal thickness over time in magmatic arcs. *Scientific Reports* 5: 17786. https://doi.org/10.1038/srep17786.

Pu X, Lange RA, and Moore G (2017) A comparison of olivine-melt thermometers based on D_{Mg} and D_{Ni}: The effects of melt composition, temperature, and pressure with applications to MORBs and hydrous arc basalts. *American Mineralogist* 102: 750–765. https://doi.org/10.2138/am-2017-5879.

Pu X, Moore GM, Lange RA, Touran JP, and Gagnon JE (2021) Experimental evaluation of a new H2O-independent thermometer based on olivine-melt Ni partitioning at crustal pressure. *American Mineralogist* 106: 235–250. https://doi.org/10.2138/am-2020-7014.

Putirka K (1999) Clinopyroxene + liquid equilibria to 100 kbar and 2450 K. *Contributions to Mineralogy and Petrology* 135: 151–163. https://doi.org/10.1007/s004100050503.

Putirka KD (2005) Igneous thermometers and barometers based on plagioclase + liquid equilibria: Tests of some existing models and new calibrations. *American Mineralogist* 90: 336–346. https://doi.org/10.2138/am.2005.1449.

Putirka KD (2008) Thermometers and barometers for volcanic systems. *Reviews in Mineralogy and Geochemistry* 69: 61–120. https://doi.org/10.2138/rmg.2008.69.3.

Putirka K (2016) Amphibole thermometers and barometers for igneous systems and some implications for eruption mechanisms of felsic magmas at arc volcanoes. *American Mineralogist* 101: 841–858. https://doi.org/10.2138/am-2016-5506.

Putirka K (2017) Geothermometry and geobarometry. In: White WM (ed.) *Encyclopedia of Geochemistry: A Comprehensive Reference Source on the Chemistry of the Earth*, pp. 597–614. Cham, Switzerland: Springer International Publishing.

Putirka K, Johnson M, Kinzler R, Longhi J, and Walker D (1996) Thermobarometry of mafic igneous rocks based on clinopyroxene-liquid equilibria, 0–30 kbar. *Contributions to Mineralogy and Petrology* 123: 92–108. https://doi.org/10.1007/s004100050145.

Putirka KD, Mikaelian H, Ryerson F, and Shaw H (2003) New clinopyroxene-liquid thermobarometers for mafic, evolved, and volatile-bearing lava compositions, with applications to lavas from Tibet and the Snake River Plain, Idaho. *American Mineralogist* 88: 1542–1554. https://doi.org/10.2138/am-2003-1017.

Qin Z, Lu F, and Anderson A (1992) Diffusive reequilibration of melt and fluid inclusions. *American Mineralogist* 77(506): 565–576.

Quinn E (2014) *Experimental Determination of Pre-Eruptive Storage Conditions and Continuous Decompression of Rhyodacite Magma Erupted from Chaos Crags, Lassen Volcanic Center, California*. MSci thesis, Hombold State University.

Rader EL and Larsen JF (2013) Experimental phase relations of a low MgO Aleutian basaltic andesite at XH2O = 0.7–1. *Contributions to Mineralogy and Petrology* 166: 1593–1611. https://doi.org/10.1007/s00410-013-0944-8.

Ranero CR, Torne M, and Banda E (1995) Gravity and multichannel seismic reflection constraints on the lithospheric structure of the Canary Swell. *Marine Geophysical Researches* 17: 519–534. https://doi.org/10.1007/BF01204342.

Rasmussen DJ, Kyle PR, Wallace PJ, Sims KWW, Gaetani GA, and Phillips EH (2017) Understanding degassing and transport of CO2-rich alkalic magmas at Ross Island, Antarctica using olivine-hosted melt inclusions. *Journal of Petrology*. https://doi.org/10.1093/petrology/egx036.

Rasmussen DJ, Plank TA, Wallace PJ, Newcombe ME, and Lowenstern JB (2020) Vapor-bubble growth in olivine-hosted melt inclusions. *American Mineralogist* 105: 1898–1919. https://doi.org/10.2138/am-2020-7377.

Rasmussen DJ, Plank TA, Roman DC, and Zimmer MM (2022) Magmatic water content controls the pre-eruptive depth of arc magmas. *Science* 375: 1169–1172. https://doi.org/10.1126/science.abm5174.

Reubi O and Blundy J (2009) A dearth of intermediate melts at subduction zone volcanoes and the petrogenesis of arc andesites. *Nature* 461: 1269–1273. https://doi.org/10.1038/nature08510.

Richard A, Morlot C, Créon L, Beaudoin N, Balistky VS, Pentelei S, Dyja-Person V, Giuliani G, Pignatelli I, Legros H, and Sterpenich J (2019) Advances in 3D imaging and volumetric reconstruction of fluid and melt inclusions by high resolution X-ray computed tomography. *Chemical Geology* 508: 3–14.

Ridolfi F (2021) Amp-TB2: An updated model for calcic amphibole thermobarometry. *Minerals* 11: 324. https://doi.org/10.3390/min11030324.

Ridolfi F and Renzulli A (2012) Calcic amphiboles in calc-alkaline and alkaline magmas: Thermobarometric and chemometric empirical equations valid up to 1,130°C and 2.2 GPa. *Contributions to Mineralogy and Petrology* 163: 877–895. https://doi.org/10.1007/s00410-011-0704-6.

