

LETTER

Accuracy versus precision in locating reaction boundaries: Implications for the garnet–plagioclase–aluminum silicate–quartz geobarometer

L. W. MCKENNA, K. V. HODGES

Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology,
Cambridge, Massachusetts 02139, U.S.A.

ABSTRACT

Determination of the P - T location of a reaction boundary is a fundamental prerequisite to geologic thermobarometry and experimental petrology. Too often, experimentalists make no distinction between how well a given data set fits a linear model (precision) and how well the same data set estimates the actual P - T location of a reaction boundary (accuracy). The two are distinct features of a calibration, often of very different magnitudes, with accuracy limits generally much larger than precision limits. Precision limits are minimized by highly precise bracketing experiments; accuracy limits are minimized by large numbers of bracketing experiments. The differences between and magnitudes of these uncertainties are demonstrated by consideration of the anorthite-breakdown reaction. Using data from five precise experimental studies, we derive the position of the reaction boundary for anorthite breakdown as P (in MPa) = $(2.20 \pm 0.15)T$ (in K) - $[6.2 (\pm 3.0) \times 10^2]$, where accuracy limits are at the 95% confidence level.

INTRODUCTION

The purpose of this note is to emphasize the sources and magnitudes of uncertainties in locating reaction boundaries. This has been a popular topic with the two of us, and we stress that our intention is not to promulgate the results of earlier papers, but instead to clearly define the different types of uncertainties inherent in experimental work. First, we discuss qualitatively these uncertainties and their meanings. We then briefly outline strategies for increasing both the precision and the accuracy of thermobarometry. We conclude that the prognosis for developing high-quality thermobarometers is very good, although a great deal of additional experimental work remains to be done.

THE CALCULATION AND SIGNIFICANCE OF UNCERTAINTIES

Methods for calculating uncertainties in a linear model of experimental data include some that appear (incorrectly we would say) computationally obese (e.g., Hodges and McKenna, 1987) and some (e.g., Demarest and Haselton, 1981) that are elegantly simple. The uncertainties calculated by these two methods are generally numerically different, with the uncertainty calculated by the method described by Hodges and McKenna (1987) much greater than that calculated by the method described by Demarest and Haselton (1981). This apparent discrepancy is resolved by the subtle fact that the two methods

are measuring different uncertainties: the Demarest and Haselton method estimates the *precision* of a calibration, whereas the Hodges and McKenna method estimates the *accuracy* of a calibration. These are distinct, but complementary, quantities.

Simply put, precision is a measure of the reproducibility of observations of a sample thought to be representative of a population. Accuracy is a measure of how well the sample represents the population. In order to visualize the difference between accuracy and precision, consider the plight of fans of the Baltimore Orioles who attended every home game in April of 1988. Were these individuals to draw a conclusion about the behavior of the population (all Major League baseball teams) from a small sample with unknown systematic errors (the Orioles, with a home record of zero wins and eight losses), they might conclude that no team can win a game at home. Given the circumstances, their assessment of the situation would be precisely correct. Only observations of different samples (e.g., the Pittsburgh Pirates, with eight wins and one loss at home) would convince our friends that their ballpark estimate of the population behavior was inaccurate.

Figure 1a briefly demonstrates the method of uncertainty estimation derived by Demarest and Haselton (1981). The approach is insightful and direct, but constructs only a probability distribution (Fig. 1a, part D) of the location of the reaction boundary between two bracketing points. As Demarest and Haselton (1981) noted in

their conclusions, this technique does not provide the answer to the problem of "locating the phase boundary curve in *P, T* space." The bracketing uncertainties derived by the Demarest and Haselton method could be used to determine a set of "permissible" reaction boundaries, but this approach would require the implicit assumption that the bracketing data are accurate. In order to be statistically rigorous about the actual location of a reaction boundary, we must explore the *likelihood* that the data are accurate, not simply assume that they are. This does not mean that the Demarest and Haselton method is not useful, but merely indicates a limitation, recognized by its authors, in its applicability.

