Article

Thermodynamic insights into the reliability of mineral-based thermobarometers

Check for updates

Xudong Wang[1], Tong Hou [1]✉, Penny E. Wieser [2] & Zhaochong Zhang [1]

To understand magmatic processes within the silicate planets, accurate pressure and temperature estimates are essential. These conditions are often measured using thermobarometers based on mineral equilibria, but their precision is limited, particularly for barometers, with standard error estimates often exceeding 200 megapascals. Here we present an approach combining thermodynamics and statistics to rigorously investigate these uncertainties. Uncertainty propagation and Sobol analysis reveal that thermometer uncertainties are primarily driven by compositional analytical uncertainty, while barometer uncertainties are dominated by model parameters. Given the quality of current calibration data, even a perfect thermodynamic model has a built-in minimum uncertainty of 120–240 megapascals for barometer and 22–41 °C for thermometer. Furthermore, we demonstrate that pressure and temperature estimates are interdependent; an error in one inevitably compromises the other. Our work provides a pathway for building more reliable thermobarometric models and underscores the need for new, high-quality experimental data.

The architecture of magmatic systems within the crust and their corresponding thermal state serves as the bedrock for building a comprehensive understanding of magmatism, which is perhaps the most critical factor in understanding planetary evolution. Temperature (T) and pressure (P) are two pivotal parameters that determine phase equilibrium assemblages[1,2], the solubility of volatiles in magma[3,4], magma density, and rheology[5,6], thereby influencing the state and evolution of magma, the formation, and growth of magma reservoirs, volcanic eruption styles, and the formation of mineral deposits[7–12]. Consequently, accurately constraining P–T has long been a focal issue in Earth science research[13].

Magma physiochemical meters (MPMs) based on mineral phase equilibria, including thermometers (T), barometers (P), hygrometers (H_2O), and oxybarometers (fO_2), have offered unparalleled information on the physiochemical parameters of magmatic reservoirs[14,15], which are hard to replace by other techniques. All MPMs, in essence, can be considered algorithms involving mapping functions that relate compositional inputs to specific physiochemical parameters. This mapping relies on the underlying assumption that minerals crystallize in equilibrium at phase boundaries, with all components adhering to thermodynamic equilibrium condition:

$$\Delta G_{\mathrm{m}}^{T,P} = -A = 0 \tag{1}$$

where $\Delta G_{\mathrm{m}}^{T,P}$ is the molar Gibbs Free Energy change at P–T and A is the chemical affinity. In general, the $\Delta G_{\mathrm{m}}^{T,P}$ can be expressed as:

$$\Delta G_{\mathrm{m}}^{T,P} = \Delta H^{P_r,T_r} - T\Delta S^{P_r,T_r} + \int_{T_r}^{T} \Delta C_p \mathrm{d}T - T\int_{T_r}^{T} \frac{\Delta C_p}{T}\mathrm{d}T + \int_{P_r}^{P} \Delta V \mathrm{d}P + RT\ln K \tag{2}$$

where $\Delta H^{P_r,T_r}$ and $\Delta S^{P_r,T_r}$ represent the change of standard enthalpy of formation and entropy at the reference pressure (P_r) and temperature (T_r), ΔC_p is the molar isobaric heat capacity and, ΔV the molar volume between products and reactants. K denotes the equilibrium constant. ΔC_p, ΔV and activity coefficients within K are nonlinear functions of P–T–X. Most MPMs simplify these dependencies by treating ΔC_p and ΔV as constants and considering only the contributions of key component interactions to activity coefficients, thereby enabling the isolation of target parameters (P, T, fO_2) for calibration. Hence, all MPMs are essentially variations of the thermodynamic equilibrium equations between mineral-melt or mineral-mineral, sharing one consistent theoretical framework[13,16]. Over the last decades, considerable numbers of phase-equilibrium-based thermobarometers have been proposed, based on mineral-liquid[17–19], mineral-mineral[20–23], mineral-only[24–26] and liquid-only[27,28], and have become widely adopted within the magma and volcanology research communities. More recently, the use of machine learning in geosciences has also led to novel approaches for

[1]Frontiers Science Center for Deep-time Digital Earth and State Key Laboratory of Geological Processes and Mineral Resources, China University of Geosciences (Beijing), Beijing, China. [2]Department of Earth and Planetary Science, McCone Hall, UC Berkeley, Berkeley, CA, USA. ✉e-mail: thou@cugb.edu.cn

thermobarometers, enabling the use of much larger experimental datasets than traditional models to train high-precision, broadly applicable models[29–31]. Although machine learning algorithms may extract more detailed information from data compared to traditional calibration methods[32,33], independent testing of these models on experimental datasets not used for calibration shows that while both traditional and machine learning models provide reasonable results for temperature, they consistently fail to yield pressure estimates with uncertainties <200 MPa[34–36]. The poor performance of barometers has been explained by several factors, including 1) volume changes in the silicate liquid partially absorbing pressure variations[13], 2) enthalpy changes ($\Delta S = -\partial \Delta G / \partial T$) being more prominent than molar volume changes ($\Delta V = \partial \Delta G / \partial P$)[37] or other thermodynamic limitations of phase-equilibrium-based equilibrium barometers[35] and 3) the insufficient precision of experimental data used for calibration, and a shortage of experiments to fully cover P–T–X space. Among these, the substantial differences in ΔS and ΔV of mineral-forming reactions undoubtedly play a dominant role, and unlike point 3), are fundamental limitations that cannot be overcome. Generally, reactions with large ΔS are T-sensitive and suitable as thermometers, whereas those with large ΔV are P-sensitive and appropriate as barometers[13]. Here, the jadeite equilibrium between clinopyroxene (Cpx) and liquid (Liq) is used as an illustrative example, which is commonly used for barometer calibration due to its large molar volume change[19]:

$$NaAlSi_2O_6(Cpx) = 0.5Na_2SiO_3(Liq) + 0.5Al_2O_3(Liq) + 1.5SiO_2(Liq)$$
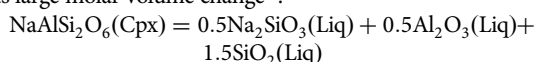
Our representation of liquid species is consistent with the thermodynamic components employed in the MELTS model. We employed Rhyolite-MELTS 1.2.0[38,39], implemented within the framework of the Enabling Knowledge Integration Project (https://gitlab.com/ENKI-portal/ThermoEngine), to calculate the entropy (ΔS) and volume (ΔV) changes associated with this reaction, using the composition and P–T conditions of experiment FM15740, yielding 139.7 J · [mol·°C]$^{-1}$ and 26.7 J · [mol·MPa]$^{-1}$ respectively. That is, for a given uncertainty in the $\Delta G_m^{T,P}$, uncertainty (in MPa) when this reaction is used as a barometer is approximately 5.2 times of the uncertainty (in °C) when used as a thermometer. If a calibrated Cpx-liquid thermometer has an root mean square error (RMSE) of 40 °C, the corresponding barometer might exhibit an RMSE of around 200 MPa. For reactions with a larger ΔS, this disparity would be even greater. Additionally, in natural magmatic systems, the complex evolution history means it is difficult to determine the relationship between phases that may have grown in equilibrium, or possibly in completely different environments[7,40,41]. Moreover, within a single hand sample of volcanic rock, crystals of a single mineral phase may be phenocrysts, antecrysts, and/or xenocrysts, according to the environment in which they formed. Individual crystals may also exhibit zoning and/or complex textures, indicating that different portions of a crystal may have grown in distinct environments[14]. This complexity in the mineral genesis, coupled with changes in water content and matrix glass chemistry due to degassing and microlite formation during magma ascent collectively contributes to a further deterioration in the reliability of mineral ±liquid thermobarometers. Owing to these limitations, single-phase thermobarometers have been widely established and applied[25–33]. Pressure miscalibration also impacts temperature and vice versa since both parameters together dictate chemical equilibrium, and neither parameter is known in most natural systems. Overall, the temperature and pressure information extracted from the mineral and/or glass (melt) compositions of igneous rocks contains several sources of inaccuracy.

We believe that the challenges encountered by phase-equilibrium thermobarometers highlight the incomplete nature of our understanding of their thermodynamic intrinsic nature. Thus, here we provide a thermodynamic perspective that can unambiguously address the following crucial questions and pave the way for the next generation of thermobarometers: 1) How do uncertainties in composition and model parameters contribute to the overall uncertainty in P–T estimates, and which variables are the most influential for thermobarometry? 2) Given known in situ analytical uncertainties in experimental datasets and assuming no additional sources of error, what is the theoretical upper limit of calibration precision? How much additional uncertainty is theoretically introduced when using single-phase thermobarometric models compared to those based on two-phase equilibria? 3) As thermobarometry often requires the simultaneous inference of both P and T from compositional data, to what extent are P and T estimates inherently correlated, regardless of whether a thermobarometric model is primarily P-dependent or T-dependent? If one parameter becomes unconstrained, is the other still robust?

Results and discussion

Integrated approach to P-T uncertainty and sensitivity analysis

We establish a statistically robust framework for evaluating thermobarometers that explicitly quantifies and propagates uncertainties. We develop a simplified thermodynamical model ensuring consistency with fundamental physical principles rather than relying solely on empirical correlations. We exemplify this framework using the jadeite (NaAlSi$_2$O$_6$) equilibrium in the clinopyroxene-liquid system. This equilibrium is expressed thermodynamically as the change in molar Gibbs free energy (ΔG_m) or chemical affinity (A):

$$\Delta G_m = -A = \mu_{Jd}^{Liq} - \mu_{Jd}^{Cpx} \qquad (3)$$

Where μ_{Jd}^{Cpx} is the chemical potential of the jadeite component in the clinopyroxene, and μ_{Jd}^{Liq} is the chemical potential of the jadeite component in the liquid:

$$\mu_{Jd}^{Liq} = 0.5\mu_{Na_2SiO_3}^{Liq} + 0.5\mu_{Al_2O_3}^{Liq} + 1.5\mu_{SiO_2}^{Liq} \qquad (4)$$

For any thermodynamic component i in a phase p, its general molar Gibbs free energy (i.e., chemical potential) can be expressed as:

$$G_m^i(P,T) = \mu_i(P,T) = \Delta_f H^{P_r,T_r} - TS^{P_r,T_r} + \int_{T_r}^{T} C_p dT - T\int_{T_r}^{T} \frac{C_p}{T} dT$$
$$+ \int_{P_r}^{P}\left\{V^{P_r,T_r} - T\left(\frac{\partial V}{\partial T}\right)_P\right\}dP - T\int_{P_r}^{P}\left(\frac{\partial V}{\partial T}\right)_P dP + RT\ln\gamma_i X_i \qquad (5)$$

$\Delta_f H^{P_r,T_r}$, S^{P_r,T_r}, V^{P_r,T_r} represent the standard enthalpy of formation, entropy, and molar volume of component i at the reference pressure (P_r, 0.1 MPa) and temperature (T_r, 298.15 K). C_p is the isobaric heat capacity, and γ_i is the activity coefficient of component i in the melt or solid solution. The expressions for C_p, $V^{P,T}$ and $\ln\gamma_i$ of clinopyroxene are detailed in the Methods section. Melt chemical potentials are calculated using Rhyolite-MELTS 1.2.0[3,38,39].