Ridolfi F, Renzulli A, and Puerini M (2010) Stability and chemical equilibrium of amphibole in calc-alkaline magmas: An overview, new thermobarometric formulations and application to subduction-related volcanoes. *Contributions to Mineralogy and Petrology* 160: 45–66. https://doi.org/10.1007/s00410-009-0465-7.

Riel N, Kaus BJP, Green ECR, and Berlie N (2022) MAGEMin, an efficient Gibbs energy minimizer: Application to igneous systems. *Geochemistry, Geophysics, Geosystems* 23. https://doi.org/10.1029/2022GC010427.

Riker JM (2005) *The 1859 Eruption of Mauna Loa Volcano, Hawai'i. Controls on the Development of Long Lava Channels*. University of Oregon. PhD thesis.

Robidoux P, Frezzotti ML, Hauri EH, and Aiuppa A (2018) Shrinkage bubbles: The C–O–H–S magmatic fluid system at San Cristóbal volcano. *Journal of Petrology* 59: 2093–2122. https://doi.org/10.1093/petrology/egy092.

Roedder E (1984) *Fluid Inclusions, Reviews in Mineralogy*. Mineralogical Society of America.

Ronga F, Lustrino M, Marzoli A, and Melluso L (2010) Petrogenesis of a basalt-comendite-pantellerite rock suite: The Boseti Volcanic Complex (Main Ethiopian Rift). *Mineralogy and Petrology* 98: 227–243. https://doi.org/10.1007/s00710-009-0064-3.

Rose-Koga EF, Bouvier A-S, Gaetani GA, Wallace PJ, Allison CM, Andrys JA, Angeles de la Torre CA, Barth A, Bodnar RJ, Bracco Gartner AJJ, Butters D, Castillejo A, Chilson-Parks B, Choudhary BR, Cluzel N, Cole M, Cottrell E, Daly A, Danyushevsky LV, DeVitre CL, Drignon MJ, France L, Gaborieau M, Garcia MO, Gatti E, Genske FS, Hartley ME, Hughes EC, Iveson AA, Johnson ER, Jones M, Kagoshima T, Katzir Y, Kawaguchi M, Kawamoto T, Kelley KA, Koornneef JM, Kurz MD, Laubier M, Layne GD, Lerner A, Lin K-Y, Liu P-P, Lorenzo-Merino A, Luciani N, Magalhães N, Marschall HR, Michael PJ, Monteleone BD, Moore LR, Moussallam Y, Muth M, Myers ML, Narváez DF, Navon O, Newcombe ME, Nichols ARL, Nielsen RL, Pamukcu A, Plank T, Rasmussen DJ, Roberge J, Schiavi F, Schwartz D, Shimizu K, Shimizu K, Shimizu N, Thomas JB, Thompson GT, Tucker JM, Ustunisik G, Waelkens C, Zhang Y, and Zhou T (2021) Silicate melt inclusions in the new millennium: A review of recommended practices for preparation, analysis, and data presentation. *Chemical Geology* 570: 120145. https://doi.org/10.1016/j.chemgeo.2021.120145.

Rosso KM and Bodnar RJ (1995) Microthermometric and Raman spectroscopic detection limits of C02 in fluid inclusions and the Raman spectroscopic characterization of C02. *Geochimica et Cosmochimica Acta* 59: 3961–3975. https://doi.org/10.1016/0016-7037(95)94441-H.

Ruscitto DM, Wallace PJ, Johnson ER, Kent AJR, and Bindeman IN (2010) Volatile contents of mafic magmas from cinder cones in the Central Oregon High Cascades: Implications for magma formation and mantle conditions in a hot arc. *Earth and Planetary Science Letters* 298: 153–161. https://doi.org/10.1016/j.epsl.2010.07.037.

Ruscitto DM, Wallace PJ, and Kent AJR (2011) Revisiting the compositions and volatile contents of olivine-hosted melt inclusions from the Mount Shasta region: Implications for the formation of high-Mg andesites. *Contributions to Mineralogy and Petrology* 162: 109–132. https://doi.org/10.1007/s00410-010-0587-y.

Ruth DCS and Costa F (2021) A petrological and conceptual model of Mayon volcano (Philippines) as an example of an open-vent volcano. *Bulletin of Volcanology* 83: 62. https://doi.org/10.1007/s00445-021-01486-9.

Rutherford MJ and Devine JD (1988) The May 18, 1980, eruption of Mount St. Helens: 3. Stability and chemistry of amphibole in the magma chamber. *Journal of Geophysical Research* 93: 11949. https://doi.org/10.1029/JB093iB10p11949.

Rutherford MJ, Sigurdsson H, Carey S, and Davis A (1985) The May 18, 1980, eruption of Mount St. Helens: 1. Melt composition and experimental phase equilibria. *Journal of Geophysical Research* 90: 2929. https://doi.org/10.1029/JB090iB04p02929.

Sack RO and Ghiorso MS (1989) Importance of considerations of mixing properties in establishing an internally consistent thermodynamic database: Thermochemistry of minerals in the system Mg2SiO4-Fe2SiO4-SiO2. *Contributions to Mineralogy and Petrology* 102: 41–68. https://doi.org/10.1007/BF01160190.

Sack RO and Ghiorso MS (1991) An internally consistent model for the thermodynamic properties of Fe?Mg-titanomagnetite-aluminate spinels. *Contributions to Mineralogy and Petrology* 106: 474–505. https://doi.org/10.1007/BF00321989.