In contrast, the method developed by Hodges and McKenna (1987), briefly described in Figure 1b, places limits on the accuracy of a calibration. This procedure determines, for a given data set, the range of *P-T* conditions that include, at some selected confidence level, the actual reaction-boundary position. This uncertainty is always greater than that calculated by the Demarest and Haselton method because even a precise calibration may not locate the reaction boundary in an absolute sense. If we are interested in knowing the accuracies of thermodynamic quantities derived from experimental data (e.g., the entropy and enthalpy of a reaction), then we must use a method, such as the one by Hodges and McKenna, that explicitly accounts for potential inaccuracies (e.g., systematic error) in the data. The inaccuracies of derived thermodynamic quantities are directly proportional to the error in the regression (the residuals of the data), the slope uncertainty (a function of the precision and collinearity of the data), and inversely proportional to the *T* range of the data. In the case that the data are perfectly collinear (the residuals are zero), these uncertainties are still substantially greater than zero for a finite-numbered data set.

DECREASING THE UNCERTAINTIES OF THERMOBAROMETRIC ESTIMATES

With these insights in mind, we now consider methods and strategies for increasing both the precision and the accuracy of reaction calibrations. Koziol and Newton (1988) recently published a set of nine experimental brackets for the anorthite-breakdown reaction,



and used the Demarest and Haselton method for determining limits for each experimental bracket. The bracketing data and the Demarest-Haselton-based uncertainties were then used to obtain ΔH and ΔS of the reaction and an estimate of the uncertainties in these quantities. Table 1 includes Koziol and Newton's results as well as the equivalent Hodges-McKenna-based values and uncertainties for the same data. Although the two treatments yield similar ΔH and ΔS (both methods use similar techniques to determine the "best-fit" model), the method used by Koziol and Newton seriously underestimates the full uncertainty in these quantities. The magnitude of the uncertainties based on the Hodges and McKenna

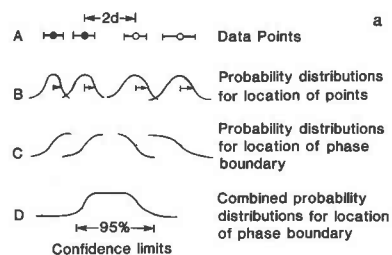


Fig. 1a. Schematic illustration of the method developed by Demarest and Haselton (1981) for estimating the probability distribution of the location of a reaction boundary between two experimental brackets. (A) Data points: filled points represent the assemblage stable to the left of the reaction boundary, unfilled points the assemblage stable to the right of the reaction boundary. The two experiments that most closely bracket the reaction are separated by the distance $2d$. (B) The uncertainty in the location of each point is assumed to be normally distributed, with a standard deviation shown by the arrowed lines. (C) The probability that the reaction boundary lies to the right (filled points) or left (unfilled) of a given position is the integral of the point's probability distribution. (D) The probability distribution of the location of the phase boundary is the product of each point's probability distribution. Redrawn from Demarest and Haselton (1981).

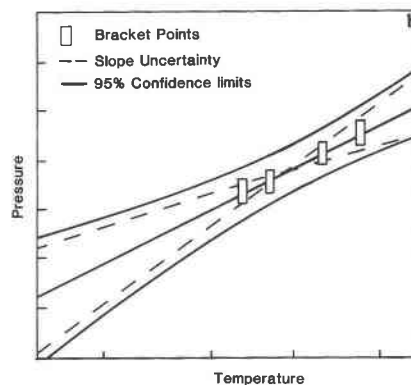


Fig. 1b. Schematic illustration of the method developed by Hodges and McKenna (1987) for determining the location and uncertainty of a reaction bracket. Bracket points have normally distributed uncertainties, with a mean equal to the midpoint of the experimentally determined bracket, and standard deviation determined by the distance between the experimental brackets. The slope uncertainty is determined by a least-squares, two-error regression after York (1969). The 95% confidence limits are determined by terms proportional to the slope uncertainty, the *P-T* spread of the data, and the number of points. The confidence limits are defined as the lines that have a 95% probability of including the actual reaction boundary.

method reflects the simple fact that nine points, although they may precisely define a straight line, may not accurately define the *correct* straight line.