Thermodynamic parameters of the clinopyroxene solid solution model, including the $\Delta_f H^{P_r,T_r}$, S^{P_r,T_r}, V^{P_r,T_r} and macroscopic interaction energies ($W_{i,j}$), were determined through Bayesian inversion using a Markov Chain Monte Carlo (MCMC) algorithm. This calibration provided full posterior probability distributions for all model parameters, along with their covariance matrix, serving as essential inputs for robust statistical uncertainty propagation analysis. Convergence of the MCMC sampling after the initial "burn-in" phase was assessed through visual inspection of trace plots, confirming that the chains effectively converged to the target posterior distribution (Fig. 1).

To thoroughly evaluate the impact of model parameter uncertainties and analytical uncertainties within clinopyroxene and liquid compositions on the calculated P–T conditions, our framework employs a multi-pronged strategy (see Methods for details): (1) individual parameter uncertainty propagation to understand the isolated impact of each parameter, (2) joint uncertainty propagation that simultaneously samples all model parameters and compositions from their full distributions to quantify the total P–T uncertainty, and (3) uncertainty contribution decomposition via Sobol analysis[42,43], which treats clinopyroxene-liquid compositions and model parameters as two macro-parameters to quantify their relative contributions to the overall P–T uncertainty.

Fig. 1 | MCMC trace plot of the log-likelihood as a function of sampling iterations. The plot demonstrates convergence to a stable posterior distribution following an initial 1000-iteration burn-in phase, with the shaded region indicating the sampled iterations used for parameter estimation. The maximum log likelihood is shown in blue dased line.
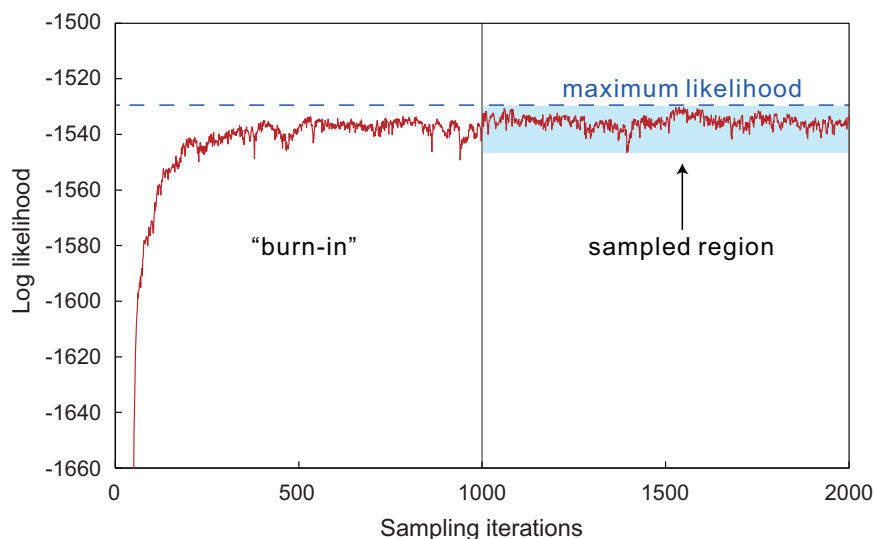
Our results reveal distinct contributions to temperature and pressure uncertainties. We consider analytical uncertainties of 10% for Na_2O, TiO_2, K_2O, Cr_2O_3 and MnO, and 5% for SiO_2–Al_2O_3–FeO–MgO–CaO for both Cpx and Liq analyses, based on reported precisions and expected offsets between laboratories[35,44]. The analytical uncertainty for H_2O in the liquid is also considered, using precision values reported in experiments or, if not available, set at 10%. For temperature calculations, these defined analytical uncertainties in clinopyroxene and liquid compositions contribute the majority of the uncertainty. The first order Sobol indices (S_1, 0.65) indicate that 65% of the total uncertainty comes from error in Cpx and Liq compositions, rising to 85% ($S_T = 0.85$) once interactions with other input parameters is considered. In contrast, the contribution from uncertainty in the model parameters ($\Delta_f H^{P_r,T_r}$, S^{P_r,T_r}, V^{P_r,T_r} and W) to the overall temperature uncertainty is secondary, with S_1 and S_T values of 0.22 and 0.23 respectively (Fig. 2b). Conversely, pressure calculation uncertainties are almost exclusively attributed to model parameters (S_1: 0.79, S_T: 0.81), with analytical uncertainties in compositions contributing only a minor fraction (Fig. 2a; S_1: 0.04, S_T: 0.04). Specifically, uncertainties stemming from W and $\Delta_f H^{P_r,T_r}$ are major sources of model uncertainty in both temperature and pressure calculations. In contrast, the impacts of uncertainties in S^{P_r,T_r} and V^{P_r,T_r} are found to be minimal and negligible, respectively.

The striking divergence in uncertainty attribution can be understood by examining the thermodynamic expressions for chemical potential. Temperature influences the Gibbs free energy primarily through the $RT\ln\gamma_i X_i$ and TS^{P_r,T_r} terms. Minor analytical deviations in mineral and liquid compositions directly affect the X_i and $\ln\gamma_i X_i$, thereby influencing the determined temperature. Conversely, pressure influences the Gibbs free energy change of the reaction predominantly through the $\int_{P_r}^{P} \Delta V dP$ term. Because ΔV is small, even small uncertainties in the calculated Gibbs energy arising from model uncertainties result in large changes in inferred P. For a given temperature, variations in $\Delta_f H^{P_r,T_r}$ and $W_{i,j}$ can cause noticeable variations in ΔG_m which in turn lead to variations in the determined P. Overall, $\Delta_f H^{P_r,T_r}$ and $W_{i,j}$ account for the largest share of model-related uncertainty through their influence on the calculated ΔG_m, while analytical errors in compositional data also substantially increase the total uncertainty in T estimates in practical applications.

Theoretical precision limit defined by analytical uncertainty
We established a thermobarometric algorithm using a Bayesian approach based on Rhyolite-MELTS (v1.2.0[3,38,39], see "Methods") and conducted inversion tests on a test dataset. The overarching conclusion from our inversion tests was the elusive nature of P estimates (Supplementary Note 1). Irrespective of phase information utilized, the predictive capability for P is consistently inferior to that of T and H_2O. Here, H_2O refers to the water

content dissolved in the liquid. Given that P–T–H_2O–fO_2 and composition (X) are all essential parameters for defining the state of a system, if a parameter exerts a substantial impact on the free energy of a system, its role in regulating equilibrium and phase transitions is expected to be pronounced. From a thermodynamic perspective, all thermobarometers aim to find the P-T solution corresponding to $A(P, T, X) = 0$. Under this framework, the influence of compositional uncertainty, and changes in pressure and temperature on the calculated affinity (in joule) can be assessed and directly compared to each other[45–47]. This provides a compelling conceptual framework; the pressure or temperature change required to produce a change in the total chemical affinity equivalent to that caused by compositional uncertainty defines the achievable limit of thermobarometric performance.

