An Inverse Method for Quantifying Petrological Parameters and Uncertainty in Phase Equilibrium Modelling

T. Mackay-Champion[1], I.P. Cawood[1,2]

1. Department of Earth Sciences, University of Oxford, South Parks Road, Oxford, OX1 3AN, UK.

2. Department of Earth and Planetary Sciences, The University of Hong Kong, Pokfulam Road, Hong Kong.

Email: ipcawood@hku.hk

Suggested short title: An Inverse Method for Phase Equilibrium Modelling

ABSTRACT

11 Phase equilibrium modelling offers a powerful quantitative framework for understanding

12 petrological processes. Yet many studies still rely on qualitative comparisons between natu-

13 ral datasets and these forward-modelled predictions to constrain model parameters, commonly

14 pressure-temperature (P-T) conditions. Compounding this, uncertainties from the observed

15 data or within the modelled predictions are rarely quantified, limiting confidence in the esti-

16 mated P-T conditions and resulting petrological interpretations. We introduce LinaForma, an

17 inverse modelling workflow that determines best-fit P-T conditions (or other petrological pa-

18 rameters) and their associated uncertainties for a given rock system by minimizing the misfit

19 between observed data (e.g., mineral compositions or modal proportions) and their forward-

20 model predictions. Uncertainty is quantified by resampling of the observed data with replace-

21 ment. Diagnostic metrics identify poorly performing variables and assess the sensitivity of the

22 inversion result to variable uncertainty. Applied to an amphibolite-facies pelite and metabasite

23 from the Greater Himalayan Sequence (Zanskar Himalaya, NW India), the approach proves ef-

24 fective across contrasting model systems, using different sets of solution models and variables to

25 produce P–T estimates consistent with classical thermobarometry.

26 The workflow offers several advantages: compatibility with outputs from any forward-

27 modelling software; flexible variable selection; systematic grid-search inversion in multidimen-

28 sional space; a robust L1-norm misfit function resistant to outliers; and sensitivity and un-

29 certainty analysis via bootstrap resampling. Limitations include the increasing computational

30 demands for high-dimensional grids (N > 2) and the absence of explicit quantification of un-

31 certainties inherited from the thermodynamic dataset and solution models. Alongside other

32 emerging quantitative methods, LinaForma enables petrologists to make more informed interpre-

33 tations of complex metamorphic systems and target improvements to thermodynamic datasets,

34 advancing the pursuit of optimal P-T estimates.

1 INTRODUCTION

35 Phase equilibrium modelling is the foundation of modern metamorphic petrology, providing

36 a powerful quantitative framework to decipher the evolution of metamorphic sequences and,

37 in turn, offering critical insights into the processes that shape Earth's crust and mantle. This

38 modelling is built on large datasets including thermodynamic end-member properties and activ-

39 ity–composition (a–X) relations of minerals, fluid and melt (e.g., Berman, 1988; Holland & Pow-

40 ell, 1998). When integrated with appropriate software, these datasets can be used to produce

41 sophisticated forward models that predict equilibrium phase assemblages, compositions, and

42 modal proportions, alongside numerous other variables across pressure-temperature-composition

43 (P-T-X) space (e.g., Spear *et al.*, 2016).

44   Despite the quantitative foundation of these techniques, petrological studies predominantly

45 reconstruct the metamorphic evolution of a rock using qualitative to semi-quantitative compar-

46 isons between the measured values from the rock and their forward-modelled predictions (Powell

47 & Holland, 2008). The quality of these comparisons depends on understanding the uncertain-

48 ties associated with both the input data and the modelling process. As is standard in most

49 scientific fields, any calculation must include a quantifiable estimate of uncertainties, alongside

50 recognition of uncertainties that cannot be readily quantified (Powell & Holland, 1994, 2008).

51 However, in petrological studies, such uncertainties are often overlooked, such that derived model

52 parameters, commonly P-T conditions, cannot be interpreted within a defined confidence range,

53 thereby limiting confidence in the resulting petrological interpretations.

54   Although several programs employ quantitative approaches for determining optimal P-T

55 conditions (e.g., Berman, 1991; Gordon, 1992; Powell & Holland, 1994; Duesterhoeft & Lanari,

56 2020; Nerone *et al.*, 2025), many of these tools are tied to specific software, are not integrated

57 with phase equilibrium modelling, or do not place a strong emphasis on uncertainty quantifi-

58 cation. Therefore, the impact of uncertainties on P-T estimates produced by current thermo-

59 barometric techniques is still difficult to estimate. To address these limitations, we present

60 LinaForma, a flexible workflow that quantitatively determines the best-fit P–T solution and

61 associated uncertainty for a given rock system. In this study, we (1) outline the method and

62 provide guidelines for using the workflow, (2) apply it to a pelite (or metapelite) and metab-

63 asite sample to test its effectiveness across contrasting metamorphic systems, (3) compare it

64 with other quantitative approaches, and (4) highlight the advantages, potential limitations, and

65 avenues for future work for quantitative methods.

2   INVERSION METHOD

66   Forward modelling is the process of predicting the data one would observe for a given set

67 of model input parameters. A pseudosection is an example of a forward model, in which the

68 equilibrium phase assemblage, phase compositions, and proportions are predicted for a rock of

69 a given chemical composition across P-T space. This is contrasted with inverse modelling in

70 which observed data is used to estimate the underlying model parameters. Single-equilibrium

71 thermobarometry (e.g., Ti-in-biotite geothermometer; Henry et $al.$, 2005) and multi-equilibrium

72 thermobarometry (Powell & Holland, 1994) are examples of inverse methods, where P–T con-

73 ditions are estimated from the observed mineral compositions.

74   Constraining the conditions of equilibration and evolution of a rock is frequently undertaken

75 using intersecting isopleths of observed phase compositions or modal proportions calculated using

76 phase equilibrium modelling. In its simplest form, this can involve using a mean or representative

77 analysis from two variables to determine a singular cross over point in P-T space (Figure 1a).

78 When the observed data distribution of two or more variables are considered, varying zones of

79 overlap may emerge (Figure 1b–d). In such cases, the intersections may be ambiguous, especially

80 when the observed distributions differ considerably or the variables have differing sensitivities to

81 P–T. Even a narrow observed population distribution can correspond to a broad calculated P–T

82 interval if the variable is highly responsive to P–T changes (Figure 1b). Additional variables may

83 make for a more "representative" result, but can also result in multiple zones of overlap (Figure

84 1c), or poor to no agreement (Figure 1d). While this method has been effectively applied to

85 semi-quantitatively analyze the tectonothermal evolution of metamorphic terranes (e.g., Vance

86 & Mahar, 1998; Hoschek, 2004; Štípská & Powell, 2005), the accuracy and precision of the

87 resulting P-T estimates, and thereby the conclusions drawn from the sample, are significantly

88 influenced by the variables selected and their associated uncertainties (Figure 1a–d).

89   To address these challenges, the workflow presented in this study, LinaForma, inverts a

90 large set of variables simultaneously and uses the uncertainties in the mineral measurements to

91 estimate variability in the derived P-T estimates (Figure 1e). For clarity, we present the method

92 for P–T conditions, though it is equally applicable to any set of model parameters of interest.

93 The workflow uses a grid-search inversion, which minimizes misfit between the mineral measure-

94 ments and their forward-modelled predictions (Figure 2a), coupled with bootstrap resampling

95 of mineral measurements (Figure 2b), to quantify the best-fit P–T conditions and associated

96 uncertainties for a given bulk rock composition. We focus specifically on uncertainties in the

97 mineral measurements because they reflect the scale of equilibrium preserved in natural samples,

98 directly constrain the precision of the results, and can be readily incorporated into uncertainty

99 analysis. Uncertainties in the thermodynamic end-member data set and in the interaction en-
100 ergies used in the activity-composition (a–X) relationships also contribute significantly to the
101 overall P–T uncertainty (Powell & Holland, 1988, 2008). However, properly accounting for these
102 effects requires recalculating the forward model for each variation—a computationally intensive
103 process—and the uncertainties are not always well constrained. Whilst these uncertainties are
104 not explicitly estimated in this workflow; the bootstrap resampling approach inherently reflects
105 some of the additional dispersion they introduce. As a large number of mineral composition
106 variables are used in this study, their definitions are included in Table 1 for reference.

2.1 Best-fit P-T conditions: the grid-search inversion

108 A grid-search inversion is used to determine the optimal or best-fit P-T conditions for a
109 given rock system. The grid-search involves drawing a large number of trial solutions from
110 a regular grid in model (e.g., P-T) space. For each point on the grid (i.e., a trial solution),
111 the difference between the forward model's predicted data and the observed data is computed
112 using an objective, or misfit, function (Figure 2a). The best-fit solution is the point on the
113 grid with the lowest value of the misfit function. In this case, the observed data may include
114 mineral compositions, mineral modal proportions, and bulk properties such as the interpreted
115 equilibrium phase assemblage and whole-rock density. The forward models for each point on the
116 grid can be calculated using programs such as Theriak-Domino (de Capitani & Brown, 1987;
117 de Capitani & Petrakakis, 2010), Perple_X (Connolly, 1990, 2005), and MAGEMin (Riel *et al.*,
118 2022).