Figure 2a illustrates the brackets and both their Demarest-Haselton and Hodges-McKenna uncertainties. The accuracy limits for these data are approximately twice the

TABLE 1. Derived thermodynamic constants

Method†	ΔH	$2\sigma_H$ ††	ΔS	$2\sigma_S$
Data set of Koziol and Newton (1988)				
DH	4.85×10^4	$+2 \times 10^4$ -1×10^4	151	+14 -7
HM	4.90×10^4	3.1×10^4	150	16
	$2\sigma_H = (1.57 \times 10^{-4})[(3.94 \times 10^{10}) + (1.55 \times 10^{10})(T - 1300)^2]^{1/2}$			
Data set from five studies*				
HM	4.1×10^4	2.0×10^4	146	10
	$2\sigma_H = (1.36 \times 10^{-4})[(1.13 \times 10^{10}) + (5.34 \times 10^9)(T - 1443)^2]^{1/2}$			

Note: ΔH and ΔS are in units of J/mol and J/(mol·K), respectively.

† DH: modified Demarest and Haselton (1981) method. Uncertainties calculated from data presented in Koziol and Newton (1988). HM: method described by Hodges and McKenna (1987).

†† Calculated at $T = 0$ K. Expressions for σ_H give the uncertainty as a function of temperature; see Hodges and McKenna (1987) for details on calculation.

* Data from Koziol and Newton (1988), Goldsmith (1980), Gasparik (1984), Hariya and Kennedy (1968), and Hays (1966).

width of the bracketing uncertainties themselves, representing the difference between the precision and accuracy of the brackets. Although the Koziol and Newton (1988) data consist of nine high-precision brackets, the accuracy limits for these data alone are similar in magnitude to those derived by Hodges and McKenna (1987) from a data set of fourteen brackets from three previous high-quality studies of this same reaction (Hays, 1966; Hariya and Kennedy, 1968; Goldsmith, 1980). The similarity reflects the strength of numbers—many *good* experiments are as accurate, although perhaps not as precise, as a few *very good* experiments.

A more accurate estimate for the location of the R1 reaction boundary can be made by combining the data from five published studies, yielding a total of 27 brackets for the reaction. The regression results for these data are listed in Table 1 and displayed in Figure 2b. The accuracy uncertainties for these data are approximately 60% of those for either the Koziol and Newton (1988) or three-study (Hodges and McKenna, 1987) data sets alone, reflecting the quality, *P-T* range, and the number of the data. In fact, the resultant uncertainty envelope is roughly tangential to the bracketing uncertainties themselves, suggesting that further improvements in the accuracy of this calibration are virtually limited by the precision of the data. This theoretical minimum in the accuracy uncertainties is intuitive: given an infinite number of points of infinite *P-T* range, the uncertainty in the position of a boundary is equal to the imprecision of the bracketing experiments, as would be expected from the work of Demarest and Haselton (1981). Further increases in the accuracy of the R1 calibration are also limited by small systematic errors in the data: Figure 2b shows that the brackets determined by Hariya and Kennedy (1968) are consistently higher than the best-fit reaction curve, whereas those of Goldsmith (1980) are consistently lower.