Sack RO and Ghiorso MS (1994) Thermodynamics of multicomponent pyroxenes: I. Formulation of a general model. *Contributions to Mineralogy and Petrology* 116: 277–286. https://doi.org/10.1007/BF00306497.

Sanfilippo A, MacLeod CJ, Tribuzio R, Lissenberg CJ, and Zanetti A (2020) Early-stage melt-rock reaction in a cooling crystal mush beneath a slow-spreading mid-ocean ridge (IODP Hole U1473A, Atlantis Bank, Southwest Indian Ridge). *Frontiers in Earth Science* 8: 579138. https://doi.org/10.3389/feart.2020.579138.

Sas M, DeBari S, Clynne M, and Rusk B (2017) *Using mineral Geochemistry to Decipher Slab, Mantle, and Crustal Input in the Generation of High-Mg Andesites and Basaltic Andesites From the Northern Cascade Arc*. msam. https://doi.org/10.2138/am-2017-5756.

Scaillet B and Macdonald R (2001) Phase relations of peralkaline silicic magmas and petrogenetic implications. *Journal of Petrology* 42: 825–845. https://doi.org/10.1093/petrology/42.4.825.

Schiavi F, Bolfan-Casanova N, Buso R, Laumonier M, Laporte D, Medjoubi K, Venugopal S, Gómez-Ulla A, Cluzel N, and Hardiagon M (2020) Quantifying magmatic volatiles by Raman microtomography of glass inclusion-hosted bubbles. *Geochemical Perspectives Letters* 16: 17–24. https://doi.org/10.7185/geochemlet.2038.

Schmidt MW (1992) Amphibole composition in tonalite as a function of pressure: An experimental calibration of the Al-in-hornblende barometer. *Contributions to Mineralogy and Petrology* 110: 304–310. https://doi.org/10.1007/BF00310745.

Scruggs MA and Putirka KD (2018) Eruption triggering by partial crystallization of mafic enclaves at Chaos Crags, Lassen Volcanic Center, California. *American Mineralogist* 103: 1575–1590. https://doi.org/10.2138/am-2018-6058.

Shamloo HI and Till CB (2019) Decadal transition from quiescence to supereruption: Petrologic investigation of the Lava Creek Tuff, Yellowstone Caldera, WY. *Contributions to Mineralogy and Petrology* 174: 32. https://doi.org/10.1007/s00410-019-1570-x.

Shedden K (2008) Gene expression–based survival prediction in lung adenocarcinoma: A multi-site, blinded validation study. *Nature Medicine* 14: 822–827. https://doi.org/10.1038/nm.1790.

Sheehan F and Barclay J (2016) Staged storage and magma convection at Ambrym volcano, Vanuatu. *Journal of Volcanology and Geothermal Research* 322: 144–157. https://doi.org/10.1016/j.jvolgeores.2016.02.024.

Shi P (1992) Basalt evolution at low pressure: Implications from an experimental study in the system CaO-FeO-MgO-Al2O3-SiO2. *Contributions to Mineralogy and Petrology* 110: 139–153. https://doi.org/10.1007/BF00310735.

Shi P (1993) Low-pressure phase relationships in the system Na2O- -CaO- -FeO- -MgO- -Al2O3- -SiO2 at 1100 C, with implications for the differentiation of basaltic magmas. *Journal of Petrology* 34: 743–762. https://doi.org/10.1093/petrology/34.4.743.

Shinohara H (2013) Volatile flux from subduction zone volcanoes: Insights from a detailed evaluation of the fluxes from volcanoes in Japan. *Journal of Volcanology and Geothermal Research* 268: 46–63. https://doi.org/10.1016/j.jvolgeores.2013.10.007.

Shishkina TA, Botcharnikov RE, Holtz F, Almeev RR, Jazwa AM, and Jakubiak AA (2014) Compositional and pressure effects on the solubility of H2O and CO2 in mafic melts. *Chemical Geology* 388: 112–129. https://doi.org/10.1016/j.chemgeo.2014.09.001.

Sides, Edmonds M, Maclennan J, Houghton BF, Swanson DA, and Steele-MacInnis MJ (2014a) Magma mixing and high fountaining during the 1959 Kīlauea Iki eruption, Hawai'i. *Earth and Planetary Science Letters* 400: 102–112. https://doi.org/10.1016/j.epsl.2014.05.024.

Sides, Edmonds M, Maclennan J, Swanson DA, and Houghton BF (2014b) Eruption style at Kīlauea Volcano in Hawai'i linked to primary melt composition. *Nature Geoscience* 7: 464–469. https://doi.org/10.1038/ngeo2140.

Siegburg M, Klügel A, Rocholl A, and Bach W (2018) Magma plumbing and hybrid magma formation at an active back-arc basin volcano: North Su, eastern Manus basin. *Journal of Volcanology and Geothermal Research* 362: 1–16. https://doi.org/10.1016/j.jvolgeores.2018.07.001.

Sisson TW and Grove TL (1993) Temperatures and H2O contents of low-MgO high-alumina basalts. *Contributions to Mineralogy and Petrology* 113: 167–184. https://doi.org/10.1007/BF00283226.

Sisson TW, Ratajeski K, Hankins WB, and Glazner AF (2005) Voluminous granitic magmas from common basaltic sources. *Contributions to Mineralogy and Petrology* 148: 635–661. https://doi.org/10.1007/s00410-004-0632-9.

