American Mineralogist, Volume 72, pages 671–680, 1987

Realistic propagation of uncertainties in geologic thermobarometry

K. V. Hodges, L. W. McKenna

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

Abstract

Two of the most significant sources of uncertainty in geologic thermobarometry are analytical imprecision and the systematic error associated with experimental calibration techniques. Analytical uncertainties are sample-specific and dictate the precision of a *P-T* estimate. Calibration uncertainties are reaction-specific and effectively limit the accuracy of an estimate. We describe a systematic method of propagating both types of uncertainty through thermobarometric calculations in order to place realistic confidence limits on *P-T* estimates. As an example, we evaluate the accuracy and precision of garnet-biotite, garnet-plagioclase-kyanite-quartz, and garnet-rutile-kyanite-ilmenite-quartz thermobarometry for a pelitic sample from the Funeral Mountains of southeastern California. Calibration and analytical uncertainties together propagate into absolute pressure and temperature uncertainties (95% confidence level) of several hundred megapascals and more than 100 K. Analytical imprecision accounts for only 10–20% of the pressure uncertainty and less than 30% of the temperature uncertainty. Our capacity to confidently calculate equilibration pressures and temperatures for geologic samples seems rather limited, but it can be improved significantly through additional careful experimental work. Comparative thermobarometry, which involves applying a single set of thermobarometers to different samples in order to calculate *differences* in *P-T* conditions, eliminates the systematic error associated with experimental calibrations. Through careful analytical work, it is possible to confidently resolve *P-T* differences of as little as a few tens of degrees and a few tens of megapascals.

Introduction

In recent years, quantitative thermobarometry and thermodynamic modeling of mineral zoning have led to a dramatic increase in the contributions made by metamorphic petrology to a better understanding of tectonic processes (e.g., Hollister, 1979; Tracy and Robinson, 1980; Hodges and Royden, 1984; Selverstone, 1985). Unfortunately, many of us are negligent (or at least optimistic) about the assignment of uncertainties to *P-T* estimates, despite the significant effect these uncertainties may have on the believability of a tectonic interpretation. The sources of uncertainty in thermobarometry have been reviewed by Essene (1982) and Powell (1985), but few attempts have been made to propagate these uncertainties into realistic confidence limits for actual *P-T* estimates. In this paper, we describe a systematic approach to expressing uncertainties in *P-T* estimates derived through the simultaneous solution of two experimentally calibrated thermobarometers.

Sources of uncertainty

There are four basic sources of uncertainty in thermobarometry.

1. Disequilibrium effects. It is impossible to "prove" that a given metamorphic assemblage ever achieved equilibrium, although textural and chemical evidence may be cited to support the assumption. In most applications of thermobarometry, only the outermost rims of adjacent phases are assumed to reflect equilibrium conditions. This assumption reduces some of the uncertainty involved, but an increasing body of kinetic data for metamorphic phases argues that disequilibrium is a problem to be reckoned with (e.g., Loomis, 1983; Carpenter and Putnis, 1986). Many thermobarometric assemblages from regionally metamorphosed terranes "re-equilibrate" during uplift and slow cooling such that the calculated pressures and temperatures underestimate "peak" metamorphic conditions (e.g., Tracy et al., 1976; Hodges and Spear, 1981). In these instances, the thermobarometric results may help define the *P-T* trajectory of a terrane during uplift and cooling (e.g., Hodges and Royden, 1984), but often differences between the kinetics of retrograde exchange and net-transfer reactions result in partial equilibrium (e.g., Loomis, 1976) or outright disequilibrium. *P-T* uncertainties associated with deviations from equilibrium may be extremely significant, but they are impossible to quantify at present.

2. Calibration uncertainties. All of the commonly used methods of calibration introduce significant uncertainties

0003–004X/87/0708–0671$02.00

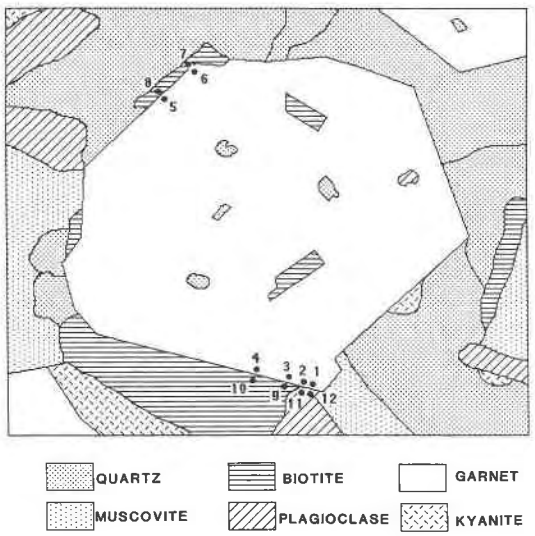

Fig. 1. Line drawing of a hypothetical domain within a pelitic schist. Numbered points indicate spot analyses referred to in text.