119 This workflow employs a normalized L1-type misfit function (Φ), expressed as the sum of
120 absolute residuals scaled by the observed values (Equation 1):

$$\Phi = \sum_{i=1}^{N} \frac{|x_i^{\mathrm{obs}} - x_i^{\mathrm{mod}}|}{x_i^{\mathrm{obs}}} \tag{1}$$

121 where N is the total number of variables, x_i^{mod} is the modelled value of variable i predicted
122 at the trial solution, and x_i^{obs} is the observed value of variable i. The residuals are scaled relative
123 to the observed value of each variable to ensure that variables with large magnitudes do not have
124 an overwhelming influence on the result. Large values of Φ suggest the model predictions poorly
125 fit the observed data, whilst low values of Φ show the observed data and the model predictions

126 are similar (i.e., the data "residuals" are small). The trial solution with the lowest value of Φ is

127 the best-fit solution.

2.2   Uncertainty analysis: bootstrap resampling

129   Bootstrap resampling is used to estimate the uncertainty associated with the best-fit P-T

130 solution. This involves resampling the observational data with replacement, and then computing

131 the required statistics for each resampled dataset, in this case the best-fit solution of the grid-

132 search inversion (Figure 2b). When repeated multiple times (e.g., $N \geq 1000$), the distribution

133 of inversion solutions obtained from the resampled datasets provides an estimate of the true

134 mean or median solution and the population distribution (Menke, 1984). This assumes the

135 method itself is inerrant, the model formulation is appropriate, and that the phases from which

136 the variables are selected are in mutual equilibrium—the limitations of these assumptions are

137 considered in Section **??**.

138   Bootstrap resampling can be either non-parametric or parametric. Non-parametric boot-

139 strapping involves resampling with replacement from the original observational dataset. Para-

140 metric bootstrapping generates samples according to an assumed distribution of the obser-

141 vational data (Efron, 1979), in this case a normal distribution (Figure 2b). Given that the

142 probability density function in non-linear problems is typically non-Gaussian (Menke, 1984),

143 percentile-based confidence intervals such as the interquartile range are often more informative

144 than the standard deviation, and are therefore preferred for results where possible. Bootstrap-

145 resampling is also used to assess the sensitivity of the inversion results to uncertainty in each

146 variable (Section 2.3.3, Figure 2c).

2.3   Diagnostics

148   The workflow includes three diagnostics with which to validate and refine the best-fit solu-

149 tion: (1) Quality of data fit (f_i, f_{total}) between the best-fit solution and the observations; (2)

150 the influence of each variable ($f_{\text{total}(i)}$) on the best-fit solution, and (3) the sensitivity of the

151 best-fit solution (ΔP_i, ΔT_i) to uncertainty in each of the chosen variables. The values of these

152 diagnostic metrics and full results for each solution are automatically output in table format at

153 the end of the inversion calculations. The definitions for each value reported in the results table

154 are included in Table 2.

2.3.1   Quality of data fit (f_i, f_{total})

155

The data fit metric assesses the level of agreement between the median best-fit solution,
post-bootstrap re-sampling, and the observations. This is quantified by assessing whether the
model's predictions at the median best-fit conditions (e.g., the modelled value of X_{Alm} garnet at
PT_{best}) fall within two standard deviations of the observed values. This assessment is provided
in the workflow by a score for each variable (f_i) and a total score (f_{total}). These are calculated
following Equations 2 and 3, respectively:

$$f_i = \frac{|x_i^{\text{obs}} - x_i^{\text{mod}}|}{2\sigma_i^{\text{obs}}} \tag{2}$$

$$f_{\text{total}} = \frac{1}{N}\sum_{i=1}^{N} f_i \tag{3}$$

where N is equal to the total number of variables, x_i^{mod} represents the modelled value of
variable i predicted at the best-fit solution, x_i^{obs} represents the observed value of variable i,
and σ_i^{obs} represents the observed standard deviation for variable i. A value of > 1 for both
the total score and the individual variable scores indicates a poor fit. A total score of ≤ 1
demonstrates that the inversion fits the data acceptably well, and an individual variable score
of ≤ 1 indicates that the modelled prediction fits the observations within 2 standard deviations.
The workflow provides options to visualize this relationship (e.g., Figure 3d). The scoring
strategy above provides a similar function to the σ_{fit} metric (square-root of the mean square
weighted deviation) used in avPT (Powell & Holland, 1994).

2.3.2   Influence ($f_{\text{total}(i)}$)

Influence analysis quantifies the impact of individual variables on the inversion result. Certain variables can exert strong leverage on the best-fit solution because of their behavior or scale
of composition range in model space. To identify whether a variable exerts disproportionate control, we compute a leave-one-out total fit, obtained by recalculating the best-fit solution after
excluding variable i and calculating the new total misfit value ($f_{\text{total}(i)}$). This metric demonstrates, for each variable, whether it disproportionately affects the solution and its exclusion
improves overall inversion performance. The influence analysis provides a similar function to
the dimensionless hat value (also called the leverage value) used to quantify the influence of

7

181 individual end-members on the final result in avPT (Powell & Holland, 1994).

2.3.3 Sensitivity (ΔP_i, ΔT_i)

183   Whereas influence analysis quantifies the effect of removing a variable on the best-fit solu-
184 tion, sensitivity analysis estimates how uncertainty in the value of a variable affects the inversion
185 result. This is performed using bootstrap resampling, in which each variable is resampled in
186 turn while the remaining variables are fixed at the appropriate mean value (Figure 2c). The
187 resultant spread in the best-fit solutions thus provides a measure of the inversion's sensitivity
188 to the uncertainty in the resampled variable, and the precision with which it constrains model
189 space. The sensitivity values are given as the maximum absolute temperature (ΔT_i) and pressure
190 (ΔP_i) difference between the mean solution of the inversion when all the variables are re-sampled
191 ($T^{\mathrm{mean}}, P^{\mathrm{mean}}$) and the range of solutions when only a single variable, i, is re-sampled N times.
192 These are shown in Equations 4–5:

$$T_i^{\mathrm{max/min}} = \bar{T}_i \pm 2\,\sigma_{T_i} \tag{4}$$

193

$$\Delta T_i = \max \left| (T^{\mathrm{mean}} - T_i^{min}), (T_i^{max} - T^{mean}) \right| \tag{5}$$

194 where $\bar{T}_i$ represents the mean best-fit temperature from the re-sampling of variable i and σ_{T_i}
195 represents the standard deviation temperature for variable i. The same equations apply for
196 pressure.

197   The comparison is made with the mean solution rather than the median solution because
198 the mean is less sensitive to the discretization of the model grid. A high sensitivity may be
199 caused by a large spread in the value of observed data variables, or a large percentage change in
200 the model values over small areas of P-T space (meaning that small differences in input value
201 results in a large change in pressure or temperature estimate). Sensitivity values are displayed
202 using tornado plots (e.g., Figure 3e).

2.4 Applying the workflow

2.4.1 Observations setup

205   Any variable type can be used in the inversion, provided an appropriate forward model can
206 be calculated, and the data can be tabulated into grid format. The type and choice of variables

(e.g., mineral compositions and/or modal proportions) used should consider the interpreted scale of chemical and textural equilibrium, the method of bulk composition acquisition, the geological problem, model precision, and the desired sensitivity of individual variables. Modal proportion variables may be used for major phases but should be applied cautiously for minor or accessory phases (Weller *et al.*, 2024). Compositional variables can be expressed as full element concentrations (apfu) or as derived compositional variables (e.g., end-members, ratios, and substitution vectors). Using apfu retains full compositional information and allows more direct uncertainty assessment, whereas derived compositional variables can reduce sensitivity to systematic model offsets where applied appropriately. Table 1 lists some recommended compositional variables for pelitic and metabasic lithologies in sub-solidus systems.

When selecting variables for inversion, maximize the number of constraints by including as many robust variables as possible from the widest range of phases interpreted to be in mutual equilibrium. At least one degree of of freedom should be preserved for each set of non-independent variables. For example, for plagioclase, including either X_{An} or X_{Ab} is sufficient—both should not be used simultaneously. For each major phase within the model, including at least one variable will ensure that the misfit function accommodates its presence, and for those with solid solution models, the major substitutions should be considered in the choice of variables. This ensures that the derived P-T uncertainty estimate reflects the largest proportion of the system. Zoned minerals such as garnet are a possible exception, as core compositions may be interrogated independently where appropriate. Note that the inferred equilibrium assemblage field is not inherently enforced as a constraint unless it is used explicitly as a variable or stable phases are assigned a non-zero mode (or vice-versa). The advantages and limitations of this choice are discussed in Section 4.2.