Skirius C, Peterson J, and Anderson AT Jr (1990) Homogenizing rhyolitic glass inclusions from the Bishop Tuff. *American Mineralogist* 75: 1381–1398.

Smith EM, Krebs MY, Genzel PT, and Brenker FE (2022) Raman identification of inclusions in diamond. *Reviews in Mineralogy and Geochemistry* 88(1): 451–473.

Solaro C, Martel C, Champallier R, Boudon G, Balcone-Boissard H, and Pichavant M (2019) Petrological and experimental constraints on magma storage for large pumiceous eruptions in Dominica island (Lesser Antilles). *Bulletin of Volcanology* 81: 55. https://doi.org/10.1007/s00445-019-1313-x.

Sorby HC (1858) On the microscopical, structure of crystals, indicating the origin of minerals and rocks. *Quarterly Journal of the Geological Society* 14: 453–500. https://doi.org/10.1144/GSL.JGS.1858.014.01-02.44.

Span R and Wagner W (1996) A new equation of state for carbon dioxide covering the fluid region from the triple-point temperature to 1100 K at pressures up to 800 MPa. *Journal of Physical and Chemical Reference Data* 25: 1509–1596. https://doi.org/10.1063/1.555991.

Spandler C and O'Neill HSC (2010) Diffusion and partition coefficients of minor and trace elements in San Carlos olivine at 1,300 C with some geochemical implications. *Contributions to Mineralogy and Petrology* 159: 791–818.

Spencer KJ and Lindsley DH (1981) A solution model for coexisting iron–titanium oxides. *American Mineralogist* 66: 1189–1201.

Steele-Macinnis M, Esposito R, and Bodnar RJ (2011) Thermodynamic model for the effect of post-entrapment crystallization on the H2O-CO2 systematics of vapor-saturated, silicate melt inclusions. *Journal of Petrology* 52: 2461–2482. https://doi.org/10.1093/petrology/egr052.

Sterner SM and Pitzer KS (1994) An equation of state for carbon dioxide valid from zero to extreme pressures. *Contributions to Mineralogy and Petrology* 117: 362–374. https://doi.org/10.1007/BF00307271.

Stock MJ, Bagnardi M, Neave DA, Maclennan J, Bernard B, Buisman I, Gleeson MLM, and Geist D (2018) Integrated petrological and geophysical constraints on magma system architecture in the western Galápagos Archipelago: Insights from Wolf Volcano. *Geochemistry, Geophysics, Geosystems* 19: 4722–4743. https://doi.org/10.1029/2018GC007936.

Sugawara T (2001) Ferric iron partitioning between plagioclase and silicate liquid: Thermodynamics and petrological applications. *Contributions to Mineralogy and Petrology* 141: 659–686. https://doi.org/10.1007/s004100100267.

Sun C and Liang Y (2012) Distribution of REE between clinopyroxene and basaltic melt along a mantle adiabat: Effects of major element composition, water, and temperature. *Contributions to Mineralogy and Petrology* 163: 807–823. https://doi.org/10.1007/s00410-011-0700-x.

Sun C and Lissenberg CJ (2018) Formation of fast-spreading lower oceanic crust as revealed by a new Mg–REE coupled geospeedometer. *Earth and Planetary Science Letters* 487: 165–178. https://doi.org/10.1016/j.epsl.2018.01.032.

Tamblyn R, Hand M, Morrissey L, Zack T, Phillips G, and Och D (2020) Resubduction of lawsonite eclogite within a serpentinite-filled subduction channel. *Contributions to Mineralogy and Petrology* 175: 74. https://doi.org/10.1007/s00410-020-01712-1.

Taracsák Z, Hartley ME, Burgess R, Edmonds M, Iddon F, and Longpré M-A (2019) High fluxes of deep volatiles from ocean island volcanoes: Insights from El Hierro, Canary Islands. *Geochimica et Cosmochimica Acta* 258: 19–36. https://doi.org/10.1016/j.gca.2019.05.020.

Taura H, Yurimoto H, Kurita K, and Sueno S (1998) Pressure dependence on partition coefficients for trace elements between olivine and the coexisting melts. *Physics and Chemistry of Minerals* 25: 469–484. https://doi.org/10.1007/s002690050138.

Teplow W, Marsh B, and Hulen (2009) Dacite melt at the Puna Geothermal Venture Wellfield, Big Island of Hawaii. *GRC Transactions* 33.

Thermoengine Code Contributors (2022) *ThermoEngine: Software for Model Building and Computational Thermodynamics Supporting Applications in the Earth Sciences*. https://doi.org/10.5281/ZENODO.6527840.

Thomas JB and Bruce Watson E (2012) Application of the Ti-in-quartz thermobarometer to rutile-free systems. Reply to: A comment on: 'TitaniQ under pressure: The effect of pressure and temperature on the solubility of Ti in quartz' by Thomas et al. *Contributions to Mineralogy and Petrology* 164: 369–374. https://doi.org/10.1007/s00410-012-0761-5.

Thomas WM and Ernst' WG (1990) *The Aluminum Content of Hornblende in Calc-Alkaline Granitic Rocks: A Mineralogic Barometer Calibrated Experimentally to 12 kbars*, Special Publication No. 2, pp. 59–63. The Geochemical Society.

Thomas JB, Bruce Watson E, Spear FS, Shemella PT, Nayak SK, and Lanzirotti A (2010) TitaniQ under pressure: The effect of pressure and temperature on the solubility of Ti in quartz. *Contributions to Mineralogy and Petrology* 160: 743–759. https://doi.org/10.1007/s00410-010-0505-3.