that propagate as systematic errors in P-T estimates. Fortunately, these uncertainties can be quantified: Anderson (1977a, 1977b) has discussed the uncertainties associated with calculating the position of an equilibrium curve in P-T space using thermochemical data, and Hodges and Crowley (1985) have examined the uncertainties inherent in empirical calibration. Relatively little attention has been paid to the uncertainties associated with direct experimental calibration, even though experimentally calibrated equilibria are inherently the most reliable thermobarometers.

3. Analytical uncertainties. Analytical uncertainties associated with electron-microprobe analysis fall into two categories: (a) standardization, X-ray counting, and correction uncertainties inherent to the technique and (b) variations in composition associated with variable scales of equilibrium. Uncertainties of the first kind have been treated routinely for some time (e.g., Smith, 1976), but actual compositional variability within samples is seldom dealt with in systematic fashion.

4. Solution modeling. Many of the phases commonly used for thermobarometric calculations exhibit nonideal solution behavior, in some cases necessitating major corrections of measured compositions to arrive at reasonable activities. Phase-equilibrium experiments (e.g., Eugster et al., 1972), calorimetric measurements (e.g., Newton et al., 1977), and empirical observations (e.g., Ganguly and Kennedy, 1974; Ghent et al., 1979; Hodges and Spear, 1982) have been used to formulate activity coefficients.

Unfortunately, there is often disagreement among workers about the magnitude of various ΔG_{excess} terms and even the mathematical form of solution models. This has led to the widespread practice of reporting multiple P-T estimates reflecting the effects of different solution models. For example, garnet-biotite temperatures have been calculated using an ideal solution model for garnet (Ferry and Spear, 1978) and nonideal models proposed by Newton and Haselton (1981), Hodges and Spear (1982), and Ganguly and Saxena (1984). In some cases, P-T estimates calculated using different solution models can vary widely, implying that uncertainties in solution behavior constitute a major source of error in thermobarometry. In general, this error cannot be quantified because the assumptions inherent in different models are often mutually exclusive.

For the remainder of this paper, we will treat only the effects of analytical and calibration uncertainties on our confidence in P-T estimates, because these are the only uncertainties that, at present, lend themselves to rigorous error-propagation techniques. Because analytical uncertainties pertain to individual samples, they propagate into minimum estimates of the *precision* of a P-T calculation. Calibration uncertainties reflect our confidence in the positions of end-member reactions in P-T space and thus propagate into minimum estimates of the *systematic error* associated with the use of specific thermobarometers. Together, the two types of uncertainty provide a sense of the *accuracy* of a P-T estimate.

Propagation of analytical uncertainties

Because it is impractical to analyze every segment of every rim of every pertinent grain in a sample, we must use some systematic (yet cost-efficient!) technique to estimate compositional variability. First, we select two to three domains in each probe mount. A hypothetical "domain" in a quartz-muscovite-biotite-garnet-plagioclase-kyanite schist is illustrated in Figure 1. Knowing a priori that we want to simultaneously solve the equilibria

$$\underset{\text{garnet}}{Mg_3Al_2Si_3O_{12}} + \underset{\text{biotite}}{KFe_3AlSi_3O_{10}(OH)_2} = \underset{\text{garnet}}{Fe_3Al_2Si_3O_{12}} + \underset{\text{biotite}}{KMg_3AlSi_3O_{10}(OH)_2} \quad \text{(R1)}$$

and

$$\underset{\text{garnet}}{Ca_3Al_2Si_3O_{12}} + 2Al_2SiO_5 + SiO_2 = \underset{\text{plagioclase}}{3CaAl_2Si_2O_8} \quad \text{(R2)}$$

for rim temperature and pressure, we have chosen this domain such that the non-end-member phases participating in the equilibria (garnet, biotite, and plagioclase) share sharp mutual boundaries. We would analyze minerals along each of the critical boundaries twice, as near as possible to the "rim" (points 1–6 for garnet, points 7–10 for biotite, and points 11 and 12 for plagioclase). In order to account for narrow concentration gradients in these phases, we might define the "rim" by analyzing points along short traverses across each boundary.