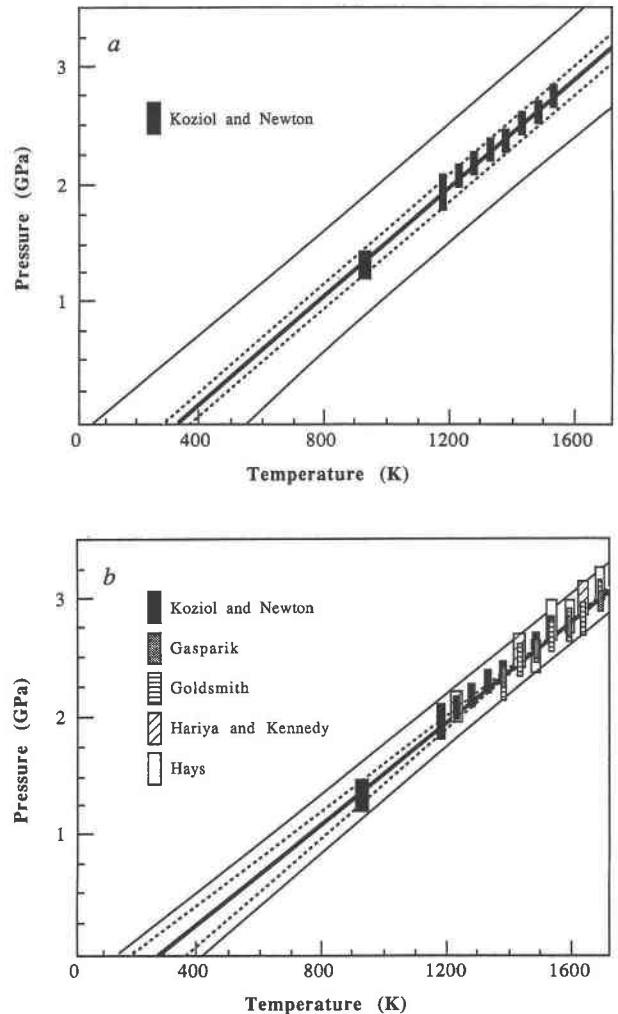


Fig. 2. The location of the anorthite-breakdown reaction (see Eq. R1). (a) The data of Koziol and Newton (1988), with uncertainty limits determined by Koziol and Newton (1988) (modified Demarest and Haselton method, dashed lines) and by the method described by Hodges and McKenna (1987) (solid lines, 95% confidence level). The best-estimate position is shown by the heavy, central, solid line. In effect, the dashed lines give precision limits, whereas the solid lines give accuracy limits, for the data. (b) The location of the anorthite-breakdown reaction determined by 27 brackets from five different studies. Solid curves give 95% confidence limits derived by the Hodges and McKenna method, whereas dashed lines give approximate Demarest and Haselton uncertainties. Note that the accuracy limits at the 95% confidence level for this combined data set are 60% of those of the Koziol and Newton (1988) data set alone (even though the Koziol and Newton experiments were of high precision), illustrating the importance of large numbers of data in accurately locating the position of the reaction boundary. The uncertainty in Fig. 2b is limited by both the systematic scatter between the results of different studies and the imprecision of the brackets.

For use in thermobarometry, the 27-point data set gives paleopressures from the relationship

$$P = (2.20 \pm 0.15)T - [6.2 (\pm 3.0) \times 10^2], \quad (1)$$

where *P* is in megapascals and *T* is in kelvins and where the uncertainties are at the 95% confidence level. These inaccuracies propagate into paleopressure uncertainties of $\pm 2.5 \times 10^8$ Pa (2.5 kbar). Although the precision of an individual study might give an apparently smaller uncertainty (e.g., Koziol and Newton (1988): $\pm 0.6 \times 10^8$ Pa), accuracy is the important parameter in absolute thermobarometry. Although a potential inaccuracy of $\pm 2.5 \times 10^8$ Pa in an estimated paleopressure may seem large, we stress that the addition of the new Koziol and Newton (1988) data to the pre-existing data set improved the accuracy of this geobarometer to the point that it is now likely to be the best-calibrated thermobarometric reaction available.

DISCUSSION AND CONCLUSIONS

Accurate calibration of the positions of reaction boundaries require many experiments, preferably done in different laboratories, on different presses and over a wide range of *P-T* conditions. In most cases, "many experiments" may be as high as twenty or thirty brackets. For the anorthite-breakdown reaction, usually described by its investigators as an "important" reaction, the present excellent calibration required at least six separate studies and 22 years of effort.