Tommasini S, Bindi L, Savia L, Mangler MF, Orlando A, and Petrone CM (2022) Critical assessment of pressure estimates in volcanic plumbing systems: The case study of Popocatépetl volcano, Mexico. *Lithos* 408–409: 106540. https://doi.org/10.1016/j.lithos.2021.106540.

Towbin WH, Plank T, Klein E, and Hauri E (2023) Measuring H2O concentrations in olivine by secondary ion mass spectrometry: Challenges and paths forward. *American Mineralogist: Journal of Earth and Planetary Materials* 108: 928–940. https://doi.org/10.2138/am-2022-8247.

Trela J, Gazel E, Sobolev AV, Moore L, Bizimis M, Jicha B, and Batanova VG (2017) The hottest lavas of the Phanerozoic and the survival of deep Archaean reservoirs. *Nature Geoscience* 10: 451–456. https://doi.org/10.1038/ngeo2954.

Tucker JM, Hauri EH, Pietruszka AJ, Garcia MO, Marske JP, and Trusdell FA (2019) A high carbon content of the Hawaiian mantle from olivine-hosted melt inclusions. *Geochimica et Cosmochimica Acta* 254: 156–172. https://doi.org/10.1016/j.gca.2019.04.001.

Tuohy RM, Wallace PJ, Loewen MW, Swanson DA, and Kent AJR (2016) Magma transport and olivine crystallization depths in Kīlauea's east rift zone inferred from experimentally rehomogenized melt inclusions. *Geochimica et Cosmochimica Acta* 185: 232–250. https://doi.org/10.1016/j.gca.2016.04.020.

Ubide T, Caulfield J, Brandt C, Bussweiler Y, Mollo S, Di Stefano F, Nazzari M, and Scarlato P (2019a) Deep magma storage revealed by multi-method elemental mapping of Clinopyroxene Megacrysts at Stromboli Volcano. *Frontiers in Earth Science* 7: 239. https://doi.org/10.3389/feart.2019.00239.

Ubide T, Mollo S, Zhao J, Nazzari M, and Scarlato P (2019b) Sector-zoned clinopyroxene as a recorder of magma history, eruption triggers, and ascent rates. *Geochimica et Cosmochimica Acta* 251: 265–283. https://doi.org/10.1016/j.gca.2019.02.021.

Ubide T, Larrea P, Becerril L, and Galé C (2022) Volcanic plumbing filters on ocean-island basalt geochemistry. *Geology* 50: 26–31. https://doi.org/10.1130/G49224.1.

Van Den Kerkhof AM (1990) Isochoric phase diagrams in the systems CO2CH4 and CO2N2: Application to fluid inclusions. *Geochimica et Cosmochimica Acta* 54: 621–629. https://doi.org/10.1016/0016-7037(90)90358-R.

van Gerve T, Neave DA, Wieser P, Lamadrid HM, Hulsbosch N, and Namur O (2023, July) Crystallisation driven deep volatile degassing in ocean island volcanoes: Integrating 3D imaging with chemical microanalysis of olivine-hosted melt inclusions from Pico (Azores). In: *Goldschmidt 2023 Conference*. GOLDSCHMIDT.

Venezky DY and Rutherford MJ (1999) Petrology and Fe–Ti oxide reequilibration of the 1991 Mount Unzen mixed magma. *Journal of Volcanology and Geothermal Research* 89: 213–230. https://doi.org/10.1016/S0377-0273(98)00133-4.

Venugopal S, Moune S, Williams-Jones G, Druitt T, Vigouroux N, Wilson A, and Russell JK (2020a) Two distinct mantle sources beneath the Garibaldi Volcanic Belt: Insight from olivine-hosted melt inclusions. *Chemical Geology* 532: 119346.

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

Villiger S (2004) The liquid line of descent of anhydrous, mantle-derived, tholeiitic liquids by fractional and equilibrium crystallization—An Experimental Study at 1{middle dot}0 GPa. *Journal of Petrology* 45: 2369–2388. https://doi.org/10.1093/petrology/egh042.

Villiger S, Müntener O, and Ulmer P (2007) Crystallization pressures of mid-ocean ridge basalts derived from major element variations of glasses from equilibrium and fractional crystallization experiments. *Journal of Geophysical Research* 112: B01202. https://doi.org/10.1029/2006JB004342.

Vogt JHL (1931) On the terms eutectic, cotectic, peritectic, anchi-eutectic, anchi-cotectic, etc., and their importance in petrogenesis. *The Journal of Geology* 39: 401–431. https://doi.org/10.1086/623862.

Voigt M, Coogan LA, and von der Handt A (2017) Experimental investigation of the stability of clinopyroxene in mid-ocean ridge basalts: The role of Cr and Ca/Al. *Lithos* 274–275: 240–253. https://doi.org/10.1016/j.lithos.2017.01.003.

Voigt A, Cassidy M, Castro JM, Pyle DM, Mather TA, Helo C, Abdurrachman M, and Kurniawan IA (2022) Experimental investigation of trachydacite magma storage prior to the 1257 eruption of Mt samalas. *Journal of Petrology* 63: egac066. https://doi.org/10.1093/petrology/egac066.