If the bulk rock composition varies significantly across the analyzed volume, the associated uncertainty can be quantified by computing a suite of forward models generated from Monte Carlo variations of the input composition used for forward modelling. The composition range may be defined from multiple analyses from different sample domains (e.g., Palin *et al.*, 2016; Duesterhoeft & Lanari, 2020), or a range characteristic of the observed scale of variation (e.g., 5%; Forshaw *et al.*, 2019).

2.4.2 Predictions setup

The P-T grid should be broad enough to avoid boundary effects influencing the best-fit solutions, whilst maintaining sufficient resolution to capture the true solution between grid points. The spacing should reflect the required precision of the problem, the precision of the input data, the model parameters, and the model's sensitivity. If results converge on the boundary of the P-T range, a different range should be chosen if possible.

The method of bootstrap resampling should be chosen based on the data distribution or availability. Parametric bootstrapping, using a mean and standard deviation, is suitable for normally distributed data, while non-parametric bootstrapping, using the original data, is more appropriate for non-normal distributions. A minimum of 1000 resamples is recommended, though the required number can be evaluated by monitoring how the distribution of P and T solutions stabilizes with increasing resamples (Figure S1).

2.4.3 Solution assessment

Prior to inversion, the relationships among variables in P-T space should be interrogated to assess whether mutual equilibrium is likely for the selected variables for the range of values chosen for the modelling parameters (e.g., P-T range, $X_{Fe^{3+}}$, aH_2O). This can be visualized for all variables (e.g., Figure 3a) or for individual phases or variables (e.g., Figure 3b). A preliminary assessment can identify inconsistency between variables that may stem from clear geological or model uncertainty. However, in most cases it is preferable to let the variables be removed through the diagnostics.

Fit diagnostics enable critical assessment of the reliability of inversion results (Table 2). If the total score of the quality of data fit (f_{total}) is above 1, variables with fit scores (f_i) exceeding 1 should be removed sequentially, starting with the most severe, until f_{total} falls below or is equal to 1. Further removal is generally unnecessary and not advised, as the L1-norm objective function is intrinsically robust to outliers (Claerbout & Muir, 1973; Li *et al.*, 2015; Ibraheem *et al.*, 2021), and excessive pruning risks discarding meaningful information. In cases where no single outlier dominates but many variables contribute moderate misfit, the leave-one-out score ($f_{\text{total}(i)}$) may be used to identify the variable whose exclusion most effectively reduces f_{total} (lowest $f_{\text{total}(i)}$). The user may also examine the dispersion of the leave-one-out best-fit solutions to assess which variables, when removed, decrease dispersion and increase stability of

10

the inversion. However, only variables that exhibit poor fit ($f_i > 1$) should be considered for removal in these cases. The above steps may then be followed. If widespread scatter persists despite these measures, it may reflect disequilibrium or retrogression (Powell & Holland, 1994); and sequential removal is unlikely to resolve the issue. High sensitivity values (ΔP_i, ΔT_i) are not inherently problematic, but when they coincide with poor variable fit ($f_i > 1$), they may justify exclusion of the offending variables.

Once a satisfactory fit has been obtained, the best-fit solution and its uncertainty should ideally overlap the interpreted equilibrium assemblage field. Discrepancies may indicate inconsistencies between the analyzed rock volumes, the scale of equilibrium, or the definition of the equilibrium assemblage. Bootstrap resampling can reveal whether the uncertainty distribution is unimodal or multi-modal. The distribution of the solution uncertainty can be assessed using histograms or the heatmap of solutions (Figure 3c). A multi-modal distribution suggests that additional observations or external constraints (e.g., single-equilibrium thermobarometers) may be needed to discriminate among competing solutions. Significant local minima will increase the estimated uncertainty, reflected in a higher interquartile range. Final results should generally be reported using the median and inter quartile range (IQR), which are suitable for non-normal distributions. Other recommendations for result reporting are included in the Supplementary text.

3   NATURAL EXAMPLES

To illustrate the effectiveness of the workflow, we present results from analyses of a pelitic and a metabasic rock, demonstrating its performance across different model systems, solution models, and sets of variables. Specifically, we present a kyanite-staurolite schist (ICSV13) and a garnet amphibolite (ICSV117) from the Greater Himalayan Sequence in the Zanskar Himalaya, NW India. A characteristic photomicrograph of each sample is found in Figure S2, and detailed petrographic descriptions are included in the Supplementary text. Mineral compositions for each sample were inverted to determine best-fit P–T conditions and compared with classical thermobarometric methods. Details of the analytical setup, mineral recalculation procedure, formulation of bulk composition, and forward modelling procedure are included in the Supplementary Text. The bulk rock composition for each sample, measured by X-ray fluorescence (XRF) under the assumption of chemical equilibrium at the rock scale, is included in Table 3.

295 The mineral composition input and forward model for each sample are included in the Supple-
296 mentary material (Tables S1–S4).

3.1 ICSV13

3.1.1 Conventional methods of analysis

299 The inferred equilibrium assemblage at peak conditions for sample ICSV13 is kyanite–
300 staurolite–garnet–biotite–muscovite–plagioclase–quartz–rutile–ilmenite–H_2O. Although the
301 kyanite- and staurolite-present field directly represents the peak assemblage of all major phases,
302 the narrow temperature range is at odds with the regular occurrence of this assemblage across
303 numerous metamorphic terrains and wider temperature conditions (Pattison & Spear, 2018).
304 This discrepancy likely reflects either sluggish reaction kinetics that yield the metastable
305 persistence of staurolite, or the small free energy difference between staurolite and kyanite
306 nucleation that enables kyanite to nucleate earlier, or staurolite later, than that predicted
307 by the thermodynamic models (Pattison & Spear, 2018; Cawood, 2024). Therefore, the peak
308 assemblage could also be assigned to either the staurolite-present or the kyanite-present field
309 (Figure 4a). In this case, the predominance of nearby assemblages being kyanite-bearing rather
310 than staurolite- or kyanite + staurolite-bearing (Cawood *et al.*, 2025), favors a kyanite-present
311 field as the most likely choice for the peak assemblage field.

312 The peak assemblage field shows a wide *P-T* range over which the given rock composition
313 may have equilibrated. Temperature ranges from ~600 °C to in excess of the upper modelled
314 temperature (725 °C) and pressure is poorly constrained owing to the dependence on the Ti-rich
315 phases to define the phase boundaries, which are avoided here as a primary *P-T* constraint (e.g.,
316 Starr *et al.*, 2020). Consequently, the broad range provided by the peak field in this sample is well
317 suited to applying the workflow to further constrain the conditions of metamorphism. Using the
318 Ti-in-biotite geothermometer calibration of Henry *et al.* (2005), the metamorphic temperature
319 is estimated at 594 ± 24 °C. At this temperature, pressure—calculated with the av*P* function of
320 THERMOCALC (Powell & Holland, 1994)—is constrained to 9.3 ± 0.9 kbar (1 σ), consistent
321 with the peak assemblage field (Figure 4a).

3.1.2   Grid-search inversion and uncertainty analysis

The workflow was performed on a 100 x 100 grid in P-T space from 425–725 °C and 5–13 kbar with 1000 random sets of samples drawn from the mineral composition data using bootstrap re-sampling assuming a normal distribution. The salient standard deviations and mean values for each mineral composition variable are provided in Table 4. Based on the interpreted peak mineral assemblage, 11 variables were used for this analysis: X_{Grs}, X_{Sps}, X_{Prp}, X_{Mg} of garnet, Si, Ti, and X_{Mg} in biotite, Si content, X_{Pa}, and X_{Cel} of muscovite, and X_{Ab} of plagioclase (see Table 1 for definitions). Staurolite and kyanite were not used as variables (the latter having no solution model) to avoid biasing the modelled stability fields of these phases. The P-T results are reported as median and IQR.

The mineral compositional variables produce a best-fit solution of 598 °C (IQR 592–604 °C) and 9.61 kbar (IQR 9.28–9.93 kbar; Figure 4c). This result and associated uncertainty show excellent agreement with the independent av$P$ and Ti-in-biotite results (thermobarometer intersection) and overlap with the kyanite-present peak assemblage field (Figure 4b). During solution refinement, X_{Mg} in biotite and X_{Cel} were sequentially removed, the former due to poor fit ($f_{\mathrm{total}} > 1$), and the later due to poor fit coupled with strong influence on the solution (low $f_{\mathrm{total}(i)}$). For the remaining nine variables, the residuals show the best-fit solution successfully fits seven within the 2σ range of the observed data (Figure 3d). X_{Prp} and X_{Mg} have poor fit, but the solution has low sensitivity to these variables. The best-fit solution is particularly sensitive to X_{Sps} and X_{Grs} with a ΔP_i of 1.58 kbar and 0.49 kbar, respectively (Table 5). The high sensitivity to X_{Sps} is consistent with the close correlation of the end-member value with the proportion of garnet — a result of Mn mass-balance in a system where garnet is the dominant Mn-host (Spear, 1993; Waters, 2019).