Total chemical affinity variations from different parameters (P–T–H_2O–fO_2) were investigated using the solid and liquid compositions from an experiment FM157[48], run at $T = 1050$ °C and $P = 600$ MPa with a liquid composition of 49.4 wt% SiO_2 and 6.1 wt% MgO in equilibrium with plagioclase (Pl, An_{94}), and clinopyroxene (Cpx, Mg# = 81.1). The variability in affinity caused by analytical uncertainties in composition data was calculated under fixed experimental P–T–H_2O–fO_2 conditions for comparison (Fig. 3, green bar). T and H_2O exhibit a strong control over the affinity, with an average variation of 71 J · °C^{-1} (Fig. 3a) and 1.8 kJ · [wt.%]$^{-1}$ for plagioclase (Supplementary Fig. S5a), and 64.3 J · °C^{-1} (Fig. 3b) and 1.4 kJ · [wt.%]$^{-1}$ for clinopyroxene (Supplementary Fig. S5c). Considering that the range of temperature and H_2O values typically span >300 °C and 0–10 wt.%, respectively, for magmatic systems[13,32], we believe the impact of T and H_2O variations can surpass 20 kJ and 14 kJ, respectively. In comparison, the effect of P is markedly diminished, especially on plagioclase (Fig. 3a, c). Considering a variation of 1 GPa (i.e. a ~ 30 km thick crust), the affinity changes by less than 1 kJ. This is comparable to the variation seen for only 10 °C temperature change. The impact of P on clinopyroxene is notably stronger; 1 GPa of variation leads to an affinity change of approximately 10.26 kJ (Fig. 3b), corresponding to an average change of 10.26 J·MPa^{-1}. Even though clinopyroxene shows a marked molar volume response to pressure variations[19,37,49], pressure still exerts a secondary influence on phase equilibria compared with temperature (Fig. 3b, d). Very similar results were observed in completely different systems, with a more chemically evolved liquid from experiment AB62[50] containing 66.59 wt.% SiO_2 and 1.98 wt.% MgO (Supplementary Fig. S6). This points to the general fact that chemical equilibrium in magmatic systems is primarily controlled by temperature and H_2O. This effectively explains why calibrating barometers is often fraught with difficulty, as pressure, in comparison to temperature and H_2O, plays a secondary role in controlling Gibbs energy of reaction and, consequently, chemical equilibrium within magmatic systems, making its influence on chemical composition more susceptible to being obscured by
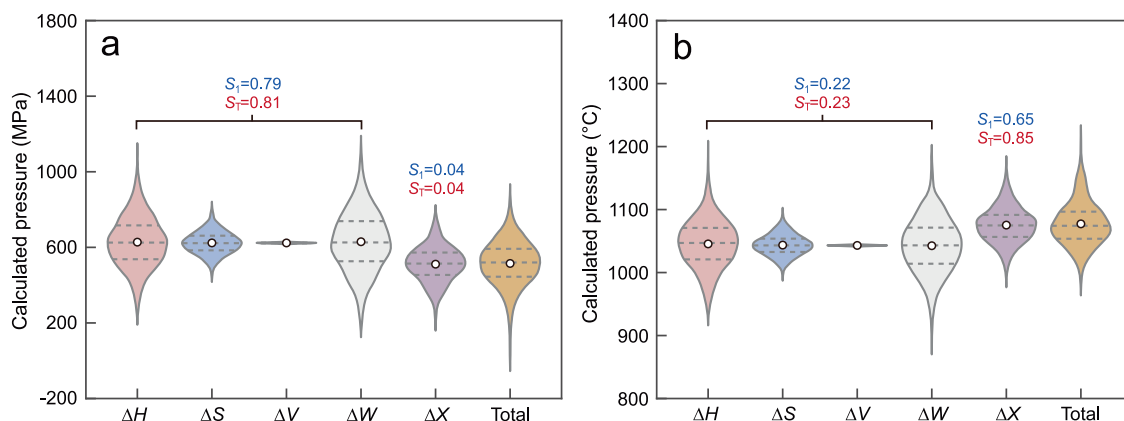
Fig. 2 | Uncertainty propagation and decomposition results for the calibrated model using jadeite-equilibrium. a Violin plots illustrate the distribution of calculated pressure results arising from individual and joint parameter uncertainty propagation. ΔH, ΔS, ΔV represent outcomes from individual uncertainty propagation of $\triangle_f H^{P_r,T_r}$, S^{P_r,T_r}, V^{P_r,T_r} repectively. ΔW, ΔX demonstrate uncertainty propagation when specific groups of parameters are treated as single macro-parameters: ΔW for all macroscopic interaction energies within the clinopyroxene thermodynamic model, and ΔX for the clinopyroxene and liquid compositions. The Total category corresponds to the result of joint parameter uncertainty propagation across all parameters and compositions. **b** uncertainty propagation results for temperature calculations. In each violin plot, the white circle represents the mean of the distribution, and the upper and lower dashed lines indicate the quartiles. S_1 and S_T denote first-order Sobol indices and total-effect Sobol indices, respectively. Precisions of each oxide are set as follows: 10% for Na_2O, TiO_2, K_2O, MnO, Cr_2O_3 and 5% for the others. The experiment selected is FM157[48].
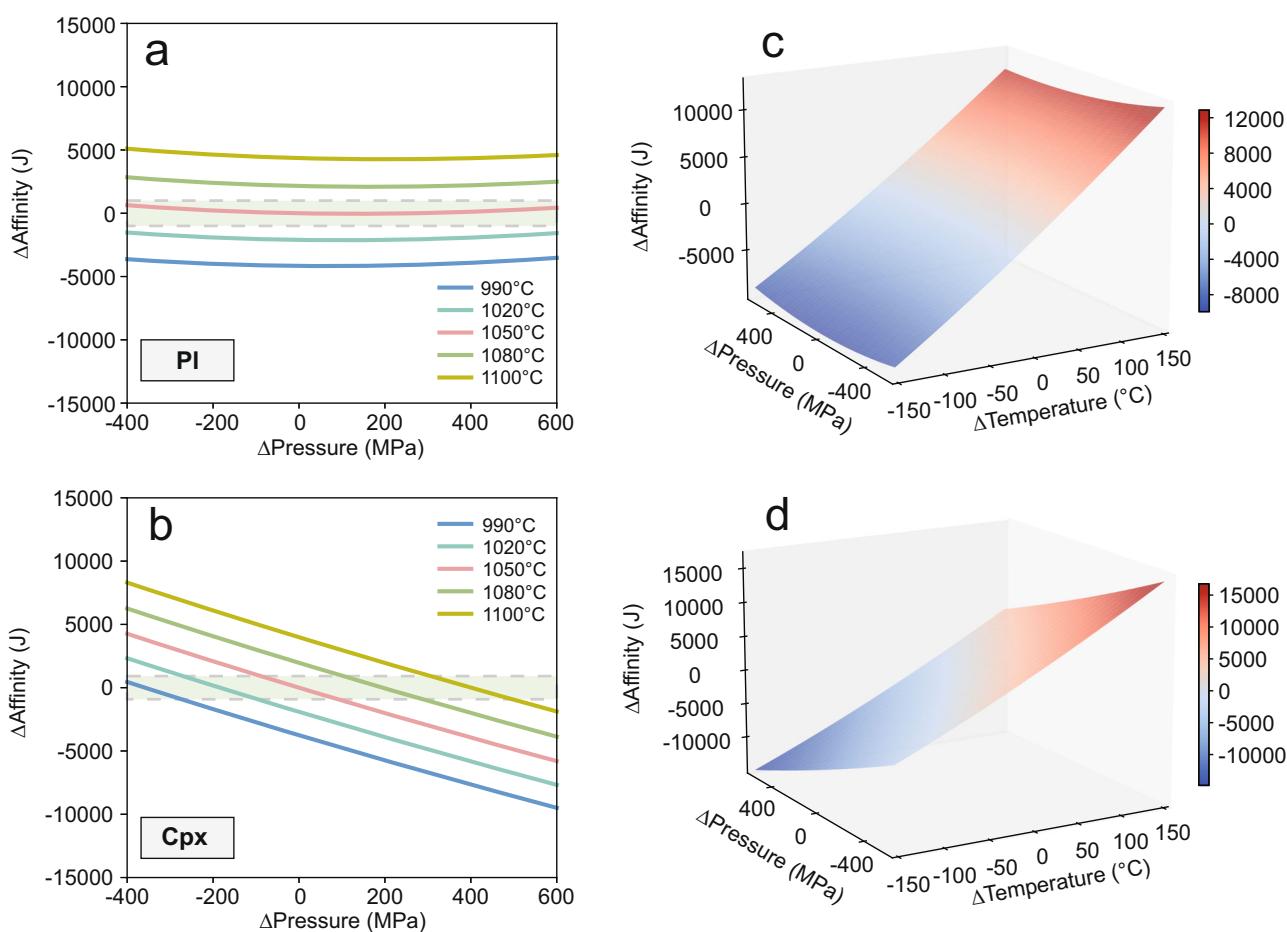
Fig. 3 | Chemical affinity of plagioclase and clinopyroxene formation reaction variations respond to P–T changes. a, b Chemical affinity responds to P–T for plagioclase and clinopyroxene formation. The green shaded area represents the $\pm 2\sigma$ range of affinity variations obtained by 2000 random synthetic solid and liquid composition sampling drawn from Gaussian distribution with synthetic analytical precisions (1σ/mean; wt.%) under the experimental condition. The water content of the liquid used in the calculations is the reported value of the experiment. Precisions of each oxide are set as follows: 10% for Na_2O, TiO_2, K_2O, MnO, Cr_2O_3 and 5% for the others. Lines in different colors signify affinity contour lines for different temperatures. **c, d** 3D representation of the chemical affinity as a function of pressure and temperature. The experiment selected is FM157[48].
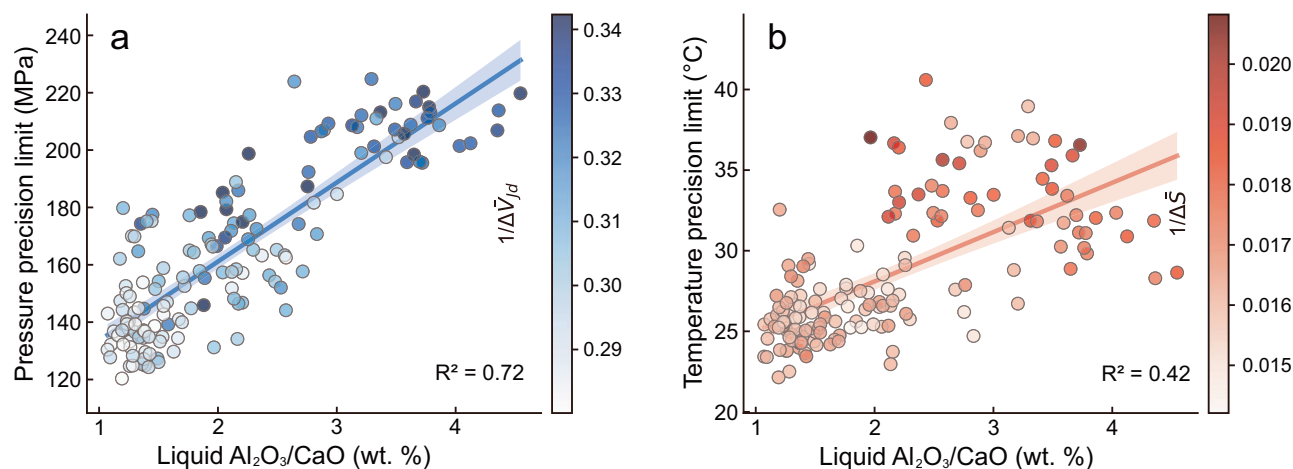
Fig. 4 | Relationship between theoretical precision limits caused by analytical uncertainty and liquid Al$_2$O$_3$/CaO for clinopyroxene-liquid thermobarometer on the test dataset. The pressure (**a**) and temperature (**b**) precision limits are the values corresponding to the ±2σ range of affinity variations, which were calculated from 2000 random synthetic clinopyroxene and liquid composition samples drawn

from a Gaussian distribution with synthetic analytical precisions (1σ/mean; wt.%) under experimental conditions. Precisions of each oxide are the same as described in Fig. 2. Solid lines indicate fitted regressions, and 1σ confidence intervals are represented by shaded regions.

analytical uncertainty, defined as the uncertainty in experimental data. For example, plagioclase has its pressure information completely overridden by analytical uncertainty (Fig. 3a), making it perilous to use as a barometer.

For each experiment in the test dataset, we conduct simulations calculating the range of total chemical affinity variations (±2σ) induced by the synthetic analytical uncertainty. This value, when divided by the change in affinity per °C and per MPa, defines the theoretical precision limit of the thermobarometer under the specified data uncertainty. Calculated precision limits of the Cpx-liquid thermobarometer range from 22 to 41 °C and 120 to 240 MPa, respectively (Fig. 4), with an average of 28 °C and 165 MPa. Many calibration studies for specific systems report SEE within this range[17,19,26], while SEE or RMSE values from global datasets are typically larger[29,31]. An interesting result is that the precision limits of the Cpx-based thermobarometer are dependent on liquid composition; theoretical precision limits are positively correlated with the Al$_2$O$_3$/CaO ratio in the liquid, particularly for barometers (Fig. 4). Thus, for the same assumed % error in oxides, some systems could consistently yield better calibrations than others. This suggests that for global models with broad compositional coverage, standard metrics such as SEE or RMSE do not rigorously reflect the uncertainty for model outputs. Instead, uncertainty propagation must be performed using the posterior means and covariance matrices of model parameters, as demonstrated above. We also observe that $1/\Delta \bar{V}_{Jd}$ and $1/\Delta \bar{S}$ calculated using the P–T obtained via inversion of Rhyolite-MELTS, are positively correlated with the precision limits of thermobarometers (Fig. 4). Here, $\Delta \bar{V}_{Jd}$ represents the molar volume change of the jadeite-forming reaction, and $\Delta \bar{S}$ denotes the average molar entropy change of the clinopyroxene-liquid reaction, defined as $\Delta \bar{S} = \partial A_{tot}/\partial T$, where A_{tot} is the total chemical affinity (see "Methods" section). This correlation arises from their relationships with P–T within the thermodynamic expression. Smaller values of $\Delta \bar{V}_{Jd}$ and $\Delta \bar{S}$ require greater variations in P and T to offset changes in the chemical potential from compositional uncertainty, thereby preserving $A = 0$. Consequently, when evaluating the results of thermobarometric calculations, $1/\Delta \bar{V}_{Jd}$ and $1/\Delta \bar{S}$ can serve as reliability indicators. Larger values of these ratios indicate greater uncertainty in P–T estimates. Additionally, although the chemical equilibrium of plagioclase is primarily controlled by temperature, its performance as a thermometer is not superior to that of clinopyroxene. It tends to exhibit higher uncertainties, as its total chemical affinity variation is higher for identical compositional uncertainty (Fig. 3). It is noteworthy that our adopted analytical uncertainties—10% for Na$_2$O, TiO$_2$, K$_2$O, Cr$_2$O$_3$ and MnO, and 5% for others—represent a highly conservative assumption. This stems from the pervasive phase variability observed in experimental products, expressed through sector zoning,