Wade JA, Plank T, Hauri EH, Kelley KA, Roggensack K, and Zimmer M (2008) Prediction of magmatic water contents via measurement of H2O in clinopyroxene phenocrysts. *Geology* 36: 799. https://doi.org/10.1130/G24964A.1.

Walker BA, Klemetti EW, Grunder AL, Dilles JH, Tepley FJ, and Giles D (2013) Crystal reaming during the assembly, maturation, and waning of an eleven-million-year crustal magma cycle: Thermobarometry of the Aucanquilcha Volcanic Cluster. *Contributions to Mineralogy and Petrology* 165: 663–682. https://doi.org/10.1007/s00410-012-0829-2.

Wallace PJ (2005) Volatiles in subduction zone magmas: Concentrations and fluxes based on melt inclusion and volcanic gas data. *Journal of Volcanology and Geothermal Research* 140: 217–240. https://doi.org/10.1016/j.jvolgeores.2004.07.023.

Wallace PJ, Anderson AT, and Davis AM (1999) Gradients in H_2O, CO_2, and exsolved gas in a large-volume silicic magma system: Interpreting the record preserved in melt inclusions from the Bishop Tuff. *Journal of Geophysical Research* 104: 20097–20122. https://doi.org/10.1029/1999JB900207.

Wallace PJ, Kamenetsky VS, and Cervantes P (2015) Melt inclusion CO2 contents, pressures of olivine crystallization, and the problem of shrinkage bubbles. *American Mineralogist* 100: 787–794. https://doi.org/10.2138/am-2015-5029.

Wallace PJ, Plank T, Bodnar RJ, Gaetani GA, and Shea T (2021) Olivine-hosted melt inclusions: A microscopic perspective on a complex magmatic world. *Annual Review of Earth and Planetary Sciences* 49: 465–494.

Walowski KJ, Wallace PJ, Clynne MA, Rasmussen DJ, and Weis D (2016) Slab melting and magma formation beneath the southern Cascade arc. *Earth and Planetary Science Letters* 446: 100–112. https://doi.org/10.1016/j.epsl.2016.03.044.

Walowski KJ, Wallace PJ, Cashman KV, Marks JK, Clynne MA, and Ruprecht P (2019) Understanding melt evolution and eruption dynamics of the 1666 C.E. eruption of Cinder Cone, Lassen Volcanic National Park, California: Insights from olivine-hosted melt inclusions. *Journal of Volcanology and Geothermal Research* 387: 106665. https://doi.org/10.1016/j.jvolgeores.2019.106665.

Wan Z, Coogan LA, and Canil D (2008) Experimental calibration of aluminum partitioning between olivine and spinel as a geothermometer. *American Mineralogist* 93(7): 1142–1147. https://doi.org/10.2138/am.2008.2758.

Wanamaker BJ and Evans B (1989) Mechanical re-equilibration of fluid inclusions in San Carlos olivine by power-law creep. *Contributions to Mineralogy and Petrology* 102: 102–111. https://doi.org/10.1007/BF01160194.

Wanamaker BJ, Wong T-F, and Evans B (1990) Decrepitation and crack healing of fluid inclusions in San Carlos olivine. *Journal of Geophysical Research* 95: 15623. https://doi.org/10.1029/JB095iB10p15623.

Wang X, Chou I-M, Hu W, Burruss RC, Sun Q, and Song Y (2011) Raman spectroscopic measurements of CO2 density: Experimental calibration with high-pressure optical cell (HPOC) and fused silica capillary capsule (FSCC) with application to fluid inclusion observations. *Geochimica et Cosmochimica Acta* 75: 4080–4093. https://doi.org/10.1016/j.gca.2011.04.028.

Wang W, Caumon M-C, Tarantola A, Pironon J, Lu W, and Huang Y (2019) Raman spectroscopic densimeter for pure CO2 and CO2-H2O-NaCl fluid systems over a wide P-T range up to 360 °C and 50 MPa. *Chemical Geology* 528: 119281. https://doi.org/10.1016/j.chemgeo.2019.119281.

Wang X, Hou T, Wang M, Zhang C, Zhang Z, Pan R, Marxer F, and Zhang H (2021) A new clinopyroxene thermobarometer for mafic to intermediate magmatic systems. *European Journal of Mineralogy* 33: 621–637. https://doi.org/10.5194/ejm-33-621-2021.

Wark DA and Watson EB (2006) TitaniQ: A titanium-in-quartz geothermometer. *Contributions to Mineralogy and Petrology* 152: 743–754. https://doi.org/10.1007/s00410-006-0132-1.

Waters LE and Lange RA (2015) An updated calibration of the plagioclase-liquid hygrometer-thermometer applicable to basalts through rhyolites. *American Mineralogist* 100: 2172–2184. https://doi.org/10.2138/am-2015-5232.

Weber G and Blundy J (2023) A machine learning-based thermometer, barometer and hygrometer for magmatic liquids (preprint). *Physical Sciences and Mathematics*. https://doi.org/10.31223/X5NW9P.

Weber G and Castro JM (2017) Phase petrology reveals shallow magma storage prior to large explosive silicic eruptions at Hekla volcano, Iceland. *Earth and Planetary Science Letters* 466: 168–180. https://doi.org/10.1016/j.epsl.2017.03.015.

White RS, McKenzie D, and O'Nions RK (1992) Oceanic crustal thickness from seismic measurements and rare earth element inversions. *Journal of Geophysical Research* 97: 19683. https://doi.org/10.1029/92JB01749.