The inversion shows that the natural data reproduce conditions most consistent with peak metamorphism in the modelled kyanite-present field, though the uncertainty distribution indicates that the staurolite-present or more temperature-restricted kyanite+staurolite fields are also viable. Given the restricted P-T range defined by the overlap of the uncertainty distribution and shared field boundaries, all solutions are broadly consistent with the microstructure and shared inclusion suite for both minerals that imply coeval growth under similar conditions (Figure S2a; Cawood *et al.*, 2025). More generally, the inversion framework demonstrates its value for precisely such cases, where ambiguities between multiple plausible assemblage fields

13

353 can be tested and narrowed using the derived uncertainty distribution.

3.2   ICSV117

3.2.1   Conventional methods of analysis

356 The inferred equilibrium assemblage at peak conditions for sample ICSV117 is: garnet–
357 hornblende–biotite–plagioclase–quartz–ilmenite–titanite–H_2O. Petrographic observations show
358 that titanite growth dominantly post-dates ilmenite growth, however it is unclear whether ti-
359 tanite represents peak conditions or growth during decompression from peak conditions. In
360 either scenario, the relationship of Ti-bearing phases is not reproduced in a clockwise P-T path
361 on the modelled diagram, therefore as applied for ICSV13, the stability fields of the Ti-rich
362 phases are not prioritised. Furthermore, the modelled presence or absence of clinopyroxene
363 is not deemed relevant to phase boundaries in clinopyroxene-absent assemblages owing to its
364 predicted overstability in modelled metabasic systems (Forshaw *et al.*, 2019). Therefore, the
365 upper-pressure boundary of the peak assemblage field is delimited by muscovite at higher pres-
366 sures and the absence of garnet at lower pressures. The peak assemblage field extends for
367 a greater range of temperatures than modelled (620–750 °C) and from ~10.0 kbar to higher
368 pressures than modelled (> 14 kbar; Figure 4f). Using the hornblende-plagioclase exchange
369 geothermometer calibration of Holland & Blundy (1994), the peak temperature is constrained
370 to 644 $\pm$ 40 °C, and the pressure, using the av$P$ function of THERMOCALC at the matching
371 temperature, to a value of 10.8 $\pm$ 1.2 kbar (1 σ). These results show excellent agreement with
372 the peak assemblage field (Figure 4f).

3.2.2   Grid-search inversion and uncertainty analysis

374 The inversion workflow was performed on a 100 x 100 grid in P-T space from 550–750 °C
375 and 8–14 kbar with 1000 random sets of samples drawn from the mineral composition data
376 using bootstrap re-sampling assuming a normal distribution. The salient standard deviations
377 and mean values for each mineral composition variable are provided in Table 4. Based on the
378 interpreted peak mineral assemblage, 11 variables were used for this analysis: X_{Grs}, X_{Prp}, X_{Mg}
379 of garnet; Ti, Si and X_{Mg} in biotite; X_{Ab} of plagioclase; Ts (tschermakite, Al(T)-Na(A)-K(A)),
380 Ed (edenite, Na(A)+K(A)), Gln (glaucophane, Na(M4)) vectors and Ti content of amphibole
381 (see Table 1 for variable definitions). Amphibole vectors were used to reduce dependency of
382 the result on overestimated (Al and A-site Na) or underestimated (Si, Ca and A-site K) cations

383 (Forshaw *et al.*, 2019).

384 The mineral compositional variables produce a median best-fit solution at 667 °C (IQR =
385 661–673 °C) and 11.76 kbar (IQR = 11.64–11.94 kbar; Figure 4g, h; Table 6), in agreement with
386 the thermobarometer intersection and peak assemblage field. Si in biotite, X_{Grs}, and the Ed
387 vector were sequentially removed due to poor fit ($f_i > 1$; Table 6). For the final eight variables,
388 the residuals show the best-fit solution successfully fits five within the 2σ range of the observed
389 data (Figure S3a). X_{Ab} and Ti in amphibole and the Ts vector show poor fit, but the inversion
390 has low sensitivity to these variables (Table 6, Figure S3b). The best-fit temperature is most
391 sensitive to X_{Prp} and Ti in Bt (ΔT_i of 10 and 14 °C respectively) and pressure is sensitive to the
392 Gln vector and Ti in biotite, with ΔP_i values of 0.47 and 0.21 kbar, respectively (Figure S3b).
393 Although the Ti in biotite variable is relatively pressure insensitive, the value determines the
394 intersection point with variables which are more pressure sensitive (i.e., sloping P-T gradients),
395 and therefore has an important role to play in determining the pressure of the solution (Figure
396 4j).

397 Initial thermobarometric analyses were performed for ICSV117 assuming a garnet-absent
398 assemblage, based on the lack of garnet observed at outcrop, hand-sample, and thin-section
399 scales. Given the sample's location within a garnet-bearing metamorphic zone of the metaba-
400 sic sequence (Cawood *et al.*, 2025), this absence was attributed to subtle bulk compositional
401 variation, as observed in other samples from the region. However, the inversion consistently re-
402 turned best-fit solutions within the garnet-bearing stability field across varied parameter ranges,
403 contradicting the a priori assemblage. Subsequent targeted reconnaissance identified accessory
404 garnet grains ranging from 140–180 μm, validating that the inversion result was resilient despite
405 the initial misclassification of the equilibrium assemblage. This outcome aligns with prior find-
406 ings that have demonstrated good agreement between predicted and observed garnet modes in
407 metabasic systems (Forshaw *et al.*, 2019). As previously outlined, the space in which the misfit
408 function identifies a best-fit solution is inherently governed by the modelled stability ranges of
409 the phases of the chosen variables. Consequently, phases (and their variables) that are either not
410 included in the inversion or not present in the observed assemblage do not necessarily impact the
411 solution. This underscores the need to critically evaluate how the best-fit result relates to the in-
412 dependently constrained peak assemblage field. We note that such resilience is case-specific and
413 in this situation was contingent on the acquisition of whole-rock XRF analyses for the bulk rock

15

composition; erroneous assemblage assumptions could have more serious consequences where effective or reactive bulk compositions are used.

4 DISCUSSION

4.1 Comparison to other quantitative workflows

Using petrological data to quantitatively invert for pressure and temperature is well established. The avPT function of THERMOCALC (Powell & Holland, 1994) uses the activities of end-members involved in balanced independent chemical equilibria from an internally consistent thermodynamic dataset to calculate the temperature and/or pressure. An iterative, least-squares inversion is used to find the optimal P-T condition which maximises the consistency between each independent chemical equilibrium involved within the inversion (Powell & Holland, 1994). The same approach applies for av$P$ and av$T$ calculations, although these are calculated directly without iteration (Powell & Holland, 1988). This approach calculates uncertainties stemming from the model, in contrast to observational uncertainties derived from the mineral measurements considered in this method. TWEEQU (Thermobarometry With Estimation of EQUilibration state) follows a similar multi-equilibrium approach, calculating all possible equilibria for the selected phase components from an internally consistent thermodynamic database, with uncertainties estimated from the weighted scatter of equilibria intersections (Berman, 1991). The subsequent winTWQ refined this by emphasizing a more robust independent set of equilibria (Berman, 2007). Whilst these methods are powerful and have been widely used, they remain independent from phase equilibrium modelling.

Bingo-Antidote uses the forward models calculated by Theriak-Domino to perform a global optimization of P and T (Duesterhoeft & Lanari, 2020) for the observed phase assemblage, mineral modal proportions, and mineral compositions. Bingo-Antidote is particularly powerful due to its integration with XMapTools (Lanari *et al.*, 2019) and thus its ability to link directly to compositional phase maps, and examine the impact of local bulk composition variation on resulting P-T estimates. IntersecT (Nerone *et al.*, 2025) builds on a similar principle, using outputs from Perple_X to provide a quantitative framework for evaluating the quality of fit between model-predicted isopleths and measured mineral compositions. Some key differences between these workflows and the method proposed here are discussed in the Supplementary Text. Most notably, uncertainty is estimated quite differently. In Bingo-Antidote, measurement uncertain-

443 ties are incorporated into its definition of the objective function. Lower values of their objective
444 function (quality factor, Q_{total}) indicate that the inversion is closer to the mean observed value
445 for any given data variable. Uncertainty is calculated as the region of P–T space yielding Q_{total}
446 scores within 2% of the optimum. This can be thought of the goodness-of-fit error, which quan-
447 tifies how well different potential solutions describe the mean value for each variable. IntersecT
448 combines the quality factor metric of Bingo-Antidote for compositional variables (Q_{cmp}) with
449 a reduced χ^2 statistic, which explicitly weights each phase according to how well its measured
450 composition is reproduced within its observed uncertainty. In contrast, LinaForma does not
451 define uncertainty from a single optimum solution. Instead, it estimates uncertainties through
452 bootstrap resampling of mineral measurements, generating multiple realizations to produce a
453 distribution of possible P–T solutions. This method emphasises how uncertainties in the under-
454 lying data translate into variability in the inferred conditions.