compositional zoning, and disequilibrium crystallization[51], which reflect inherent chemical heterogeneity within individual mineral phases. These factors, combined with systematic interlaboratory biases in electron microprobe analyses[35,44], indicate that the actual analytical uncertainties in datasets currently utilized for thermobarometer calibration are likely to substantially exceed our assumed values. Our simulations demonstrate that precision limits of Cpx-Liq models will exceeds 300 MPa and 60 °C when assigning 20% uncertainty to low-concentration elements (Na$_2$O, K$_2$O, TiO$_2$, MnO, Cr$_2$O$_3$) and 10% to other elements. Thus, regardless of the algorithms or theoretical models employed, without rigorous data filtering and correction for systematic interlaboratory biases in analytical data, the SEE or RMSE for Cpx-Liq thermobarometer will remain at levels comparable to 120–300 MPa and 20–60 °C. To achieve a SEE/RMSE for the thermobarometer that matches the uncertainty of the experimental apparatus, with the piston-cylinder apparatus commonly exhibiting a lower bound of ±50 MPa and thermocouple a precision of ±10 °C, analytical precision should be at 2%. With this level of precision, the theoretical limits for the Cpx-based thermobarometer are about 8–16 °C and 45–90 MPa, respectively, according to our simulations.

Precision limits of insufficiently constrained thermobarometric models

All MPMs are built upon a fundamental assumption: sufficient constraints are required to uniquely pinpoint an optimal solution. To constrain both T and P, a minimum of two effective constraints are necessary, which translates into at least two thermodynamic equilibrium equations ($\mu_i^\alpha = \mu_i^\beta$). Partial absence of information would lead to incomplete constraints, exposing the model to the risk of non-uniqueness. Intersections on a P–T diagram where mineral phase boundaries with different slopes meet[27,28] or where equilibrium boundaries between two or more thermodynamic components in a mineral solid solution intersect[18,19], can offer complete constraints. It should be noted that the reliability does not seem to improve with more constraints. Inverting Pl-Cpx-liquid is not more accurate than Pl-liquid and Cpx-liquid models (Supplementary Note 1).

The uncertainty associated with non-uniqueness is defined by the span of equivalent solutions. Previous work has suggested that temperature and plagioclase compositions can be extracted using only liquid compositions[32,52] and single-phase compositions of clinopyroxene can also effectively constrain P–T and liquid compositions[30]. At equilibrium, the chemical affinity of each thermodynamic component in the mineral must be zero. Hence, the number of thermodynamic components in a mineral is equal to the number of equations required to describe its equilibrium. For
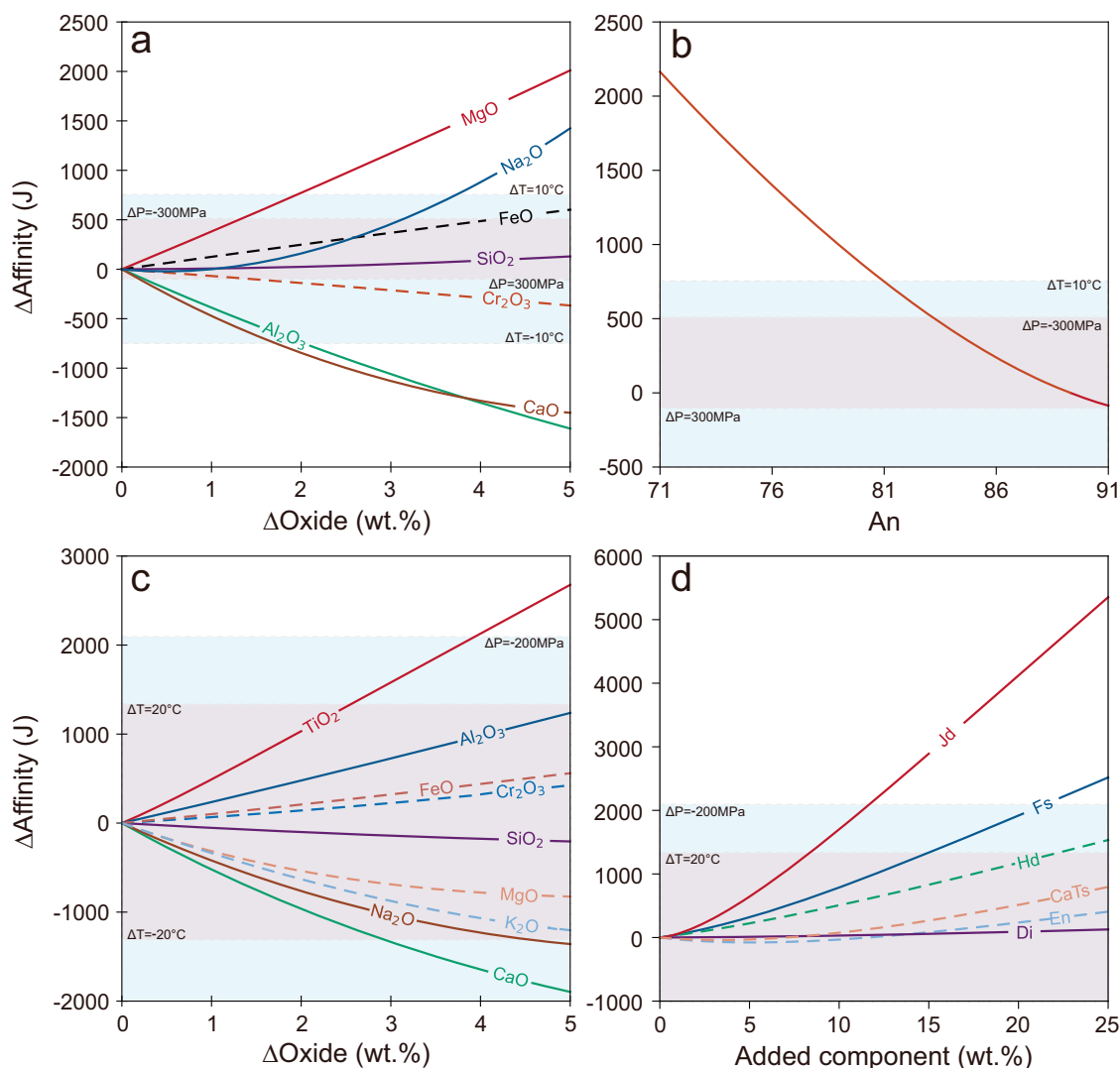
Fig. 5 | Chemical affinity changes as a function of oxide content increase in the liquid phase and varying component additions in the solid phase. a Chemical affinity changes for plagioclase formation as a function of oxide content increase in liquid. The pink and pale blue shaded regions denote the ranges of affinity variations due to ±300 MPa pressure and ±10 °C temperature changes, respectively. **b** Chemical affinity changes as a function of An component. **c, d** The same for clinopyroxene. The pink and pale blue shaded regions represent the ranges of affinity variations due to ±20 °C of temperature and ±200 MPa of pressure changes, respectively. The experiment selected is FM157[48].

models based on the plagioclase phase boundary, five unknowns exist, three endmember components (anorthite, albite, orthoclase) and P–T. The effective constraints are four in total: three equations and the normalization condition that the sum of three endmember components equals 1. For Cpx-only models, there are 11 unknowns, including nine oxide concentrations in the liquid (we exclude MnO and K_2O because they do not constitute primary components of the pyroxene) and P–T. The effective constraints consist of nine equations for the nine thermodynamic components (i.e., enstatite, ferrosilite, diopside, hedenbergite, jadeite, Ca-Tschermak, CaCr-Tschermak, CaTi-Tschermak, CaFe-Tschermak/acmite). To minimize the total deviation during calibration, unknown parameters with the least impact on equilibrium should be set to appropriate fixed values, ensuring the number of effective constraints equals the number of remaining unknowns. This principle also applies to other mineral solid solutions, where the number of thermodynamic components in the solid solution equals to the number of constraints it can provide. The uncertainty imposed on the calibrated model is characterized by the range of solution multiplicity caused by the fluctuations in the unknowns constrained to fixed values.

The effect of each component/oxide in solid and liquid phases on equilibrium was investigated by using total chemical affinity as a metric,

with comparisons made to P–T (Fig. 5). A 5 wt.% or less variation in CaO, Al_2O_3, MgO, and Na_2O in the liquid can produce a comparable variation in total chemical affinity to a 10 °C temperature change. When only the liquid composition (including H_2O) is known, estimating the equilibrium P–T conditions for plagioclase requires implicitly solving for its anorthite (An), albite (Ab), and orthoclase (Or) components. Given that there are five unknown variables (An, Ab, Or, P, T) but only four effective constraints, a unique solution cannot be obtained directly. Therefore, to achieve a unique solution, one of these five variables must be fixed or assumed, which then allows for the determination of the other four. The variable chosen to be assumed should be the one with the least influence on the overall equilibrium, thereby minimizing the total error. If An is chosen as the variable to be fixed, an error of 10 units in the assumed An content will induce a temperature error of approximately 10 °C, or a pressure error exceeding 600 MPa (Fig. 5b). When performing parameter calibration, minimizing the sum of squared residuals over the entire calibration dataset will necessarily position the assumed An value near the center of the An range of the calibration dataset. For instance, if the An range is 10–90 and the assumed An is 50 (assuming P is correctly calculated), this can lead to an uncertainty of up to ±40 °C for the calibrated thermometer. In comparison, if P is held

constant at 500 MPa and the plagioclase composition is perfectly calculated, fluctuations within the 0–1 GPa range induce uncertainty in calculated temperature of just ±15 °C (Fig. 5a). Hence, for plagioclase-saturated liquid-only thermometers, it is necessary to fix P at a constant value to ensure that the number of thermodynamic constraints equals the number of unknown parameters, thereby minimizing the uncertainty in the derived T. Considering a pressure range of 0–1 GPa and assuming no additional sources of uncertainty, the maximum uncertainty for plagioclase-saturated liquid-only thermometers is limited to ±15 °C. Furthermore, because variations in An and T both strongly affect P, plagioclase-saturated liquid-only barometers are inherently unstable. This makes them particularly unsuitable for use as barometers, especially within relatively small pressure ranges (0–1 GPa). For the clinopyroxene-liquid equilibrium, our results indicate that the liquid oxides that most strongly control total affinity are TiO_2, Al_2O_3, Na_2O, and CaO. Variation in each of these 4 oxides causes a change in calculated temperature of up to ±20 °C and ±200 MPa. In contrast, SiO_2, Cr_2O_3, and FeO exert the weakest influence, with the concentration of the latter showing minimal variation in magma systems, causing changes no greater than ±10 °C and ±50 MPa, respectively (Fig. 5c). However, we acknowledge that this sensitivity testing relies on the accuracy of the MELTS model to affinity changes caused by minor components; MELTS does not have a Cr end member in its Cpx, which may explain the lack of sensitivity to Cr despite experimental work showing the opposite[53]. Considering SiO_2 and FeO within a specific interval of the system (e.g., basic or intermediate, with a range of about 10 wt.%), and specifying both SiO_2 and FeO in the liquid, the composition of clinopyroxene can uniquely constrain P–T and the other seven liquid oxide compositions theoretically. In this case, assuming that the liquid SiO_2 and FeO composition is accurately specified, the solution multiplicity range for P–T is approximately ±10 °C and ±50 MPa, which represents the theoretical precision limit of Cpx-only thermobarometers. Any attempts to expand the applicability of such insufficiently constrained thermobarometers will increase solution multiplicity and add uncertainty to the model, rendering them inappropriate for calibration using global datasets.