The same procedure is repeated for each domain, resulting in a suite of rim analyses for each non-end-member phase. Using the formula basis for each analysis, we

calculate the mean ($\bar{X}$) and sample variance (Var_X) for each element (see Table 1 for notation of variables),

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \tag{1}$$

$$\mathrm{Var}_X = n\sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i\right)^2, \tag{2}$$

where X_i is the composition measured at $i = 1$ to n points. Note that Var_X is related to the commonly used estimated standard deviation (esd) of X, s_X by

$$\mathrm{Var}_X = s_X^2. \tag{3}$$

Mole fractions of various components (e.g., almandine in the garnet solution) are calculated using measured compositional data, and their uncertainties are obtained by applying the basic error-propagation equation (Larsen and Marx, 1981). For a function F of variables $z_1, z_2, \ldots, z_n$,

$$\mathrm{Var}_F = \left(\frac{\delta F}{\delta z_1}\right)^2 \mathrm{Var}_{z_1} + \left(\frac{\delta F}{\delta z_2}\right)^2 \mathrm{Var}_{z_2} + \ldots + \left(\frac{\delta F}{\delta z_n}\right)^2 \mathrm{Var}_{z_n} + 2\sum_{i=1}^{n}\sum_{j=1}^{n} \rho_{z_i z_j} \frac{\delta F}{\delta z_i}\frac{\delta F}{\delta z_j} \mathrm{Var}_{z_i}^{1/2} \mathrm{Var}_{z_j}^{1/2}, \tag{4}$$

where variables $z_1, z_2, \ldots z_n$ have normally distributed random uncertainties, and $\rho_{z_i z_j}$ are covariance coefficients describing the interdependence between the variances of any two compositional variables.

At this point, we wish to calculate a pressure and a temperature by simultaneously solving two equilibrium-condition equations of the form

$$0 = \Delta H - T\Delta S + (P - 10^5)\Delta V + RT \ln K, \tag{5}$$

where R is the gas constant, P is pressure, T is temperature, and ΔH, ΔS, and ΔV are the enthalpy, entropy, and volume change for the reaction. The equilibrium constant (K) is a function of various mole fractions and (if nonideal solutions are involved) T and P. If we assume that all of the uncertainty in T and P is a consequence of analytical uncertainty, then we could use equations such as 4 to propagate uncertainties in mole fractions into s_T and s_P. Although rigorous, this approach can be cumbersome if complex solution models are involved.

A computationally simpler approach was suggested by Steltenpohl and Bartley (1984) and involves Monte Carlo propagation of the compositional uncertainties (Anderson, 1976, 1977b; Hodges and Crowley, 1985). The procedure is (1) create a 200-element array for each mole fraction, consisting of a population of values normally distributed about the accepted value; (2) choose a "seed" pressure and temperature (e.g., 500 MPa and 500 K) for the calculation; (3) calculate both equilibrium constants using a randomly chosen element from each mole-fraction array, as well as the "seed" P and T, if necessary; (4) calculate P and T as a function of the various thermodynamic variables and the equilibrium constants; (5) iterate step 4 until P and T converge; and (6) iterate steps 2 through 5 100 times in order to produce a total of 100 P-T pairs for the sample. The Monte Carlo method of error propagation is strictly valid only when the variables subject to uncertainty are independent. Because mole fractions are constrained to sum to unity for a single phase, this condition is not completely satisfied in our application. Covariances between mole fraction variables will be dominantly negative; consequently, Equation 4 indicates that the Monte Carlo method will slightly overestimate the P-T uncertainty.

TABLE 1. Statistical and thermodynamic variables

X_i = the ith datum in a set $X_1, X_2, \ldots, X_n$
$\bar{X}$ = sample mean of X
Var_X = sample variance of X
n = the number of data in a sample
s_X = sample esd of X
F = any function of variables $z_1, z_2, \ldots, z_n$
$\rho_{X_i X_j}$ = covariance coefficient of X_i and X_j
ω_T, ω_P = uncertainty in calculated T and P
E = ellipticity of an uncertainty ellipse
M = slope of a line in two-dimensional space
B = ordinate intercept of a line in two-dimensional space
X_k = a particular value of X
$\hat{Y}_k$ = expected value of Y from a regression at X_k
W = confidence interval of a regression
t = Student's distribution coefficient
α = confidence interval for the Student's t distribution

H = enthalpy (J/mol) S = entropy [J/(mol·K)]
T = temperature (K) P = pressure (Pa)
V = volume (m³/mol) K = equilibrium constant
R = gas constant [8.314 J/(mol·K)]

In P-T space, the Monte Carlo array defines an elliptical cloud that reflects the estimated uncertainty in P and T as well as the correlation between these uncertainties. We extend the approach of Steltenpohl and Bartley (1984) by calculating a best-fit, $2s$ (95% confidence) ellipse for the data. We begin by defining the long axis of the ellipse as parallel to a best-fit line passing through the P-T array. In practice, we use the mean of least-squares linear regressions of P on T and T on P. The slope of this line is used to define a rotation matrix that permits reorientation of the elliptical cloud with principal axes parallel to P and T axes. The resultant cloud is translated to the origin in P-T space, and an ellipse is defined such that its radii span $\pm 2s$ variations in P and T from the origin. This ellipse is then translated and rotated back to its proper position in P-T space.