455 ## 4.2 Advantages and limitations of the new workflow

456 **Grid-search inversion** The grid-search inversion offers a systematic approach to identify the
457 global minimum of the misfit function and determine the best-fit solution within the selected
458 model space. By discretizing P-T space into a grid of forward models and calculating the misfit at
459 each grid point (Figure 2a), this method ensures the global minimum is located, provided the grid
460 range and precision are appropriate. The grid-search is also advantageous because it is adaptable
461 to various combinations of model parameters, such as T-X and P-X, making it applicable
462 across a broad range of petrological modelling scenarios. Furthermore, the workflow is equally
463 applicable to higher-dimensional analyses (e.g., P-T-X). However, the grid-search approach
464 introduces computational challenges as the number of trial solutions increases exponentially with
465 increasing dimensionality. Recent advances in forward modelling software (e.g., Riel *et al.*, 2022)
466 have significantly reduced the computational time required for these calculations, allowing the
467 extension of this workflow to higher dimensions and reducing the need for assigning parameters
468 a priori (e.g., $X_{Fe^{3+}}$, aH_2O).

469 **Misfit/objective function** The workflow employs the L1-norm misfit function, a robust
470 measure of fit that is widely used across scientific disciplines because of its resilience to outliers,
471 outperforming L2-norm (least-squares) or chi-squared functions in this regard (Claerbout &
472 Muir, 1973; Li *et al.*, 2015; Ibraheem *et al.*, 2021). This makes the method particularly well-

suited for petrological systems where significant outliers are common. In addition, the leave-one-out calculation allows the user to diagnose variables that disproportionately affect the inversion and ensure that high-leverage and poorly-fitting variables do not bias the final solution.

The inversion does not directly resolve or use the stable phase assemblage unless the variables are specifically defined to do so, and can therefore also be applied with caution in metastable systems. The misfit function seeks solutions in which all phases included in the selected variables are of the correct composition and/or modal proportion, thereby defined by where these phases are predicted to be stable. Therefore, the boundaries of the result are inherently dependent on the variables employed in the inversion. The lack of strict relationship between the misfit function and a user-defined equilibrium assemblage field has several benefits: (1) The assemblage field corresponding to a mineral assemblage in a pseudosection may be smaller than the estimated uncertainty (Powell & Holland, 2008; Waters, 2019); (2) field boundaries defined by small modal proportion or accessory phases may not be reliable (Weller *et al.*, 2024); and (3) assemblage field boundaries, mineral modal proportion contours, and mineral compositional isopleths show relatively decreasing uncertainties and increasing precision (Waters, 2019). However, where mismatch occurs between an inversion result and the inferred equilibrium assemblage field, this should in all cases be critically assessed.

The misfit function does not inherently penalize inversions using a small number of variables. However, the number of variables will have a significant impact on the P-T uncertainty derived from bootstrap resampling, discussed below. A notable feature of this workflow is its deliberate avoidance of variable weighting, such that all variables are treated as equally valid in the inversion. This approach is advantageous in scenarios where the relative importance of variables is difficult to determine or where weighting would introduce significant bias into the results. A limitation of this approach is that highly influential variables can disproportionately affect the solution; however, this is mitigated by quantifying variable leverage with the influence analysis.

Uncertainty analysis Monte Carlo methods, such as bootstrap resampling, are particularly well suited to estimating uncertainty in non-linear problems such as P-T inversion, where the combined effects of random and systematic errors cannot be calculated directly (Menke, 1984). The application of bootstrap resampling to the observed data allows users to empirically estimate the variability, or uncertainty, of the inverse solution (Figure 2b). It may also be used to

504 identify any significant local minima that may provide alternative hypotheses and to evaluate
505 the sensitivity of the final result to the different input variables (Figure 3e).

506   Although the workflow incorporates multiple sources of uncertainty stemming from the
507 observations, not all uncertainties are fully addressed. Uncertainty in the parameters used during
508 modelling and the underlying thermodynamic end-member data sets and a-X relations are not
509 directly examined by bootstrap resampling. Nevertheless, the method inherently incorporates
510 some of these uncertainties due to reliance on Monte Carlo-style methods. With the reduced
511 computational times of forward modelling (e.g., Riel *et al.*, 2022), it is now possible to include
512 Monte Carlo simulations of the forward model itself, systematically varying input parameters,
513 solution models and end-member properties to evaluate the influence of model uncertainty on
514 the resulting P-T solutions.

5   CONCLUSIONS

515   Accurately quantifying the petrological parameters of metamorphic rocks is crucial for un-
516 derstanding a broad range of processes in both the solid and surficial Earth, including identify-
517 ing geothermal gradients in the subsurface and sources of economically valuable raw materials
518 (e.g., critical metals), understanding climate and Earth system feedbacks, reconstructing the
519 tectonometamorphic evolution of terrains, and informing broader geodynamic models. The
520 ability to understand and quantify uncertainties in derived model parameters, such as P-T con-
521 ditions, is essential to ensuring that results can be interpreted within a defined confidence range.
522 This is particularly important in scenarios where geological or petrological interpretations hinge
523 on relatively small variations in pressure and/or temperature (e.g., Pattison & DeBuhr, 2015).
524 Although many sources of uncertainty in phase equilibrium modelling cannot be estimated or
525 are difficult to quantify, for those that can be, there should be an attempt to do so (Powell &
526 Holland, 2008). The key conclusions of this study are:

527 1. The new workflow, LinaForma, provides quantitative constraints for best-fit P-T con-
528   ditions (or other petrological parameters) and associated uncertainties for a given rock
529   system. The technique identifies the best-fit solution by comparing observed data, such as
530   mineral compositions and/or modal proportions, with forward model predictions across a
531   user-defined P-T grid. Bootstrap resampling (repeated sampling with replacement) quan-
532   tifies the uncertainty of the inverse solution and assesses its sensitivity to input variable

uncertainty. Three diagnostic metrics—quality of data fit, variable influence, and variable sensitivity—are incorporated to validate and refine the results.

2. Application to natural samples demonstrates that peak P-T conditions and uncertainty estimates for both pelitic and metabasic compositions align with classical thermobarometric methods. Local discrepancies help assess model limitations and refine prior interpretations, such as the equilibrium assemblage.

3. The workflow is compatible with any forward-modelling software, supports flexible variable selection, performs a systematic grid-search inversion in multidimensional parameter space, and uses a robust L1-norm misfit function that is relatively resistant to outliers, while also providing uncertainty and sensitivity analysis through bootstrap resampling. Its main limitations are the high computational cost of exploring grids in more than two or three dimensions, and the lack of explicit quantification of uncertainties inherited from the underlying thermodynamic end-member data set and solution models.

SOFTWARE AVAILABILITY

The outlined workflow, LinaForma, is available via the GitHub respository https://github.com/quantPT/LinaForma.

ACKNOWLEDGEMENTS

REFERENCES

Berman, R. G. (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(2), 445–522, doi:10.1093/petrology/29.2.445.

Berman, R. G. (1991), Thermobarometry using multi-equilibrium calculations; a new technique, with petrological applications, *The Canadian Mineralogist*, 29, 833–855.

Berman, R. G. (2007), winTWQ (version 2.3): A software package for performing internally-consistent thermobarometric calculations, Geological Survey of Canada, Open File, 5462.

Cawood, I. P. (2024), Structural and metamorphic evolution of the Zanskar Himalaya, Suru Valley region, NW India, Ph.D. thesis, University of Oxford.

Cawood, I. P., M. R. St-Onge, O. M. Weller, M. P. Searle, D. J. Waters, & T. Ahmad (2025), Structural and metamorphic architecture of the Zanskar Himalaya, Suru Valley region, NW India: Implications for the evolution of the Himalayan metamorphic core, *GSA Bulletin*, doi:10.1130/B37241.1.

Claerbout, J. F., & F. Muir (1973), Robust modeling with erratic data, *Geophysics*, 38(5), 826–844, doi:10.1190/1.1440378.

Connolly, J. (1990), Multivariable phase diagrams: An algorithm based on generalized thermodynamics, *American Journal of Science*, 290, 666–718, doi:10.2475/ajs.290.6.666.