In both liquid-only and mineral-only models, when solving for P–T, the unknown compositional variables can also be determined, such as An in plagioclase or the Al_2O_3, Na_2O, and CaO concentrations in the liquid. The accuracy of the compositional variables is strongly linked to P–T. If compositional variables are accurately calculated, P–T should also be accurate. Otherwise, both will deviate from true values. This indicates, from a thermodynamic perspective, that liquid-only and mineral-only thermobarometers can be equivalent to liquid-mineral ones. Conversely, relying solely on plagioclase composition to constrain the liquid and P–T is not reliable, as four key variables in the liquid strongly affect the equilibrium, and any sacrifice in their precision results in substantial uncertainty (Fig. 5a). Given the negligible influence of the diopside, enstatite, and Ca-Tschermak components on affinity (Fig. 5d), predicting clinopyroxene equilibrium P–T from the liquid composition remains viable, reducing the number of unknowns to align with the number of constraints. However, most liquid compositions in magmatic processes are transient, and the magmatic evolution is primarily preserved in the zoning of compositional complex minerals such as clinopyroxene rather than as diversity in erupted liquid compositions. Therefore, mineral-only models are more widely applicable than liquid-only or liquid-mineral models.

Deviations arising from composition to P–T inference

The end member equilibrium curve is defined by the P–T solutions where the chemical affinity of a specific thermodynamic end member in a solid solution equals zero on the P–T diagram. All thermobarometers can be considered equivalent to finding intersections between a phase boundary (e.g. Cpx-in) and an endmember equilibrium curve on the P–T diagram (Fig. 6). In this manner, the inference from composition to P–T is entirely controlled by these two curves, with any change or inaccuracy causing a shift in the intersection. We choose an experimental Pl-Cpx-liquid assemblage, considering the endmember equilibrium curves and phase boundaries

calculated from their compositions as "correct" solutions. Variations in the mineral and liquid components/oxides are used to simulate the deviations induced by model uncertainty. Phase boundaries are approximated by curves where total chemical affinities equal zero on the P–T diagram since these curves remain almost fixed when mineral composition undergoes minimal changes (Fig. 6a, c).

Thermobarometric models can be classified into two categories; liquid-dependent which relies on liquid compositional terms and liquid-independent which does not use liquid composition. Since liquid composition is considered, we assume that the phase boundary (Pl-in or Cpx-in) implied by the liquid-dependent models is reliable, then variations in the P–T solution align with the phase boundary. The Pl-in phase boundary has a relatively steep slope ($\approx$18.6 MPa · °C^{-1}). Changing the composition of the plagioclase by 3 wt.% Ab causes a large shift in the Ab equilibrium curve and a minimal shift in the Pl-in phase boundary (red vs. blue curves, Fig.6a). The calculated pressure shifts by 130 MPa, while temperature by only 7 °C (Fig. 6a). For the Cpx-liquid barometer, the moderate slope of the Cpx-in phase boundary ($\approx$6.5 MPa · °C^{-1}) means both pressure and temperature calculations are influenced; a shift in the Jd equilibrium curve by 1 wt.% yields a pressure error of 280 MPa, and a temperature error of 43 °C (Fig. 6c). For liquid-independent models or mineral-liquid models that yields incorrect phase boundaries, the intersections will shift along the direction between the phase boundary and the equilibrium curve of the most pressure-sensitive component (Fig. 6b, d). The calculated temperature for plagioclase demonstrates a variation of 19 °C when the liquid MgO is increased by 3 wt.%, while the pressure shifts by only 70 MPa (Fig. 6b). For Cpx, both pressure and temperature undergo a notable shift after adding 3 wt.% CaO to the liquid, with more pronounced temperature fluctuations compared to adding 1 wt% Jd (Fig. 6d vs. 6c). Clearly, errors in phase boundary calculations will substantially influence temperature results. Furthermore, as the entirety of phase boundary information is encapsulated within the liquid composition, liquid-independent thermobarometers, such as Cpx-only models, are consequently highly susceptible to errors when constraining temperature. Mineral composition, however, mainly affects the pressure results, as phase boundaries generally have steep slopes in magma systems. This explains the observation made using the machine learning method, where temperature information is predominantly stored in the liquid phase, while pressure information is primarily contained in the mineral phase[29].

If the focus is on accurately determining temperature, a plagioclase-based thermometer is quite reliable, with only minimal influence from the pressure. However, if the objective is to accurately calculate both temperature and pressure, pressure-sensitive minerals like clinopyroxene are more reliable. Thermobarometers based on P-sensitive minerals are mutually dependent on the P–T accuracy, making it unlikely for the temperature calculation to be unreliable while the pressure remains accurate, or vice versa. As a result, with the typical SEE/RMSE range of modern thermobarometric models (200–500 MPa and 20–50 °C)[35,36], it is difficult to derive both reliable temperature and pressure from natural samples where both parameters are uncertain.

Conclusions

Our results highlight a critical issue inherent to all current mineral-based thermobarometers: the uncertainties in model outputs are not adequately evaluated, and the generally large uncertainties, coupled with the inherent interdependence of P–T calculations, render the results unreliable for many applications in the crustal range. We propose the following principles to guide the development of next-generation mineral-based thermobarometers: 1) Develop thermodynamically robust models rather than empirical ones. This approach ensures reliable extrapolation and allows the use of metrics such as $1/\Delta \bar{V}_{Jd}$, $1/\Delta \bar{S}$ to assess the reliability of calculated results. 2) Employ rigorous statistical methods to incorporate all sources of uncertainty into the standard deviations and covariance matrices of model parameters, which should be used in practice for error propagation to obtain robust uncertainty estimates. 3) Construct a high-quality experimental
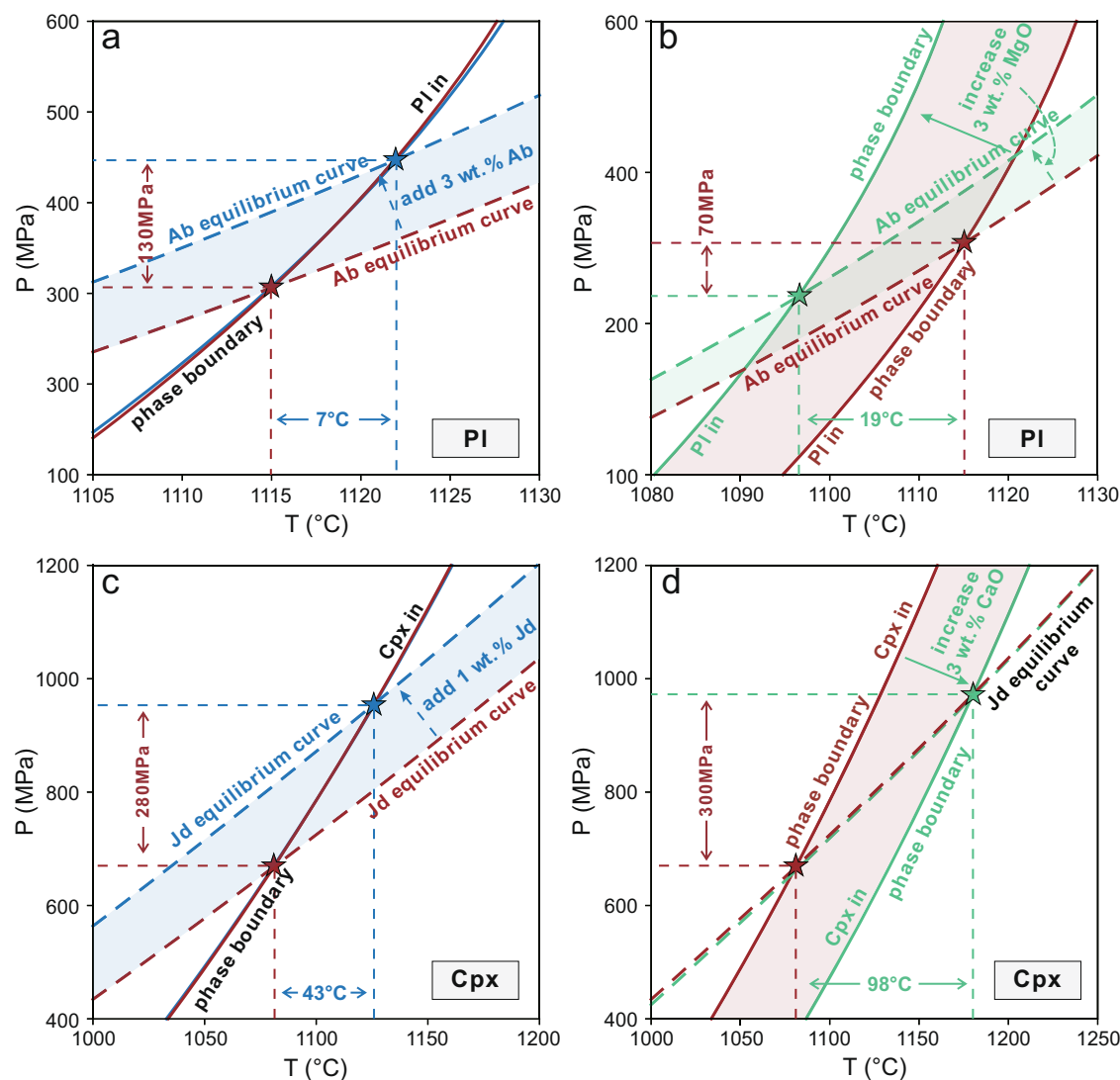
Fig. 6 | P–T diagrams showing how P–T estimates change with phase boundaries and endmember equilibrium curve variations. a Change in Ab equilibrium curves and their intersection points due to the addition of 3 wt.% Ab to plagioclase, with liquid composition fixed. **b** Results obtained by increasing MgO by 3 wt.% in liquid, with the plagioclase composition remaining fixed. **c, d** Variations of intersection points due to the addition of 1 wt.% Jd to clinopyroxene as well as 3 wt.% CaO

increase in liquid, respectively. Red stars represent the intersection points corresponding to the experimental composition, while blue and green stars represent the intersection points after the composition is modified. Solid lines represent P–T solution curves where the total chemical affinity equals 0, approximating phase boundaries. Dashed lines represent the equilibrium curves of corresponding end members. The experiment selected is AB62[50].

dataset to serve as the foundation for model calibration and validation. 4) The parameters $\Delta_f H^{P_r,T_r}$ and $W_{i,j}$ for mineral solid solutions are of critical importance. More precise measurements of $\Delta_f H^{P_r,T_r}$ should be incorporated into models to derive more accurate $W_{i,j}$ thereby improving overall model precision. For solid solutions such as clinopyroxene and plagioclase, calibration should encompass the equilibria of all thermodynamic components, producing parameters and covariance matrices applicable to both mineral-liquid and single-phase thermobarometers. 5) Ensure high analytical data quality when analyzing natural samples to minimize uncertainties stemming from compositional data. For clinopyroxene-based thermobarometers, if either the calculated pressure or temperature is deemed unreliable, both P–T estimates should be considered unreliable.