Wieser P and DeVitre C (2023) DiadFit: An open-SourcePython3 tool for peak fitting of Raman data from silicate melts and CO2 fluids (preprint). *Earth Science*. https://doi.org/10.31223/X5CQ1F.

Wieser PE, Edmonds M, Maclennan J, Jenner FE, and Kunz BE (2019) Crystal scavenging from mush piles recorded by melt inclusions. *Nature Communications* 10: 5797. https://doi.org/10.1038/s41467-019-13518-2.

Wieser PE, Jenner F, Edmonds M, Maclennan J, and Kunz BE (2020) Chalcophile elements track the fate of sulfur at Kīlauea Volcano, Hawai'i. *Geochimica et Cosmochimica Acta*. https://doi.org/10.1016/j.gca.2020.05.018. S0016703720303239.

Wieser PE, Lamadrid H, Maclennan J, Edmonds M, Matthews S, Iacovino K, Jenner FE, Gansecki C, Trusdell F, Lee RL, and Ilyinskaya E (2021) Reconstructing magma storage depths for the 2018 K⊠lauean eruption from melt inclusion CO_2 contents: The importance of vapor bubbles. *Geochemistry, Geophysics, Geosystems* 22. https://doi.org/10.1029/2020GC009364.

Wieser PE, Edmonds M, Gansecki C, Maclennan J, Jenner FE, Kunz B, Antoshechkina P, Trusdell F, Lee RL, and Eimf (2022a) Explosive activity on Kīlauea's Lower East Rift Zone fueled by a Volatile-Rich, Dacitic Melt. *Geochemistry, Geophysics, Geosystems* 23. https://doi.org/10.1029/2021GC010046.

Wieser PE, Iacovino K, Matthews S, Moore G, and Allison CM (2022b) VESIcal: 2. A critical approach to volatile solubility modeling using an open-source Python3 engine. *Earth and Space Science* 9. https://doi.org/10.1029/2021EA001932.

Wieser PE, Petrelli M, Lubbers J, Wieser E, Ozaydin S, Kent A, and Till C (2022c) Thermobar: An open-source Python3 tool for thermobarometry and hygrometry. *Volcanica* 5: 349–384. https://doi.org/10.30909/vol.05.02.349384.

Wieser P, Kent A, Till C, and Abers G (2023a) Geophysical and geochemical constraints on magma storage depths along the cascade arc: Knowns and unknowns (preprint). *Earth Science*. https://doi.org/10.31223/X5KX00.

Wieser PE, Kent AJ, and Till CB (2023b) Barometers behaving badly II: A critical evaluation of Cpx-only and Cpx-Liq thermobarometry in variably-hydrous arc magmas. *Journal of Petrology* 64(8). egad050.

Wieser P, Till C, Kent A, and Gleeson M (2023c) Comment on 'The magmatic architecture and evolution of the Chang'e-5 lunar basalts' Penny E. Wieser1, Christy Till2, Adam Kent3, Matthew Gleeson1 (preprint). *Earth Science*. https://doi.org/10.31223/X5MM3B.

Wieser PE, Kent AJ, Till CB, Donovan J, Neave DA, Blatter DL, and Krawczynski MJ (2023d) Barometers behaving Badly I: Assessing the influence of analytical and experimental uncertainty on clinopyroxene thermobarometry calculations at crustal conditions. *Journal of Petrology* 64(2). egac126.

Wilke S, Holtz F, Neave DA, and Almeev R (2017) The effect of anorthite content and water on Quartz–Feldspar cotectic compositions in the rhyolitic system and implications for geobarometry. *Journal of Petrology* 58: 789–818. https://doi.org/10.1093/petrology/egx034.

Wilke S, Holtz F, Li X, Neave DA, and Almeev RR (2019) Rhyolite-MELTS vs DERP—Reply to Comment by Gualda et al. on 'The effect of anorthite content and water on Quartz–Feldspar cotectic compositions in the rhyolitic system and implications for geobarometry' by Wilke et al. (2017), Journal of Petrology, 58, No. 4, 789–818. *Journal of Petrology* 60: 865–870. https://doi.org/10.1093/petrology/egz002.

Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten J-W, da Silva Santos LB, Bourne PE, Bouwman J, Brookes AJ, Clark T, Crosas M, Dillo I, Dumon O, Edmunds S, Evelo CT, Finkers R, Gonzalez-Beltran A, Gray AJG, Groth P, Goble C, Grethe JS, Heringa J, 't Hoen PAC, Hooft R, Kuhn T, Kok R, Kok J, Lusher SJ, Martone ME, Mons A, Packer AL, Persson B, Rocca-Serra P, Roos M, van Schaik R, Sansone S-A, Schultes E, Sengstag T, Slater T, Strawn G, Swertz MA, Thompson M, van der Lei J, van Mulligen E, Velterop J, Waagmeester A, Wittenburg P, Wolstencroft K, Zhao J, and Mons B (2016) The FAIR guiding principles for scientific data management and stewardship. *Scientific Data* 3: 160018. https://doi.org/10.1038/sdata.2016.18.

Wilson CJN, Seward TM, Allan ASR, Charlier BLA, and Bello L (2012) A comment on: 'TitaniQ under pressure: The effect of pressure and temperature on the solubility of Ti in quartz', by Jay B. Thomas, E. Bruce Watson, Frank S. Spear, Philip T. Shemella, Saroj K. Nayak and Antonio Lanzirotti. *Contributions to Mineralogy and Petrology* 164: 359–368. https://doi.org/10.1007/s00410-012-0757-1.