Connolly, J. (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(1-2), 524–541, doi:10.1016/j.epsl.2005.04.033.

de Capitani, C., & T. H. Brown (1987), The computation of chemical equilibrium in complex systems containing non-ideal solutions, *Geochimica et Cosmochimica Acta*, 51(10), 2639–2652, doi:10.1016/0016-7037(87)90145-1.

de Capitani, C., & K. Petrakakis (2010), The computation of equilibrium assemblage diagrams with Theriak/Domino software, *American Mineralogist*, 95(7), 1006–1016, doi:10.2138/am.2010.3354.

Duesterhoeft, E., & P. Lanari (2020), Iterative thermodynamic modelling—Part 1: A theoret-

584   ical scoring technique and a computer program (Bingo-Antidote), *Journal of Metamorphic*
585   *Geology*, 38(5), 527–551, doi:10.1111/jmg.12538.

586 Efron, B. (1979), Bootstrap methods: Another look at the jackknife, *The Annals of Statistics*,
587   7(1), 1–26, doi:10.1007/978-1-4612-4380-9_41.

588 Forshaw, J. B., D. J. Waters, D. R. M. Pattison, R. M. Palin, & P. Gopon (2019), A comparison
589   of observed and thermodynamically predicted phase equilibria and mineral compositions in
590   mafic granulites, *Journal of Metamorphic Geology*, 37(2), 153–179, doi:10.1111/jmg.12454.

591 Gordon, T. M. (1992), Generalized thermobarometry: Solution of the inverse chemical equi-
592   librium problem using data for individual species, *Geochimica et Cosmochimica Acta*, 56(5),
593   1793–1800, doi:10.1016/0016-7037(92)90310-F.

594 Henry, D. J., C. V. Guidotti, & J. A. Thomson (2005), The Ti-saturation surface for low-to-
595   medium pressure metapelitic biotites : Implications for geothermometry and Ti-substitution
596   mechanisms, *American Mineralogist*, 90, 316–328, doi:10.2138/am.2005.1498.

597 Holland, T., & J. Blundy (1994), Non-ideal interactions in calcic amphiboles and their bearing
598   on amphibole-plagioclase thermometry, *Contributions to Mineralogy and Petrology*, 116(4),
599   433–447, doi:10.1007/BF00310910.

600 Holland, T. J. B., & R. Powell (1998), An internally consistent thermodynamic data set for
601   phases of petrological interest, *Journal of Metamorphic Geology*, 16(3), 309–343, doi:10.1111/
602   j.1525-1314.1998.00140.x.

603 Hoschek, G. (2004), Comparison of calculated P-T pseudosections for a kyanite eclogite from
604   the Tauern Window, Eastern Alps, Austria, *European Journal of Mineralogy*, 16(1), 59–72,
605   doi:10.1127/0935-1221/2004/0016-0059.

606 Ibraheem, I. M., B. Tezkan, & R. Bergers (2021), Integrated interpretation of magnetic and
607   ERT data to characterize a landfill in the north-west of Cologne, Germany, *Pure and Applied*
608   *Geophysics*, 178(6), 2127–2148, doi:10.1007/s00024-021-02750-x.

609 Lanari, P., S. Ferrero, P. Goncalves, & E. G. Grosch (2019), Metamorphic geology: Progress
610   and perspectives, *Geological Society, London, Special Publications*, 478(1), 1–12, doi:10.1144/
611   SP478-2018-186.

612 Li, C.-N., Y.-H. Shao, & N.-Y. Deng (2015), Robust L1-norm two-dimensional linear discrimi-
613   nant analysis, *Neural Networks*, 65, 92–104, doi:10.1016/j.neunet.2015.01.003.

614 Menke, W. (1984), Geophysical Data Analysis: Discrete Inverse Theory, Academic Press, Or-
615   lando.

616 Nerone, S., P. Lanari, H. Dominguez, J. B. Forshaw, C. Groppo, & F. Rolfo (2025), IntersecT:
617   A Python script for quantitative isopleth thermobarometry of equilibrium and disequilibrium
618   systems, *Computers & Geosciences*, 202, 105,949, doi:10.1016/j.cageo.2025.105949.

619 Palin, R. M., O. M. Weller, D. J. Waters, & B. Dyck (2016), Quantifying geological uncertainty
620   in metamorphic phase equilibria modelling; a Monte Carlo assessment and implications for
621   tectonic interpretations, *Geoscience Frontiers*, 7(4), 591–607, doi:10.1016/j.gsf.2015.08.005.

622 Pattison, D. R. M., & C. L. DeBuhr (2015), Petrology of metapelites in the Bugaboo aureole,
623   British Columbia, Canada, *Journal of Metamorphic Geology*, 33(5), 437–462, doi:10.1111/
624   jmg.12128.

625 Pattison, D. R. M., & F. S. Spear (2018), Kinetic control of staurolite–Al_2SiO_5 mineral as-
626   semblages: Implications for Barrovian and Buchan metamorphism, *Journal of Metamorphic
627   Geology*, 36(6), 667–690, doi:10.1111/jmg.12302.

628 Powell, R., & T. Holland (1988), An internally consistent dataset with uncertainties and corre-
629   lations: 3. Applications to geobarometry, worked examples and a computer program, *Journal
630   of Metamorphic Geology*, 6(2), 173–204, doi:10.1111/j.1525-1314.1988.tb00415.x.

631 Powell, R., & T. Holland (1994), Optimal geothermometry and geobarometry, *American Min-
632   eralogist*, 79, 120–133.

633 Powell, R., & T. J. B. Holland (2008), On thermobarometry, *Journal of Metamorphic Geology*,
634   26(2), 155–179, doi:10.1111/j.1525-1314.2007.00756.x.

635 Riel, N., B. J. P. Kaus, E. C. R. Green, & N. Berlie (2022), MAGEMin, an efficient Gibbs
636   energy minimizer: Application to igneous systems, *Geochemistry, Geophysics, Geosystems*,
637   23(7), {e2022GC010,427}, doi:10.1029/2022GC010427.

638 Spear, F. S. (1993), Metamorphic Phase Equilibria And Presure-Temperature-Time-Paths,
639   Monograph Series, Mineralogical Society of America, Washington, D.C.

640 Spear, F. S., D. R. Pattison, & J. T. Cheney (2016), The metamorphosis of metamorphic petrol-

641 ogy, in The Web of Geological Sciences: Advances, Impacts, and Interactions II: Geological
642 Society of America Special Paper 523, edited by M. Bickford, The Geological Society of Amer-
643 ica, doi:10.1130/2016.2523(02).

644 Starr, P. G., D. R. M. Pattison, & D. E. Ames (2020), Mineral assemblages and phase equilibria
645 of metabasites from the prehnite–pumpellyite to amphibolite facies, with the Flin Flon Green-
646 stone Belt (Manitoba) as a type example, *Journal of Metamorphic Geology*, 38(1), 71–102,
647 doi:10.1111/jmg.12513.

648 Štípská, P., & R. Powell (2005), Constraining the P–T path of a MORB-type eclogite using
649 pseudosections, garnet zoning and garnet-clinopyroxene thermometry: An example from the
650 Bohemian Massif, *Journal of Metamorphic Geology*, 23(8), 725–743, doi:10.1111/j.1525-1314.
651 2005.00607.x.

652 Vance, D., & E. Mahar (1998), Pressure-temperature paths from P-T pseudosections and zoned
653 garnets: Potential, limitations and examples from the Zanskar Himalaya, NW India, *Contri-*
654 *butions to Mineralogy and Petrology*, 132, 225–245.

655 Waters, D. J. (2019), Metamorphic constraints on the tectonic evolution of the High Himalaya in
656 Nepal: The art of the possible, *Geological Society London, Special Publication*, 483, 325–375,
657 doi:10.1144/sp483-2018-187.

658 Weller, O. M., T. J. B. Holland, C. R. Soderman, E. C. R. Green, R. Powell, C. D. Beard, &
659 N. Riel (2024), New thermodynamic models for anhydrous alkaline-silicate magmatic systems,
660 *Journal of Petrology*, 65(10), {egae098}, doi:10.1093/petrology/egae098.

FIGURE CAPTIONS

661 Figure 1: Schematic summary of (a–d) common problems associated with application of inter-
662 secting mineral composition variables commonly applied in constraining the $P\text{-}T$ conditions from
663 forward models. (e) Using a large and diverse range of robust mineral measurement variables
664 (composition, modal proportion, or both) to provide a more representative $P\text{-}T$ constraint using
665 grid-search inversion and bootstrap resampling. Definitions of mineral composition variables are
666 included in Table 1 for reference.

667 Figure 2: Schematic summary of the new workflow and methods. From observations and forward
668 modelling of a given rock system to (a) grid-search non-linear inversion to determine the best-fit
669 $P\text{-}T$ conditions, to (b) bootstrap-resampling to assess the uncertainty of the $P\text{-}T$ solution, and
670 (c) result diagnostic using bootstrap resampling of the employed variables to estimate sensitivity
671 of the best-fit result to uncertainty in the individual variables. Variables 1, 2, and 3 are shown
672 with a normalised propability distributions, with colour representing trial or iteration number.
673 Thin-section schematic from Palin *et al.* (2016).