Methods
Chemical affinity calculations using rhyolite-MELTS
Chemical affinity is a measure of the extent of approaching equilibrium, where a value of zero denotes complete equilibrium. Deviations from zero indicate an increasing degree of disequilibrium. The chemical affinity of a

specific endmember in a mineral-liquid pair with a given composition may be written as[46]:

$$-A = \sum_{i=1}^{P} v_{i,l}\mu_i^{\text{liq}} - \mu_l^{\text{sol}} \tag{6}$$

where $v_{i,l}$ are stoichiometric reaction coefficients, μ_i^j refers to the chemical potential of the ith component in the jth phase, and P is the number of components in liquid. At equilibrium, the left-hand expression for each endmember equals zero. The total chemical affinity is defined as the sum of the chemical affinities of each endmember at a specific pressure and temperature, weighted by their respective mole fractions:

$$A_{\text{tot}} = \sum_{l=1}^{n} X_l A_l \tag{7}$$

where X_l is the mole fraction of the lth endmember, and n is the number of endmembers in the mineral solid solution. It represents the overall

equilibrium state of the mineral solid solution, reaching a value of zero at equilibrium. The MELTS model includes internally consistent thermodynamic properties of silicate melt and minerals suitable for a wide compositional range[38,39], enabling us to calculate chemical affinity under specific P–T–X conditions[47]. Thus, we utilize Rhyolite-MELTS version 1.2.0[3,38,39] built within the framework of the Enabling Knowledge Integration Project (https://gitlab.com/ENKI-portal/ThermoEngine) to implement thermodynamic calculations.

Thermodynamic model for clinopyroxene
Accurate thermobarometry necessitates a robust thermodynamic description of mineral-melt equilibria across relevant P–T–X ranges. The general expression for the molar Gibbs free energy (i.e., chemical potential) of any thermodynamic component i in a phase p is presented in the main text. Isobaric heat capacity C_p is typically represented by a polynomial function of temperature, if lambda transition is ignored[54]:

$$C_p = k_0 + k_1 T^{-0.5} + k_2 T^{-2} + k_3 T^{-3} \qquad (8)$$

The constants k_0 to k_3 are empirical parameters. $V^{P,T}$ is the molar volume at P and T conditions. Its pressure and temperature dependence are approximated by[55]:

$$\frac{V^{P,T}}{V^{P_r,T_r}} = 1 + v_1(P - P_r) + v_2(P - P_r)^2 + v_3(T - T_r) + v_4(T - T_r)^2 \qquad (9)$$

The constants v_1 to v_4 describe the volumetric response. $\ln\gamma_i X_i$ accounts for the non-ideal mixing behavior, where R is the ideal gas constant, X_i is the mole fraction of i, and γ_i is the activity coefficient of component i. For the clinopyroxene solid solution, we employ a symmetric formalism to describe activity coefficients[56]:

$$RT\ln\gamma_l = -\sum_{i}^{n-1}\sum_{j>i}^{n}(p_i' - p_i)(p_j' - p_j)W_{i,j} \qquad (10)$$

In which $p_i' = 1$ when $i = l$ and $p_i' = 0$ when $i \neq l$, p_i is the molar fraction of component i in the solid solution. $W_{i,j}$ is the macroscopic interaction energy between component i and j. In this study, $W_{i,j}$ are treated as constant parameters. While more complex models allow $W_{i,j}$ to be P and T-dependent[57], this simplification is adopted given the constraints of our dataset and the primary focus on the uncertainty quantification framework. The prior and posterior distributions for each parameter, along with their covariance matrix of the posterior distribution, are presented in Supplementary Table S3.

Bayesian inversion theoretical framework
Bayesian inversion has been employed to estimate the magma physicochemical parameters and corresponding uncertainties. This approach represents parameters using a particular probability distribution and infers the joint posterior probability distribution of the parameters of interest through prior information and observational data. Bayesian theory can be outlined as follows[58]:

$$P(\Theta,|,D,H) = \frac{P(D|\Theta,H) \cdot P(\Theta,H)}{P(D|H)} \qquad (11)$$

where D refers to the observational data, Θ is the parameter vector, and H represents the model assumptions. Specifically, D includes the measured mineral compositions and the equilibrium liquid compositions. Θ encompasses the parameters to be inferred to determine the equilibrium state of the system, including T, P, H_2O, fO_2, and liquid composition (X). H comprises the thermodynamic models for minerals and melts that have been selected, along with a basic assumption that the system is in equilibrium between

minerals and melts. $P(\Theta, H) \equiv \pi(\Theta)$ denotes the prior distributions for Θ, representing the information about the parameters obtained through other means or empirical assumptions before the posterior distribution is derived. $P(\Theta,|,D,H)$ is the posterior distribution, representing the conditional probability of the parameters Θ given the observational data D and model assumptions H. $P(D,|,\Theta,H) \equiv L(\Theta)$ is the likelihood function, which measures the degree to which the model results match the observed data given the parameters Θ and assumptions H. A higher likelihood function value indicates a better match between the model predictions and the observed data. Theoretically, all parameters can be inverted provided that the likelihood function offers enough constraints and contains valid information regarding the parameter. Here, we define the likelihood function L as:

$$\ln L = -\frac{\varnothing}{N} \qquad (12)$$

Φ differs across inversion modes: for single-parameter inversion, it is the square of the total chemical affinity, whereas for dual inversions, it is the sum of the squares of the chemical affinities of the two endmembers (anorthite and albite for plagioclase, diopside and jadeite for clinopyroxene). Corrections of 100 °C and 30 °C were conducted respectively for plagioclase and clinopyroxene, to eliminate the systematic deviations present in the test, given our observation of systematic biases in the MELTS model when inverting P-T.

$P(D,|,H) = \int L(\Theta) \cdot \pi(\Theta)d\Theta \equiv Z$ represents the normalization constant, also known as Bayesian evidence, used for comparing different model assumptions. This integral is often analytically and computationally intractable in high-dimensional and nonlinear cases[59]. Nonetheless, this study does not focus on the differences in predictive performance between various thermodynamic models. Our interest lies in the posterior distribution of the parameters, and since there is only one model assumption, the exact value of Z is not required. In such cases, Bayesian inversion can be efficiently approximated using sampling-based algorithms.

Bayesian inversion for calibration of model parameters
Here, we define the likelihood function $L(\Theta)$ based on the calculated chemical affinity for each experimental product:

$$L = \prod_{k=1}^{N} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{A_k^2}{2\sigma^2}\right) \qquad (13)$$

where A_k is the chemical affinity calculated for the k-th experiment given the parameter vector Θ. N is the total number of experiments used for calibration. σ is set to 4000 to ensure sampling efficiency. The prior and posterior distributions of each parameter are shown in Supplementary Tables S1 and S2.

Implement Bayesian inversion via Gibbs sampling
We employed the Gibbs sampling algorithm to approximate the posterior distribution of inversion. This method provides a powerful general approach to exploring complicated high-dimensional shapes of distributions where analytic techniques are infeasible. Gibbs sampling is an iterative procedure used to generate a sequence of samples from the joint posterior distribution of a set of parameters. It decomposes the joint distribution into conditional distributions, allowing each parameter to be sampled sequentially while conditioning on other parameters[60]. The steps for Gibbs sampling are as follows:

Initialization. Randomly generate the initial state of parameters from their prior distributions:

$$\Theta^0 = (\theta_1^0, \theta_2^0, \cdots, \theta_n^0)$$

Iterative Sampling. For iteration t = 1, 2, …, n, make random drawings from the full conditional distributions successively as follows:

$$\theta_1^t \sim p\left(\theta_1, |, \theta_2^{t-1}, \theta_3^{t-1}, \cdots, \theta_n^{t-1}\right)$$

$$\theta_2^t \sim p\left(\theta_2, |, \theta_1^t, \theta_3^{t-1}, \cdots, \theta_n^{t-1}\right)$$

$$\vdots$$

$$\theta_n^t \sim p\left(\theta_n, |, \theta_{-n}^t\right)$$

By this manner, the state transition from Θ^{t-1} to Θ^t is completed. Running this cycle for sufficiently long can generate a Markov chain that converges to the target distribution.

Within the conditional distribution for each step, direct sampling is always impractical if it is analytically intractable. Consequently, the Metropolis-Hastings (M-H) algorithm is frequently embedded within the Gibbs sampling framework to implement the sampling process. For each Metropolis step, the algorithm proceeds as follows:

Parameter proposal. Proposing a new state for the current dimension through a random walk originating from the present state, while holding the parameter values of other dimensions constant.

Calculating acceptance probability. Evaluating the corresponding likelihood of the current and proposed parameter states. To determine whether the proposed state could be accepted into the Markov chain, the M-H algorithm constructs the transition probability by calculating an "acceptance ratio", α, which is defined as the ratio of the posterior probabilities of proposed and current states:

$$\alpha = \frac{L\left(\theta_i^t, \theta_{-i}\right)\pi(\theta_i^t|\theta_{-i})}{L(\theta_i^{t-1}, \theta_{-i})\pi(\theta_i^{t-1}|\theta_{-i})} \tag{14}$$

Accepting or rejecting the proposed state. If $\alpha > 1$, the proposal is accepted and added to the Markov chain. If $\alpha < 1$, the new state is accepted with a probability of α. This is achieved by drawing a random number r from a uniform distribution between 0 and 1, and accepting the new state if $r < \alpha$. However, if $r > \alpha$, the proposed state should be rejected, and the current state should be appended to the chain again.

Repeating the above steps. Take turns sampling each dimension of parameters in the above manner until all dimensions have been traversed. To ensure the Markov chain converges to the posterior distribution, we ran the Gibbs sampler for 10000 iterations, discarding an initial burn-in period of 1000 iterations to eliminate the influence of the starting values. Upon completion of the sampling process, the inverted expected parameter estimates are derived from the median of the posterior distribution.