Figure 2 illustrates the results of this sort of analysis for a pelitic schist sample from the Monarch Canyon area of the Funeral Mountains, southeastern California (Labotka, 1980; Hodges and Walker, in prep.). The sample contains the subassemblage ilmenite-rutile-garnet-kyanite-plagioclase-muscovite-biotite-quartz, permitting the application of thermobarometers R1 and R2, as well as GRAIL,

TABLE 2. Compositional data for FM-12

Mineral	Component	Mole fraction	Uncertainty (1s)
Garnet	Almandine	0.798	7.2×10^{-3}
	Pyrope	0.119	7.7×10^{-3}
	Grossular	0.0464	8.5×10^{-4}
	Spessartine	0.0371	1.7×10^{-4}
Biotite	Annite	0.427	4.8×10^{-3}
	Phlogopite	0.399	4.3×10^{-3}
Muscovite	K/Σalkalis	0.842	3.8×10^{-2}
	Na/Σalkalis	0.157	3.8×10^{-2}
	^{VI}Al	0.911	6.6×10^{-3}
Plagioclase	Albite	0.886	8.4×10^{-3}
	Anorthite	0.110	8.2×10^{-3}
Ilmenite	Fe^{2+}	0.958	3.0×10^{-3}

$$\underset{\text{ilm}}{3FeTiO_3} + \underset{\text{kyanite}}{Al_2SiO_5} + \underset{\text{quartz}}{2SiO_2} = \underset{\text{garnet}}{Fe_3Al_2Si_3O_{12}} + \underset{\text{rutile}}{3TiO_2}, \quad \text{(R3)}$$

which was experimentally calibrated by Bohlen et al. (1983). Rim compositions for non-end-member phases were obtained using the JEOL Superprobe at MIT (Table 2) and manipulated using the techniques outlined above to obtain precisions for mole fractions. The dashed ellipse in Figure 2 corresponds to the precision of a P-T estimate made by using the Ferry and Spear (1978) and Goldsmith (1980) calibrations of R1 and R2, respectively, with solution models as described by Hodges and Spear (1982) and modified by Hodges and Royden (1984). We can calculate the approximate precisions associated with the geothermometer (ω_T) and geobarometer (ω_P) using half of the spread in T and P (respectively) enclosed by the ellipse. These uncertainties are ±45 K for R1 and ±105 MPa for R2 and are in good agreement with commonly accepted uncertainties for the equilibria (e.g., Ferry, 1980; Ghent et al., 1982). The ellipticity (E) of the uncertainty field (defined as the major radius divided by the minor radius) provides a measure of the dependence of the overall precision on the individual precisions of the thermobarometers; ellipticities greater than 10 indicate that the precision of one of the equilibria dominates the precision of the P-T estimate. In this case, $E = 12.8$, implying that most of the uncertainty represented by the ellipse is a function of the uncertainty in the composition of phases involved in equilibrium R1.

We can explore the sensitivity of the P-T estimate to the uncertainties associated with individual phases by setting selected mole-fraction uncertainties to zero and deriving new ellipses (Fig. 3a). For this sample, most of the uncertainty in R1 stems from uncertainties in the garnet composition, whereas most of the uncertainty in R2 depends on the precision of the plagioclase analyses.

The dotted ellipse in Figure 2 represents the precision of a P-T estimate calculated by using R1 and R2 and alternative garnet (Ganguly and Saxena, 1984) and plagioclase (Newton and Haselton, 1981) solution models. The resultant shift in the absolute position of the ellipse is relatively minor, but the size of the ellipse increases markedly ($\omega_T = 53$ K, $\omega_P = 135$ MPa, $E = 11.0$). This is a consequence of the increased mathematical complexity and greater dependence on compositional terms associated with the alternative solution models, and it emphasizes the point that more complex solution models lead to less precise (albeit possibly more accurate) P-T estimates.