674 Figure 3: Workflow example using natural sample ICSV13. (a) Overlapping isopleth "fields"
675 defined by 2 σ range of mineral measurements. (b) Intersection of X_{Grs}, X_{Prp}, and X_{Sps} isopleth
676 fields defined by 2 σ range of mineral measurements. (c) Grid-search best-fit solutions and
677 heatmap (error-surface) for bootstrapped mineral measurement data (n = 1000), overlain with
678 the bootstrapped mean and median best-fit result. (d) Data residuals at the median $P\text{-}T$ point
679 of 598 °C and 9.61 kbar, showing the fit between the model and the observations. (e) Sensitivity
680 analysis for T and P of mineral composition variables presented as tornado plots at the mean
681 $P\text{-}T$ point of 597 °C and 9.82 kbar.

682 Figure 4: Example of compiled final result of ICSV13 (a–e) and ICSV117 (f–j). (a, f) Pseudo-
683 section. (b, g) Heatmap (error-surface) of best-fit solutions showing the misfit of mineral com-
684 position variables, overlain with the corresponding pseudosection and the bootstrapped mean
685 best-fit result and 1 σ uncertainty. (c, h) Boxplots of bootstrapped T and P estimates with the
686 grey bar indicating the range of the selected thermometer and av$P$ results. (d, i) Log-scaled 2D-
687 histogram bin-plot of the bootstrapped data with the mean and median best-fit result marked
688 by stars. (e, j) percent overlap of compositional isopleth fields at 2σ.

25

Table 1: Mineral composition variable definitions for common pelitic and metabasic minerals at sub-solidus conditions. *N.b.* Maintain one degree of freedom where appropriate.

Mineral	Variable	Definition
Grt	X_{Alm}	$\frac{\text{Fe}^{2+}}{\text{Fe}^{2+}+\text{Mg}+\text{Ca}+\text{Mn}}$
	X_{Grs}	$\frac{\text{Ca}}{\text{Fe}^{2+}+\text{Mg}+\text{Ca}+\text{Mn}}$
	X_{Sps}	$\frac{\text{Mn}}{\text{Fe}^{2+}+\text{Mg}+\text{Ca}+\text{Mn}}$
	X_{Prp}	$\frac{\text{Mg}}{\text{Fe}^{2+}+\text{Mg}+\text{Ca}+\text{Mn}}$
	X_{Mg}	$\frac{\text{Mg}}{\text{Mg}+\text{Fe}^{2+}}$
Grt (without Mn)	X_{Alm}	$\frac{\text{Fe}^{2+}}{\text{Fe}^{2+}+\text{Mg}+\text{Ca}}$
	X_{Grs}	$\frac{\text{Ca}}{\text{Fe}^{2+}+\text{Mg}+\text{Ca}}$
	X_{Prp}	$\frac{\text{Mg}}{\text{Fe}^{2+}+\text{Mg}+\text{Ca}}$
	X_{Mg}	$\frac{\text{Mg}}{\text{Mg}+\text{Fe}^{2+}}$
St	X_{Mg}	$\frac{\text{Mg}}{\text{Mg}+\text{Fe}^{2+}}$
Crd	X_{Mg}	$\frac{\text{Mg}}{\text{Mg}+\text{Fe}^{2+}}$
Chl	X_{Mg}	$\frac{\text{Mg}}{\text{Mg}+\text{Fe}^{2+}}$
Bt	Si	apfu
	Ti	apfu
	X_{Mg}	$\frac{\text{Mg}}{\text{Mg}+\text{Fe}^{2+}}$
Ms	Si	apfu
	X_{Cel}	Mg apfu
	X_{Pa}	$\frac{\text{Na}}{\text{Na}+\text{Ca}+\text{K}}$
Pl	X_{Ab}	$\frac{\text{Na}}{\text{Na}+\text{Ca}+\text{K}}$
Kfs	X_{San}	$\frac{\text{K}}{\text{Na}+\text{Ca}+\text{K}}$
Ep	X_{Fe}	$\frac{\text{Fe}^{3+}}{\text{Al}+\text{Fe}^{3+}}$
Amph	Ti	apfu
	Ts	$\text{Al(T)} - \text{Na(A)} - \text{K(A)}$
	Ed	$\text{Na(A)} + \text{K(A)}$
	Gln	Na(M4)
Cpx	Al	apfu
	X_{Mg}	$\frac{\text{Mg}}{\text{Mg}+\text{Fe}^{2+}}$
	Ca	apfu
	X_{Jd}	$\begin{cases} \text{Na(M2)}, & \text{if Na(M2)} < \text{Al(M1)} \\ \text{Al(M1)}, & \text{if Al(M1)} < \text{Na(M2)} \end{cases}$

Table 2: Definitions of parameters used in the inversion and values presented in the result output and diagnostics table.

Parameter	Definition
f_i	Quality of fit score for each variable. Higher values indicate a poorer fit.
$f_{\text{total}(i)}$	Leave-one-out total fit for variable i; calculated by removing variable i, recomputing the best-fit solution and finding the resulting total fit while excluding i.
ΔT_i	Maximum absolute temperature variation due to bootstrap resampling of variable i, with all other variables fixed. Measured relative to the mean best-fit solution calculated using all variables; measures sensitivity.
ΔP_i	Maximum absolute pressure variation due to bootstrap resampling of variable i, with all other variables fixed. Measured relative to the mean best-fit solution calculated using all variables; measures sensitivity.
$\mu_{\text{obs}} \pm 2\sigma$	Mean observed variable value, with two standard deviations (uncertainty).
mod	Modelled value of variable for the median best-fit solution.
mean	Mean temperature and pressure of the solution distribution (1σ).
median	50th percentile of the temperature and pressure distribution.
IQR	Interquartile range (25th–75th percentile) of the temperature and pressure distribution.
f_{total}	Overall quality of fit score for the inversion; values above 1 indicate poor fit.
grid resolution	Resolution/spacing of the grid in °C and kbar
bootstrap resamples	Number of bootstrap resamples

Table 3: Measured bulk composition via XRF of ICSV13 and ICSV117 in wt%.

Sample	SiO$_2$	Al$_2$O$_3$	Fe$_2$O$_3$	MnO	MgO	CaO	Na$_2$O	K$_2$O	TiO$_2$	P$_2$O$_5$	LOI	Total
ICSV13	61.23	19.13	7.10	0.09	2.65	0.47	1.52	4.12	0.98	0.09	1.96	99.31
ICSV117	47.80	14.42	14.92	0.19	6.39	9.99	2.41	0.75	2.12	0.18	0.34	99.51

Table 4: Variables used for parametric bootstrap resampling in ICSV13 and ICSV117 (rounded to 3 decimal places).

Mineral	Garnet				Biotite			Muscovite			Plagioclase	Amphibole			
Variable	X_{Grs}	X_{Sps}	X_{Prp}	X_{Mg}	Si	Ti	X_{Mg}	Si	X_{Cel}	X_{Pa}	X_{Ab}	Ti	Ts	Ed	Gln
ICSV13															
Mean	0.059	0.019	0.164	0.177	2.714	0.104	0.490	3.068	0.079	0.230	0.888				
SD	0.007	0.004	0.011	0.011	0.014	0.007	0.003	0.023	0.011	0.015	0.036				
ICSV117															
Mean	0.281		0.101	0.140	2.764	0.147	0.490				0.792	0.085	1.189	0.487	0.150
SD	0.013		0.009	0.009	0.005	0.009	0.015				0.012	0.012	0.034	0.028	0.010

Table 5: Result output and diagnostics for ICSV13. See definitions in Table 2.

Variable	f_i	$f_{total(i)}$	ΔT (°C)	ΔP (kbar)	$\mu_{obs} + 2\sigma$	$\mu_{obs} - 2\sigma$	mod
X_{Grs}	0.4998	0.4836	12.03	0.4867	0.07243	0.04593	0.05256
X_{Sps}	0.05676	0.2883	11.9	1.579	0.02657	0.01084	0.01915
X_{Prp}	1.628	0.4	4.154	0.2916	0.1857	0.1415	0.1995
X_{Mg} Grt	1.519	0.4142	3.946	0.2674	0.1992	0.1556	0.2106
Si Bt	0.1276	0.5609	3.315	0.2133	2.741	2.686	2.717
Ti Bt	0.4031	0.5334	8.731	0.4593	0.1183	0.09032	0.09869
Si Ms	0.3182	0.5383	3.315	0.2133	3.115	3.021	3.053
X_{Pa}	0.5078	0.5887	12.61	0.3664	0.2597	0.2006	0.2451
X_{Ab}	0.07038	0.5626	3.315	0.2133	0.9596	0.8155	0.8926

mean = 597 ± 8 °C, 9.82 ± 0.80 kbar (2σ)

median = 598 °C (IQR = 592–604 °C), 9.61 kbar (IQR = 9.28–9.93 kbar)

f_{total} (median) = 0.57

# of fitted variables = 7/9

model resolution = 3.03 °C, 0.0808 kbar

bootstrap resamples = 1000

Table 6: Result output and diagnostics for ICSV117. See definitions in Table 2.