Experimental dataset for clinopyroxene model calibration and test

The experimental dataset we compiled encompasses a wide range of conditions, spanning temperatures from 800–1300 °C, pressures from 0–1300 MPa and liquid compositions from basalt to rhyolite[48,50,61–64] (Supplementary Fig. S1). These experiments were conducted after the development of Rhyolite-MELTS and form a subset of the ArcPL[35]. The quality of this compiled experimental data is ensured by the following criteria: 1) All compiled experimental studies reported products consistent with near-thermodynamic equilibrium. This was assessed through various methods, including textural evidence, mass balance calculations, mineral-melt partitioning coefficients (e.g., Fe-Mg, Ca-Na), and reversal experiments. 2) These experiments collectively cover a broad spectrum of water contents, ranging from water-poor (<0.1 wt% H_2O) to highly water-rich (up to 9 wt% H_2O). Water content in melt glasses

was either measured or quantitatively determined in all compiled studies, employing techniques such as Fourier transform infrared spectroscopy (FTIR), secondary ion mass spectrometry (SIMS), or the "by-difference" method. 3) All compiled studies reported comprehensive electron probe microanalysis (EPMA) instrument settings, including beam current and counting times for major elements. 4) Clinopyroxene structural formulae, normalized to 6 oxygen atoms, consistently yielded total cation sums within the narrow range of 3.9–4.1, confirming their high analytical quality.

Uncertainty propagation and decomposition strategy

To provide a comprehensive understanding of how uncertainties in model parameters and compositions propagate to the final P–T estimates, we employ a three-step uncertainty analysis:

1) Individual Parameter Sensitivity Analysis. This step quantifies the isolated impact of variations in each parameter. We perform Gaussian sampling for clinopyroxene and liquid compositions, as well as for each model parameter (treating clinopyroxene and liquid compositions as one macro-variable and all W as another macro-variable) based on their mean values and standard deviations obtained from their posterior distributions. For each sampled value, T or P is calculated by given the experimental P or T via inverse modeling, and the resulting distribution of P–T estimates indicates the sensitivity to that specific input.

2) Joint Uncertainty Propagation. This more comprehensive approach accounts for the correlations among parameters. We simultaneously draw samples for clinopyroxene composition, liquid composition, and all model parameters from their respective distributions. Crucially, the model parameters are sampled from a multivariate Gaussian distribution that incorporates the full covariance matrix obtained from the Bayesian inversion. This ensures that the interdependencies identified during the calibration are accurately reflected in the propagated uncertainties, yielding a more realistic total uncertainty in calculated P–T conditions.

3) Uncertainty contribution decomposition. To quantify the relative importance of different uncertainty sources, we apply Sobol analysis, a variance-based global sensitivity analysis method[43]. Let $Y = f(X_1, X_2, \ldots, X_M)$ be the model output (P or T), where X_i are the independent input factors. In this study, we group our parameters into $M = 2$ macro-variables: X_1 (clinopyroxene and liquid compositions) and X_2 (model parameters, including $\Delta_f H^{P_r, T_r}$, S^{P_r, T_r}, V^{P_r, T_r} and macroscopic interaction energies). The total variance of the output $V(Y)$ can be decomposed as:

$$V(Y) = \sum_{i=1}^{M} V_i + \sum_{i<j}^{M} V_{i,j} + \sum_{i<j<l}^{M} V_{i,j,l} + \ldots + V_{1,2,\ldots,M} \tag{15}$$

where $V_i = V(E(Y|X_i))$ represents the variance contributed by the factor X_i alone, $V_{i,j} = V\left(E\left(Y, |, X_i, X_j\right)\right) - V_i - V_j$ represents the variance due to the interaction between X_i and X_j, and so on. The first-order Sobol index (S_1) for an input factor X_i quantifies its individual contribution to the output variance[43,47]:

$$S_1 = \frac{V(E(Y|X_i))}{V(Y)} \tag{16}$$

The total-order Sobol index (S_T) for an input factor X_i quantifies its total contribution to the output variance, including both its first-order effect and all its interactions with other factors:

$$S_T = 1 - \frac{V(E(Y|X_{\sim i}))}{V(Y)} \tag{17}$$

Where $X_{\sim i}$ denotes all input factors except X_i. Sobol indices are calculated using Monte Carlo simulations. Specifically, the Saltelli method is employed for this purpose. The Saltelli method is a widely adopted and computationally efficient sampling strategy for global sensitivity analysis, designed to

accurately estimate both first-order and total-order Sobol indices from a single set of model evaluations[42,65]. The strategy is described as follows.

For D input factors (here $D = M = 2$ macro-variables) and a chosen base sample size N, sampling scheme involves constructing specific input matrices: i) Matrix A. An $N \times D$ matrix where each column corresponds to a factor X_i and each row is an independent sample drawn from the input factor's distribution. ii) Matrix B. Another N×D matrix similar to A but composed of independent samples from the same distributions. iii) Matrices AB, D additional N × D matrices. For each i from 1 to D, the matrix A_B^i is constructed by taking all columns from Matrix A, except for the i-th column, which is taken from Matrix B. This specific construction allows for efficient estimation of partial variances. The model is then evaluated for each row of matrices A, B, and all A_B^i. The total variance V(Y) can be estimated as:

$$V\left(Y_A\right) \approx \frac{1}{N} \sum_{j=1}^{N} Y_{A,j}^2 - \left(\frac{1}{N} \sum_{j=1}^{N} Y_{A,j}\right)^2 \quad (18)$$

$Y_{A,j} = f(A_j)$ denotes the output of j-th row of matrix A. The first-order Sobol index S_1 for X_i is estimated as:

$$S_1 \approx \frac{\frac{1}{N} \sum_{j=1}^{N} Y_{A,j} \cdot Y_{A_{B,j}^{(i)}} - \left(\frac{1}{N} \sum_{j=1}^{N} Y_{A,j}\right)^2}{V\left(Y_A\right)} \quad (19)$$

The total-order Sobol index S_T for X_i is estimated as:

$$S_T \approx 1 - \frac{\frac{1}{N} \sum_{j=1}^{N} Y_{B,j} \cdot Y_{A_{B,j}^{(i)}} - \left(\frac{1}{N} \sum_{j=1}^{N} Y_{B,j}\right)^2}{V\left(Y_B\right)} \quad (20)$$

Where $Y_{A_{B,j}^{(i)}} = f(A_{B,j}^{i})$ denotes the output of j-th row of matrix A_B^i. For this study, 10000 model evaluations were performed for sample size N. Clinopyroxene and liquid oxide compositions were sampled from independent Gaussian distributions, with means equal to reported values and standard deviations set at 10% for low-concentration oxides (Na_2O, K_2O, TiO_2, MnO) and 5% for others. Model parameters were drawn from a multivariate Gaussian distribution, defined by the mean and covariance matrix of their posterior distributions.

Data availability
Experimental dataset and Supplementary Table S3 are available at https://zenodo.org/records/17078877.

Code availability
All codes of results in this work are available from https://zenodo.org/records/17078877.

References
1. Ulmer, P., Kägi, R. & Müntener, O. Experimentally derived intermediate to silica-rich arc magmas by fractional and equilibrium crystallization at 1.0 GPa: an evaluation of phase relationships, compositions, liquid lines of descent and oxygen fugacity. *J. Petrol.* **59**, 11–58 (2018).
2. Waters, L. E., Cottrell, E., Coombs, M. L., Kelley, K. A. Generation of calc-alkaline magmas during crystallization at high oxygen fugacity: an experimental and petrologic study of tephras from Buldir Volcano, Western Aleutian Arc, Alaska, USA. *J. Petrol.* **62**, (2020).
3. Ghiorso, M. S. & Gualda, G. A. R. An H_2O–CO_2 mixed fluid saturation model compatible with rhyolite-MELTS. *Contrib. Mineral. Petrol.* **169**, 53 (2015).
4. Dayton, K. et al. Deep magma storage during the 2021 La Palma eruption. *Sci. Adv.* **9**, eade7641 (2023).
5. Giordano, D., Russell, J. K. & Dingwell, D. B. Viscosity of magmatic liquids: a model. *Earth. Planet. Sci. Lett.* **271**, 123–134 (2008).
6. Pirajno, F. Halogens in Hydrothermal Fluids and Their Role in the Formation and Evolution of Hydrothermal Mineral Systems. In: Harlov D. E., Aranovich L. (eds) *The Role of Halogens in Terrestrial and Extraterrestrial Geochemical Processes: Surface, Crust, and Mantle.* (Springer International Publishing, 2018) pp 759-804.
7. Cashman, K. V., Sparks, R. S. J. & Blundy, J. D. Vertically extensive and unstable magmatic systems: a unified view of igneous processes. *Science* **355**, eaag3055 (2017).
8. La Spina, G., Polacci, M., Burton, M. & de Michieli Vitturi, M. Numerical investigation of permeability models for low viscosity magmas: application to the 2007 Stromboli effusive eruption. *Earth. Planet. Sci. Lett.* **473**, 279–290 (2017).
9. Mutch, E. J. F., Maclennan, J., Shorttle, O., Edmonds, M. & Rudge, J. F. Rapid transcrustal magma movement under Iceland. *Nat. Geosci.* **12**, 569–574 (2019).
10. Pritchard, M. E., Mather, T. A., McNutt, S. R., Delgado, F. J. & Reath, K. Thoughts on the criteria to determine the origin of volcanic unrest as magmatic or non-magmatic. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **377**, 20180008 (2019).
11. Edmonds, M. RESEARCH FOCUS: flotation of magmatic minerals. *Geology* **43**, 655–656 (2015).
12. Hou, T. et al. Immiscible hydrous Fe–Ca–P melt and the origin of iron oxide-apatite ore deposits. *Nat. Commun.* **9**, 1415 (2018).
13. Putirka, K. D. Thermometers and barometers for volcanic systems. *Rev. Mineral. Geochem.* **69**, 61–120 (2008).
14. Pan, R., Hou, T., Wang, X., Encarnación, J., Botcharnikov, R. Multiple magma storage regions and open system processes revealed by chemistry and textures of the Datong tholeiitic lavas, North China Craton. *J. Petrol.* **63**, (2022).
15. Matthews, S. W. et al. A dynamic mid-crustal magma domain revealed by the 2023 to 2024 Sundhnúksgígar eruptions in Iceland. *Science* **386**, 309–314 (2024).
16. Putirka K. D. Geothermometry and Geobarometry. In: White W. M. (ed) *Encyclopedia of Geochemistry: A Comprehensive Reference Source on the Chemistry of the Earth.* (Springer International Publishing, 2018) pp 597-614.
17. Lange, R. A., Frey, H. M. & Hector, J. A thermodynamic model for the plagioclase-liquid hygrometer/thermometer. *Am. Mineral.* **94**, 494–506 (2009).
18. Masotta, M., Mollo, S., Freda, C., Gaeta, M. & Moore, G. Clinopyroxene–liquid thermometers and barometers specific to alkaline differentiated magmas. *Contrib. Mineral. Petrol.* **166**, 1545–1561 (2013).
19. Neave, D. A. & Putirka, K. D. A new clinopyroxene-liquid barometer, and implications for magma storage pressures under Icelandic rift zones. *Am. Mineral.* **102**, 777–794 (2017).
20. Lindsley, D. H. & Andersen, D. J. A two-pyroxene thermometer. *J. Geophys. Res-Sol. Ea.* **88**, A887–A906 (1983).
21. Ghiorso, M. S. & Evans, B. W. Thermodynamics of rhombohedral oxide solid solutions and a revision of the Fe-Ti two-oxide geothermometer and oxygen-barometer. *Am. J. Sci.* **308**, 957–1039 (2008).
22. Wan, Z., Coogan, L. A. & Canil, D. Experimental calibration of aluminum partitioning between olivine and spinel as a geothermometer. *Am. Mineral.* **93**, 1142–1147 (2008).
23. Molina, J. F. et al. A reassessment of the amphibole-plagioclase NaSi-CaAl exchange thermometer with applications to igneous and high-grade metamorphic rocks. *Am. Mineral.* **106**, 782–800 (2021).
24. Nimis, P. & Ulmer, P. Clinopyroxene geobarometry of magmatic rocks Part 1: an expanded structural geobarometer for anhydrous and

hydrous, basic and ultrabasic systems. *Contrib. Mineral. Petrol.* **133**, 122–135 (1998).