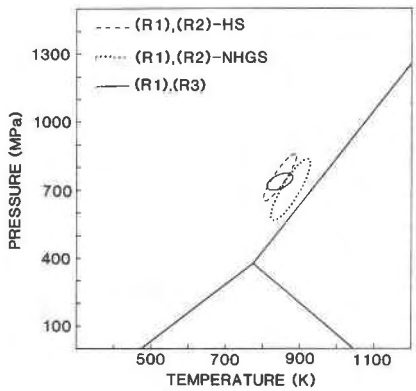

Fig. 2. Analytical uncertainty ellipses (2s) for the Funeral Mountains sample. R1, R2, and R3 refer to equilibria used. HS and NHGS indicate solution models of Hodges and Spear (1982) or Newton and Haselton (1981) and Ganguly and Saxena (1984). Aluminosilicate stability fields after Holdaway (1971).

The solid-line ellipse in Figure 2 shows the results for simultaneous solution of thermobarometers R1 and R3, using the Bohlen et al. (1983) calibration of R3 and assuming the Hodges and Spear (1982) solution model for garnet. For the ellipse, $\omega_T = 35$ K, $\omega_P = 37$ MPa, and $E = 3.6$. The precision of the R1-R3 estimate is not strongly dependent on the precision of one thermobarometer or the other, and the precision of the GRAIL geobarometer is inherently greater than that of the garnet-plagioclase-aluminosilicate-quartz geobarometer. Sensitivity analysis (Fig. 3b) reveals that almost all of the uncertainty in the R1-R3 estimate is due to uncertainty in garnet composition. Our ultimate conclusion for the Funeral Mountains sample is that the R1-R2 and R1-R3 P-T estimates (solid and dashed ellipses) are internally consistent within 2s analytical uncertainty.

PROPAGATION OF CALIBRATION UNCERTAINTIES

Equation 5 can be rewritten as

$$P = \frac{T(\Delta S - R \ln K)}{\Delta V} + \frac{-\Delta H}{\Delta V} + 10^5 \quad \text{(6a)}$$

or

$$\ln K = -\frac{1}{T}\frac{\Delta H + (P - 10^5)\Delta V}{R} + \frac{\Delta S}{R}. \quad \text{(6b)}$$

Commonly a net-transfer reaction (e.g., R2) is experimentally calibrated as a thermobarometer by bracketing

the position of the end-member reaction in P-T space, performing a linear regression on the bracketing data, and deriving values for reaction entropy and enthalpy from the slope and P intercept of the regression line by using Equation 6a. For an exchange reaction (e.g., R1), mineral pairs are allowed to equilibrate at a variety of temperatures, their compositions are measured and used to calculate equilibrium constants, and ΔH and ΔS are calculated from $1/T$ vs. ln K regression parameters by using Equation 6b. Solving Equations 6a and 6b for ΔS and ΔH yields

$$\Delta S = M_a \Delta V \qquad \Delta H = -(B_a - 10^5)\Delta V \tag{7a}$$

$$\Delta S = RB_b \qquad \Delta H = -M_b R - (P - 10^5)\Delta V, \tag{7b}$$

where M and B are the slope and Y intercept of the regression lines, respectively, and the a and b subscripts refer to the (T, P) and ($1/T$, ln K) forms of Equations 6a and 6b, respectively. Applying Equations 3 and 4 to Equations 7a and 7b produces the formulae for the uncertainties in ΔH and ΔS. For (T, P) space

$$s^2_{\Delta H} = \Delta V^2 s^2_{B_a} + (B_a - 10^5)^2 s^2_{\Delta V} \qquad s^2_{\Delta S} = \Delta V^2 s^2_{M_a}, \tag{8a}$$

and for ($1/T$, ln K) space

$$s^2_{\Delta H} = R^2 s^2_{M_b} + \Delta V^2 s^2_P + (P - 10^5)^2 s^2_{\Delta V} \qquad s^2_{\Delta S} = R^2 s^2_{B_b}, \tag{8b}$$

where s_B and s_M are the uncertainties in the Y intercept and slope of the regression line, $s_{\Delta V}$ is the uncertainty in the volume change for the reaction, and s_P is the uncertainty of the pressure at which $1/T$ vs. ln K experiments were conducted. Because ΔV is generally well known (Robie et al., 1978), $s^2_{\Delta V}$ is rather small and can be dropped from Equations 8a and 8b without introducing significant error. The experimental uncertainties that contribute to $s_{\Delta H}$ and $s_{\Delta S}$ include (1) the temperature uncertainty in each run; (2) the pressure uncertainty in each run; and (3) in the case of net-transfer reactions, the pressure and temperature differences between pairs of experiments that bracket the position of a reaction over a range of conditions.