Additional work is necessary to derive a more general method to accurately locate the position of curved reaction boundaries in *P-T* space. Chayes (1968) addressed this problem, but his technique is not amenable to analytical error analysis, and some features of his model suggest that it is not rigorously correct (Demarest and Haselton, 1981). Although many reaction boundaries that are curved in *P-T* space can be linearized by a coordinate transformation (Chayes, 1968), the general solution to this problem is not in hand.

We stress that statistical sound and fury do not mitigate the fact that not all experimental studies are of equivalent quality. Uncritical acceptance of all experimental data pertinent to the position of a reaction boundary, regardless of internal inconsistencies in the cumulative data set, can lead to incorrect determinations of thermodynamic constants or overly pessimistic accuracy estimates. As H. T. Haselton, Jr., reminded us in his review of an earlier version of this paper, "error analysis must be accompanied by a reasonable amount of insight and intelligence." The results of suspect experimental studies should

be disregarded in a statistical analysis of the position of a reaction boundary if we can objectively justify our suspicions. For example, we have ignored the results of one experimental determination of the anorthite-breakdown reaction (Schmid et al., 1978) in this paper because we feel that the study incorrectly assumed no excess volume of mixing in the garnet solid solution.

Finally, we feel that arguments such as "the reaction has been calibrated already" are not sufficient reasons to avoid repeating experiments. The operative questions must be "How well should we know this reaction?" and "How many brackets will that require?"

ACKNOWLEDGMENTS

We thank David Silverberg for stimulating our interest in this problem and for consistently asking us to "Wait a minute and prove that." We also thank H. T. Haselton and J. L. Munoz for expeditious and thoughtful reviews of the original manuscript; A. MacFarlane and T. Juster kindly reviewed our first draft. Finally, our appreciations to the editors of the *American Mineralogist* for creating this new venue for "brief, timely communications." This work was supported by National Science Foundation Grants EAR 8319768 and EAR 8407730.

REFERENCES CITED

- Chayes, F. (1968) On locating field boundaries in simple phase diagrams by means of discriminant functions. *American Mineralogist*, 53, 359–371.
- Demarest, H.H., Jr., and Haselton, H.T., Jr. (1981) Error analysis for bracketed phase equilibrium data. *Geochimica et Cosmochimica Acta*, 45, 217–224.
- Gasparik, T. (1984) Experimental study of the subsolidus phase relations and mixing properties of pyroxene in the system CaO-Al₂O₃-SiO₂. *Geochimica et Cosmochimica Acta*, 48, 2537–2545.
- Goldsmith, J.R. (1980) Melting and breakdown reactions of anorthite at high pressure and temperatures. *American Mineralogist*, 65, 272–284.
- Hariya, Y., and Kennedy, G.C. (1968) Equilibrium study of anorthite under high temperature and high pressure. *American Journal of Science*, 266, 193–203.
- Hays, J.F. (1966) Lime-alumina-silica. *Carnegie Institution of Washington Year Book* 65, 234–239.
- Hodges, K.V., and McKenna, L.W. (1987) Realistic propagation of uncertainties in geologic thermobarometry. *American Mineralogist*, 72, 671–680.
- Koziol, A.M., and Newton, R.C. (1988) Redetermination of the anorthite breakdown reaction and improvement of the plagioclase-garnet-Al₂SiO₃-quartz geobarometer. *American Mineralogist*, 73, 216–223.
- Schmid, R., Cressey, G., and Wood, B.J. (1978) Experimental determination of univariant equilibria using divariant solid-solution assemblages. *American Mineralogist*, 63, 511–513.
- York, D. (1969) Least-squares fitting of a straight line with correlated errors. *Earth and Planetary Letters*, 5, 320–324.

MANUSCRIPT RECEIVED JULY 6, 1988

MANUSCRIPT ACCEPTED JULY 15, 1988

Editor's note: For requirements of the Letter section, see the July–August 1988 issue of American Mineralogist, vol. 73, p. 940.