Wilson CJN, Barker SJ, Charlier BLA, Myers ML, and Hansen KF (2021) A comment on: magma residence and eruption at the Taupō Volcanic Center (Taupō Volcanic Zone, New Zealand)—insights from rhyolite-MELTS geobarometry, diffusion chronometry, and crystal textures, by AS Pamukçu et al., Contrib Mineral Petrol 175:48 (2020). *Contributions to Mineralogy and Petrology* 176: 79. https://doi.org/10.1007/s00410-021-01839-9.

Winpenny B and Maclennan J (2011) A partial record of mixing of mantle melts preserved in Icelandic phenocrysts. *Journal of Petrology* 52: 1791–1812. https://doi.org/10.1093/petrology/egr031.

Wong K, Ferguson D, Matthews S, Morgan D, Tadesse AZ, Sinetebeb Y, and Yirgu G (2022) Exploring rift geodynamics in Ethiopia through olivine-spinel Al-exchange thermometry and rare-earth element distributions. *Earth and Planetary Science Letters* 597: 117820. https://doi.org/10.1016/j.epsl.2022.117820.

Wong K, Ferguson D, Wieser P, Morgan D, Edmonds M, Tadesse AZ, Yirgu G, Harvey J, and Hammond S (2023) Focused mid-crustal magma intrusion during continental break-up in Ethiopia. *Geophysical Research Letters* 50: e2023GL103257. https://doi.org/10.1029/2023GL103257.

Wood BJ (1974) The solubility of alumina in orthopyroxene coexisting with garnet. *Contributions to Mineralogy and Petrology* 46: 1–15. https://doi.org/10.1007/BF00377989.

Wood BJ and Blundy J (1997) A predictive model for rare earth element partitioning between clinopyroxene and anhydrous silicate melt. *Contributions to Mineralogy and Petrology* 129: 166–181.

Wright HM, Bacon CR, Vazquez JA, and Sisson TW (2012) Sixty thousand years of magmatic volatile history before the caldera-forming eruption of Mount Mazama, Crater Lake, Oregon. *Contributions to Mineralogy and Petrology* 164: 1027–1052. https://doi.org/10.1007/s00410-012-0787-8.

Yamamoto J and Kagi H (2006) Extended micro-raman densimeter for CO_2 applicable to mantle-originated fluid inclusions. *Chemistry Letters* 35: 610–611. https://doi.org/10.1246/cl.2006.610.

Yang H-J, Kinzler RJ, and Grove TL (1996) Experiments and models of anhydrous, basaltic olivine-plagioclase-augite saturated melts from 0.001 to 10 kbar. *Contributions to Mineralogy and Petrology* 124: 1–18. https://doi.org/10.1007/s004100050169.

Yuan X and Mayanovic RA (2017) An empirical study on Raman peak fitting and its application to Raman quantitative research. *Applied Spectroscopy* 71: 2325–2338. https://doi.org/10.1177/0003702817721527.

Zanon V and Frezzotti ML (2013) Magma storage and ascent conditions beneath Pico and Faial islands (Azores archipelago): A study on fluid inclusions: Magma storage beneath pico and faial. *Geochemistry, Geophysics, Geosystems* 14: 3494–3514. https://doi.org/10.1002/ggge.20221.

Zellmer GF, Sakamoto N, Iizuka Y, Miyoshi M, Tamura Y, Hsieh H-H, and Yurimoto H (2014) Crystal uptake into aphyric arc melts: Insights from two-pyroxene pseudo-decompression paths, plagioclase hygrometry, and measurement of hydrogen in olivines from mafic volcanics of SW Japan. *Geological Society of London, Special Publication* 385: 161–184. https://doi.org/10.1144/SP385.3.

Zhang Y and Namur O (2022) A re-evaluation of the Al-in-Olivine geothermometer. In: *Presented at the Goldcshmidt 2022*. https://conf.goldschmidt.info/goldschmidt/2022/meetingapp.cgi/Paper/11383.

Zhang Y, Xu Z, Zhu M, and Wang H (2007) Silicate melt properties and volcanic eruptions: Silicate melt properties. *Reviews of Geophysics* 45. https://doi.org/10.1029/2006RG000216.

Zhang J, Humphreys MCS, Cooper GF, Davidson JP, and Macpherson CG (2017) Magma mush chemistry at subduction zones, revealed by new melt major element inversion from calcic amphiboles. *American Mineralogist* 102: 1353–1367. https://doi.org/10.2138/am-2017-5928.

Zhukova I, O'Neill H, and Campbell IH (2017) A subsidiary fast-diffusing substitution mechanism of Al in forsterite investigated using diffusion experiments under controlled thermodynamic conditions. *Contributions to Mineralogy and Petrology* 172: 53. https://doi.org/10.1007/s00410-017-1365-x.

Ziberna L (2021) Geothermobarometry of mafic and ultramafic xenoliths: examples from Hualalai and Mauna Kea volcanoes, Hawaii. In: *Crustal Magmatic System Evolution: Anatomy, Architecture, and Physico-Chemical Processes*, pp. 1–18. John Wiley & Sons.

Ziberna L, Green ECR, and Blundy JD (2017) Multiple-reaction geobarometry for olivine-bearing igneous rocks. *American Mineralogist* 102: 2349–2366. https://doi.org/10.2138/am-2017-6154.