Temperature uncertainties for calibration experiments generally range from 5 to 15 K (Schmid et al., 1978; Bohlen et al., 1983), depending on the apparatus used. The pressure uncertainties in calibration experiments vary widely (Hays, 1966; Bohlen et al., 1983). Johannes (1978) and Johannes et al. (1971) have discussed accuracies of piston presses over a range of run parameters. In general, they have concluded that experimental runs in piston-cylinder apparatus with alkali halide pressure media are accurate to $\pm 0.5 \times 10^8$ Pa, whereas the same presses run with talc or pyrophyllite cells may have inaccuracies as great as 1.5×10^8 Pa. The experimental reproducibility (i.e., the precision) of talc cells is better, and the inaccuracy can be improved to near $\pm 0.5 \times 10^8$ Pa by calibrating the cell against alkali halide cells. If experimental data are collected on the same press, the inaccuracy of the press affects only the regression line's intercept uncertainty. Because the slope of the regression line depends only on the relative position of the data in (T, P) or ($1/T$, ln K) space, a constant bias in the press cancels out of the slope calculation, making the uncertainty in the slope a function only of the imprecision of the press. Thus, different calibrations of the same reaction often have better slope agreement than intercept agreement. Unfortunately, press inaccuracy becomes important when data from different laboratories are regressed together, and then the full inaccuracy must be considered.

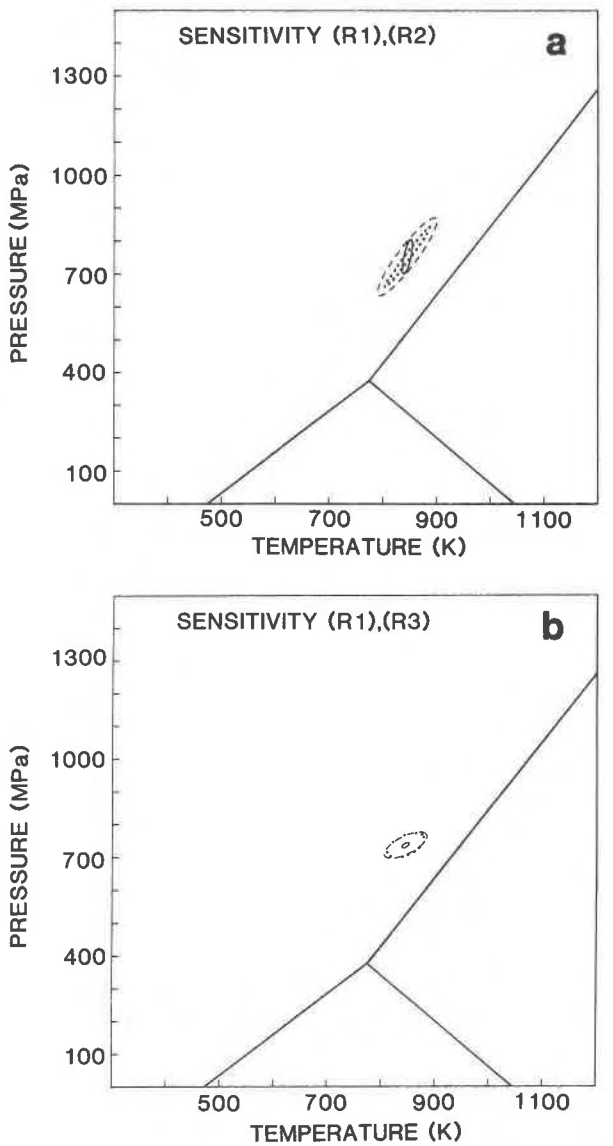

Fig. 3. (a) Sensitivity analysis for R1-R2 uncertainties. Solid ellipse assumes no uncertainty in garnet analysis. Dashed ellipse assumes no uncertainty in biotite analysis. Dotted ellipse assumes no uncertainty in plagioclase analysis. (b) Sensitivity analysis for R1-R3 uncertainties. Solid ellipse assumes no garnet uncertainty, dashed ellipse assumes no biotite uncertainty, and dotted ellipse assumes no ilmenite uncertainty.

Bracketing uncertainties for net-transfer reactions are calculated from the differences in run conditions (T_{BR},

P_{BR}) between a pair of runs. For two runs at (T_1, P_1) and (T_2, P_2),

$$T_{BR} = (T_1 - T_2)$$
$$P_{BR} = (P_1 - P_2). \quad (9)$$

Using Equation 4, the bracketing variances are

$$\mathrm{Var}_{T_{BR}} = \mathrm{Var}_{T_1} + \mathrm{Var}_{T_2}$$
$$\mathrm{Var}_{P_{BR}} = \mathrm{Var}_{P_1} + \mathrm{Var}_{P_2} \quad (10)$$

Assuming that the variances are equal at the two (T, P) conditions, then

$$s_{T_{BR}} = 1.4_{s_T}$$
$$s_{P_{BR}} = 1.4_{s_P} \quad (11)$$

where s_T and s_P are the experimental uncertainties for temperature and pressure, respectively. Thus, even when two runs nominally bracket an equilibrium *perfectly* (that is $T_1 = T_2$, $P_1 = P_2$), the bracket still has a $2s$ uncertainty of $2.8s_T$ and $2.8s_P$, due simply to the inaccuracy of the apparatus.