Variable	f_i	$f_{total(i)}$	ΔT (°C)	ΔP (kbar)	$\mu_{obs} + 2\sigma$	$\mu_{obs} - 2\sigma$	mod
X_{Prp}	0.2188	0.7681	10.46	0.1434	0.1188	0.08252	0.09671
X_{Mg} Grt	0.4122	0.7718	5.495	0.09814	0.1577	0.1223	0.1473
Ti Bt	0.02354	0.833	13.77	0.2064	0.1650	0.1298	0.1478
X_{Mg} Bt	0.2958	0.776	5.929	0.1048	0.5210	0.4596	0.4994
X_{Ab}	1.502	0.61	2.214	0.04892	0.8162	0.7686	0.7567
Ti Amph	1.504	0.6097	3.617	0.06997	0.1097	0.05974	0.04716
Gln	0.06948	0.8758	5.442	0.4722	0.1697	0.1311	0.1491
Ts	1.746	0.575	2.214	0.04892	1.257	1.121	1.308

mean = 667 ± 10 °C, 11.77 ± 0.24 kbar (2σ)

median = 667 °C (IQR = 661–673 °C), 11.76 kbar (IQR = 11.64–11.94 kbar)

f_{total} (median) = 0.721

# of fitted variables = 5/8

model resolution = 2.02 °C, 0.0606 kbar

bootstrap resamples = 1000

Figure 1

selectively chosen variables

(a) too precise

(b) different sensitivity/precision

(c) multiple zones of overlap

(d) poor or no agreement

P (kbar)

X_{Ab}

Ti in Bt

2σ

X_{Grs}

1σ

Si in Ms

1σ

2σ

X_{Pa}

X_{Prp}

Si in Bt

X_{Cel}

X_{Mg} in Bt

X_{Sps}

T (°C)

(e) all 'robust' variables

composition   mode   composition + mode

inversion and bootstrap resampling

best-fit and uncertainty

P (kbar)

P (kbar)

T (°C)

T (°C)

variables

symbols

Grt   Bt   Ms   Pl

X_{Prp} Si   Si X_{Ab}

X_{Sps} Ti X_{Cel}

X_{Grs} X_{Mg} X_{Pa}

☆ result

◯ 1 sd ellipse

⬦ assemblage field

Figure 2

(a) best-fit *P-T*: grid-search inversion

(b) uncertainty analysis: bootstrap resampling
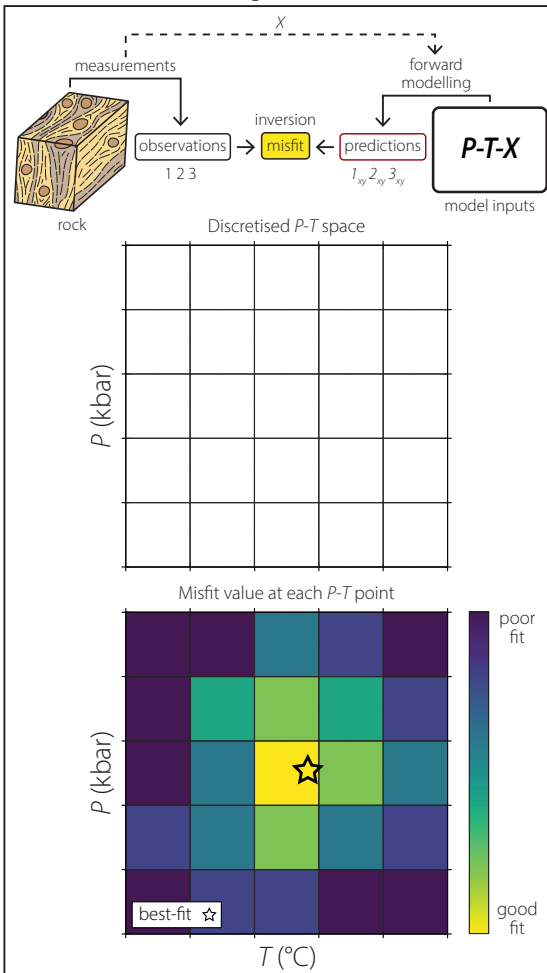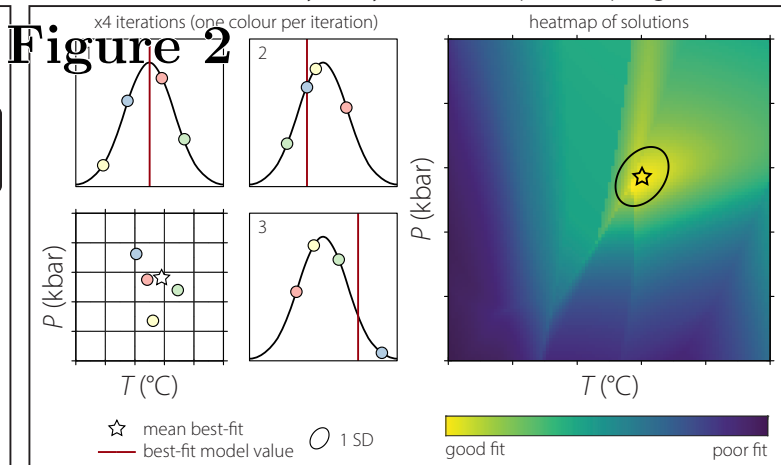
(c) sensitivity: bootstrap resampling

(a) all isopleth fields: % overlap heatmap

(b) select isopleth fields

Figure 3

(c) grid-search best-fit solutions

- ○ Resampled best-fit solutions
- — 1σ uncertainty ellipse
- ☆ Mean best-fit solution
- ☆ Median best-fit solution

Mdn = 598 °C (IQR = 592–604)   Mdn = 9.61 kbar (IQR = 9.28–9.93)

(d) residuals at best-fit *P-T* point

XGr[Grt] XSps[Grt] XPrp[Grt]

XMg[Grt]   Si[Bt]   Ti[Bt]

Si[Mu] XPa[Mu] XAb[Pl]

— Observations   - - - ± 2σ   ▬ Model result

(e) variable sensitivity at best-fit *P-T* point

597 °C   9.82 kbar

(a) pseudosection
ICSV13: Ky-St schist, all fields +Pl +Grt +Bt +Qtz +H$_2$O
XFe^{3+} = 0.065, H$_2$O mol% = 20.0
run conditions: aH$_2$O = 0.6

Figure 4

pelite

Ti-in-Bt & AvP intersection

1: Chl Ms Rt, 2: Chl Ms Ilm Rt, 3: Chl Ms Pa Ilm Rt, 4: Chl Ms Pa Ep Ilm Rt (−Pl), 5: Chl Ms Pa Ilm Rt (−Pl), 6: St Ms Pa Ilm Rt (−Pl), 7: Ms Pa Ilm Rt, 8: St Ms Rt, 9: Sil Ms Rt

(b) misfit + pseudosection

n = 9 (−X_{Mg} in Bt, −X_{Cel})

best-fit mean ☆
best-fit σ ⬭
Ti-in-Bt & av$P$ intersection ▱

log misfit

(f) pseudosection
ICSV117: Grt amphibolite, all fields +Di +Hbl +Bt +Qtz +H$_2$O,
XFe^{3+} = 0.14, H$_2$O mol % = 20.0
run conditions: aH$_2$O = 0.7

metabasite

Hbl-Pl & AvP intersection

X_{Na} Cpx > 0.35

1: Grt Omp Ep Ttn (−Di), 2: Grt Omp Ms Ep Ttn (−Di), 3: Grt Omp Ms Ep Ttn Rt (−Di), 4: Grt Omp Ms Rt

(g) misfit + pseudosection

n = 8 (−Si in Bt, −X_{Grs}, −Ed)

best-fit mean ☆
best-fit σ ⬭
Hbl-Pl & av$P$ intersection ▱

log misfit

(c) best-fit

598 °C

9.61 kbar

Ti-in-Bt   AvP

(d) 2-D hist

mean ☆
median ☆

N = 1000, max = 137

log number of points (n)

(e) % overlap

n = 9, max 100%

% overlap

(h) best-fit

667 °C

11.76 kbar

Hbl-Pl   AvP

(i) 2-D hist

mean ☆
median ☆

N = 1000, max = 63

log number of points (N)

(j) % overlap

n = 8, max 75%

% overlap