25. Ridolfi, F. & Renzulli, A. Calcic amphiboles in calc-alkaline and alkaline magmas: thermobarometric and chemometric empirical equations valid up to 1130 °C and 2.2 GPa. *Contrib. Mineral. Petrol.* **163**, 877–895 (2012).

26. Wang, X. D. et al. A new clinopyroxene thermobarometer for mafic to intermediate magmatic systems. *Eur. J. Mineral.* **33**, 621–637 (2021).

27. Gualda, G. A. R. & Ghiorso, M. S. Phase-equilibrium geobarometers for silicic rocks based on rhyolite-MELTS. Part 1: principles, procedures, and evaluation of the method. *Contrib. Mineral. Petrol.* **168**, 1033 (2014).

28. Higgins, O., Stock, M. J. A new calibration of the OPAM thermobarometer for anhydrous and hydrous mafic systems. *J. Petrol.* **65**, (2024).

29. Petrelli, M., Caricchi, L. & Perugini, D. Machine learning thermobarometry: application to clinopyroxene-bearing magmas. *J. Geophys. Res-Sol. Ea.* **125**, e2020JB020130 (2020).

30. Higgins, O., Sheldrake, T. & Caricchi, L. Machine learning thermobarometry and chemometry using amphibole and clinopyroxene: a window into the roots of an arc volcano (Mount Liamuiga, Saint Kitts). *Contrib. Mineral. Petrol.* **177**, 10 (2021).

31. Jorgenson, C., Higgins, O., Petrelli, M., Bégué, F. & Caricchi, L. A machine learning-based approach to clinopyroxene thermobarometry: Model optimization and distribution for use in Earth Sciences. *J. Geophys. Res-Sol. Ea.* **127**, e2021JB022904 (2022).

32. Cutler, K. S., Cassidy, M. & Blundy, J. D. Plagioclase-saturated melt hygrothermobarometry and plagioclase-melt equilibria using machine learning. *Geochem. Geophys. Geosyst.* **25**, e2023GC011357 (2024).

33. Weber, G., Blundy, J. A machine learning-based thermobarometer for magmatic liquids. *J. Petrol.* **65**, (2024).

34. Erdmann, S., Martel, C., Pichavant, M. & Kushnir, A. 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. *Contrib. Mineral. Petrol.* **167**, 1016 (2014).

35. Wieser, P. E. et al. Barometers behaving badly I: assessing the influence of analytical and experimental uncertainty on clinopyroxene thermobarometry calculations at crustal conditions. *J. Petrol.* **64**, egac126 (2022).

36. Wieser, P. E., Gleeson, M. L. M., Matthews, S., DeVitre, C., Gazel, E. Determining the pressure-temperature-composition (P-T-X) conditions of magma storage. In: Anbar A., Weis D. (eds). *Treatise on Geochemistry (Third edition)*. (Elsevier, 2025) pp 83−151.

37. Putirka, K. Amphibole thermometers and barometers for igneous systems and some implications for eruption mechanisms of felsic magmas at arc volcanoes. *Am. Mineral.* **101**, 841–858 (2016).

38. Gualda, G. A. R., Ghiorso, M. S., Lemons, R. V. & Carley, T. L. Rhyolite-MELTS: a modified calibration of MELTS optimized for silica-rich, fluid-bearing magmatic systems. *J. Petrol.* **53**, 875–890 (2012).

39. Ghiorso, M. S. & Sack, R. O. 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. *Contrib. Mineral. Petrol.* **119**, 197–212 (1995).

40. Humphreys, M. C. S., Blundy, J. D. & Sparks, R. S. J. Magma evolution and open-system processes at Shiveluch volcano: Insights from phenocryst zoning. *J. Petrol.* **47**, 2303–2334 (2006).

41. Jackson, M. D., Blundy, J. & Sparks, R. S. J. Chemical differentiation, cold storage and remobilization of magma in the Earth's crust. *Nature* **564**, 405–409 (2018).

42. Saltelli, A. Making best use of model evaluations to compute sensitivity indices. *Comput. Phys. Commun.* **145**, 280–297 (2002).

43. Sobol, I. M. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Math. Comput. Simul.* **55**, 271–280 (2001).

44. Gale, A., Dalton, C. A., Langmuir, C. H., Su, Y. & Schilling, J.-G. The mean composition of ocean ridge basalts. *Geochem. Geophys. Geosyst.* **14**, 489–518 (2013).

45. Ghiorso, M. S., Carmichael, I. S. E., Rivers, M. L. & Sack, R. O. The Gibbs free energy of mixing of natural silicate liquids: an expanded regular solution approximation for the calculation of magmatic intensive variables. *Contrib. Mineral. Petrol.* **84**, 107–145 (1983).

46. Ghiorso, M. S. Algorithms for the estimation of phase stability in heterogeneous thermodynamic systems. *Geochim. Cosmochim. Acta* **58**, 5489–5501 (1994).

47. Huggins, E. G., Ruprecht, P. & Ghiorso, M. S. Using chemical affinities to understand disequilibrium textures of plagioclase preserved in magmatic systems. *Geophys. Res. Lett.* **48**, e2021GL092884 (2021).

48. Marxer, F., Ulmer, P. & Müntener, O. Polybaric fractional crystallisation of arc magmas: an experimental study simulating trans-crustal magmatic systems. *Contrib. Mineral. Petrol.* **177**, 3 (2021).

49. Putirka, K. D., Johnson, M., Kinzler, R., Longhi, J. & Walker, D. Thermobarometry of mafic igneous rocks based on clinopyroxene-liquid equilibria, 0–30 kbar. *Contrib. Mineral. Petrol.* **123**, 92–108 (1996).

50. Almeev, R. R., Holtz, F., Ariskin, A. A. & Kimura, J.-I. Storage conditions of Bezymianny Volcano parental magmas: results of phase equilibria experiments at 100 and 700 MPa. *Contrib. Miner. Pet.* **166**, 1389–1414 (2013).

51. Neave, D. A. et al. Clinopyroxene–liquid equilibria and geothermobarometry in natural and experimental tholeiites: the 2014–2015 Holuhraun Eruption, Iceland. *J. Pet.* **60**, 1653–1680 (2019).

52. Neave, D. A. & Namur, O. Plagioclase archives of depleted melts in the oceanic crust. *Geology* **50**, 848–852 (2022).

53. Voigt, M., Coogan, L. A. & von der Handt, A. Experimental investigation of the stability of clinopyroxene in mid-ocean ridge basalts: the role of Cr and Ca/Al. *Lithos* **274-275**, 240–253 (2017).

54. Berman, R. G. & Brown, T. H. Heat capacity of minerals in the system $Na_2O-K_2O-CaO-MgO-FeO-Fe_2O_3-Al_2O_3-SiO_2-TiO_2-H_2O-CO_2$: representation, estimation, and high temperature extrapolation. *Contrib. Miner. Pet.* **89**, 168–183 (1985).

55. Berman, R. G. Internally-Consistent Thermodynamic Data for Minerals in the System $Na_2O-K_2O-CaO-MgO-FeO-Fe_2O_3-Al_2O_3-SiO_2-TiO_2-H_2O-CO_2$. *J. Pet.* **29**, 445–522 (1988).

56. Powell, R. & Holland, T. On the formulation of simple mixing models for complex phases. *Am. Miner.* **78**, 1174–1180 (1993).

57. Holland, T. & Powell, R. Activity–composition relations for phases in petrological calculations: an asymmetric multicomponent formulation. *Contrib. Miner. Pet.* **145**, 492–501 (2003).

58. van de Schoot, R. et al. Bayesian statistics and modelling. *Nat. Rev. Methods Prim.* **1**, 1 (2021).

59. Casella, G. & George, E. I. Explaining the Gibbs sampler. *Am. Stat.* **46**, 167–174 (1992).

60. Hara, R. B. O. & Sillanpää, M. J. A review of Bayesian variable selection methods: What, how and which. *Bayesian Anal.* **4**, 85–117 (2009).

61. Blatter, D. L., Sisson, T. W. & Hankins, W. B. Voluminous arc dacites as amphibole reaction-boundary liquids. *Contrib. Miner. Pet.* **172**, 27 (2017).

62. Melekhova, E., Blundy, J., Robertson, R. & Humphreys, M. C. S. Experimental evidence for polybaric differentiation of primitive arc Basalt beneath St. Vincent, Lesser Antilles. *J. Pet.* **56**, 161–192 (2015).

63. Parat, F., Streck, M. J., Holtz, F. & Almeev, R. Experimental study into the petrogenesis of crystal-rich basaltic to andesitic magmas at Arenal volcano. *Contrib. Miner. Pet.* **168**, 1040 (2014).

64. Husen, A., Almeev, R. R. & Holtz, F. The effect of H_2O and pressure on multiple saturation and liquid lines of descent in Basalt from the Shatsky Rise. *J. Pet.* **57**, 309–344 (2016).

65. Saltelli, A. et al. Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Comput. Phys. Commun.* **181**, 259–270 (2010).

Author contributions
T.H. and X.W. conceived the idea. X.W. developed the idea and wrote the code. X.W., T.H., and P.W. performed the data analysis. X.W., T.H., P.W., Z.Z. co-wrote the manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43247-025-02831-y.

Correspondence and requests for materials should be addressed to Tong Hou.

Peer review information *Communications Earth & Environment* thanks Oliver Higgin, Frank Spera and Wendy Bohrson for their contribution to the peer review of this work. Primary Handling Editors: Joe Aslin and Carolina Ortiz Guerrero. [A peer review file is available].

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.