Given a variety of experimental uncertainties, we require a linear regression routine that will permit us to propagate these uncertainties into entropy and enthalpy uncertainties. For a similar problem involving empirical thermobarometric calibrations, Hodges and Crowley (1985) used the York (1969) treatment, which yields simple uncertainties in the calculated slope and Y intercept. The problem of experimental calibration involves an additional factor: many experiments are conducted at P-T conditions that greatly exceed the natural conditions pertinent to crustal thermobarometry. Intuitively, thermobarometers calibrated at more realistic P-T should be better than those calibrated at less realistic P-T, and this should be reflected in uncertainty calculations. Because

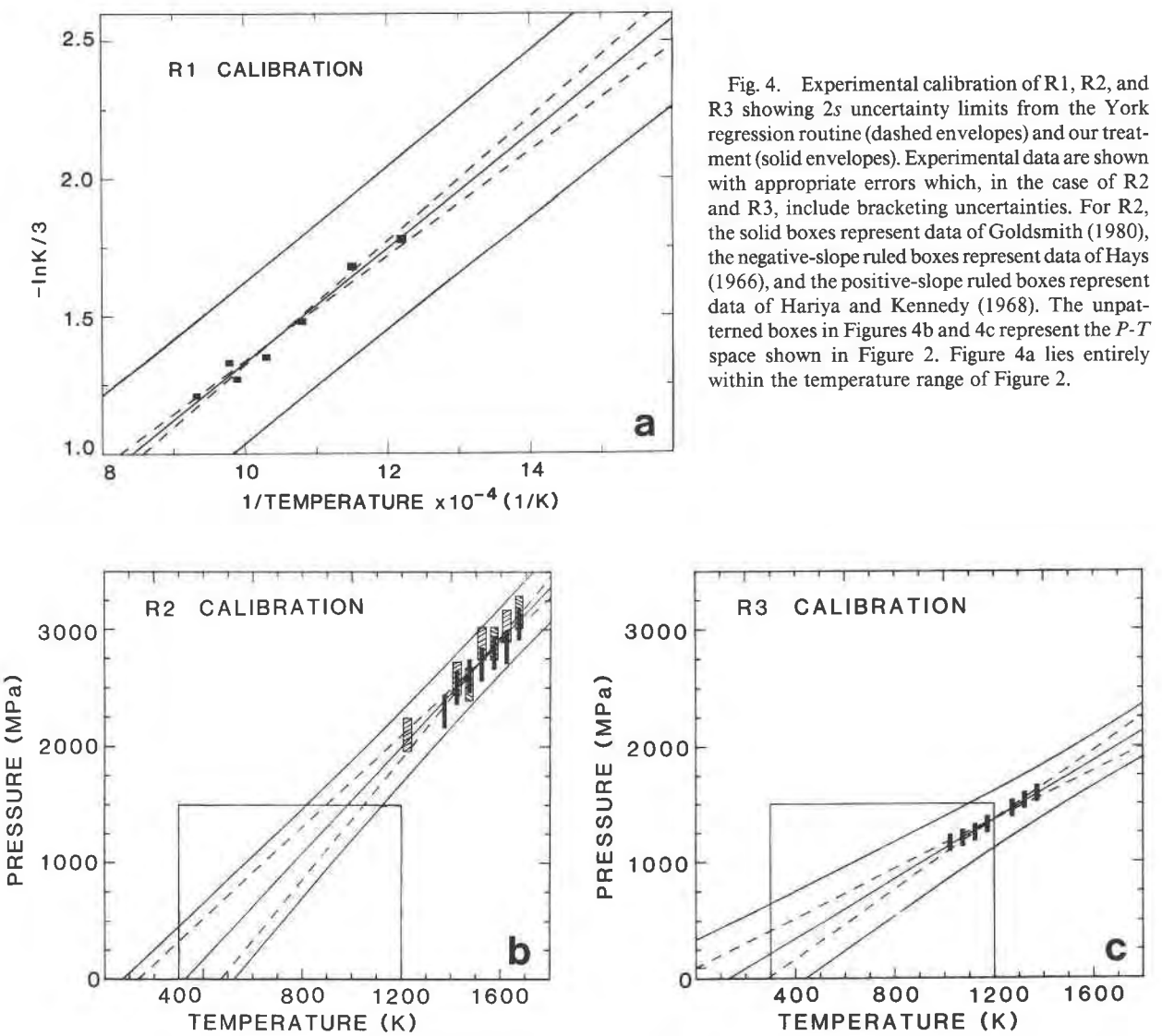

Fig. 4. Experimental calibration of R1, R2, and R3 showing $2s$ uncertainty limits from the York regression routine (dashed envelopes) and our treatment (solid envelopes). Experimental data are shown with appropriate errors which, in the case of R2 and R3, include bracketing uncertainties. For R2, the solid boxes represent data of Goldsmith (1980), the negative-slope ruled boxes represent data of Hays (1966), and the positive-slope ruled boxes represent data of Hariya and Kennedy (1968). The unpatterned boxes in Figures 4b and 4c represent the P-T space shown in Figure 2. Figure 4a lies entirely within the temperature range of Figure 2.

the York (1969) treatment does not account for this effect, we derive an explicit formula for the full uncertainty field defined by the experimental data. The regression line is defined by

$$\hat{Y}_k = \bar{Y} + M(X_k - \bar{X}), \tag{12}$$

where X_k is any X value, Y_k is the expected value of the regression at X_k, M is the slope, and $\bar{X}$ and $\bar{Y}$ are the average X and Y (Draper and Smith, 1966). Applying Equation 4 to 12 and noting that the variance of X_k is zero because X_k is an arbitrarily chosen variable, we find that

$$\begin{aligned}\mathrm{Var}_{\hat{Y}_k} = {} & \mathrm{Var}_{\bar{Y}} + (X_k - \bar{X})^2\mathrm{Var}_M + M^2\mathrm{Var}_{\bar{X}} \\ & + 2\rho_{\bar{Y}M}(X_k - \bar{X})\mathrm{Var}_{\bar{Y}}^{0.5}\mathrm{Var}_M^{0.5} \\ & + 2\rho_{\bar{Y}\bar{X}}(-M)\mathrm{Var}_{\bar{X}}^{0.5}\mathrm{Var}_{\bar{Y}}^{0.5} \\ & + 2\rho_{\bar{X}M}(-M)(X_k - \bar{X})\mathrm{Var}_M^{0.5}\mathrm{Var}_{\bar{X}}^{0.5}, \end{aligned} \tag{13}$$

where $\mathrm{Var}_{\hat{Y}_k}$ is the variance of the expected Y_k value from the regression line, Var_M is the variance of the slope, and $\mathrm{Var}_{\bar{X}}$ and $\mathrm{Var}_{\bar{Y}}$ are the variances of the mean X and Y. $\mathrm{Var}_{\bar{X}}$ and $\mathrm{Var}_{\bar{Y}}$ may be found by applying Equation 4 to Equation 1. The variance of the slope may be taken from the results of a standard York regression of an experimental data set. In effect, the first and third terms in Equation 13 account for the uncertainties in individual experimental runs. The second term accounts for the effect of the slope uncertainty on a Y_k estimate. It has a minimum of zero at $X_k = \bar{X}$ and increases symmetrically about $\bar{X}$ as a hyperbola in (X, Y) space. The next three terms in the equation account for covariances between Y, X, and M.

Once the variance of a calibration has been determined, it must be converted into a confidence interval. Through standard practice, the confidence interval is given at the 95% ("2s") level. The confidence interval has a full width of $2W$; W is given by the relation (after Larsen and Marx, 1981)

$$W = \mathrm{Var}_{\hat{Y}_k}^{0.5} t_{(\alpha/2, n-2)}, \tag{14}$$

where $t_{(\alpha/2,n-2)}$ is the Student's t distribution with $(n - 2)$ degrees of freedom at the $100(1 - \alpha)$ percent confidence level. For 95% uncertainties, $\alpha = 0.05$ and $t_{(\alpha/2,n-2)}$ ranges from ~13 ($n = 3$) to 1.96 ($n = \infty$).

The morphology of the confidence curves is dominated by the $\mathrm{Var}_{\bar{Y}}$ and $\mathrm{Var}_{\bar{X}}$ terms in the neighborhood of ($\bar{X}$, $\bar{Y}$), and the Var_M and covariance terms dominate where $|X_k - \bar{X}|$ is large. Because two of the covariance terms in Equation 13 are odd functions of $(X_k - \bar{X})$, they change signs at $X_k = \bar{X}$. This behavior leads to confidence curves that are asymmetric with respect to $\bar{X}$. The confidence interval for Y_k at $X_k = 0$ gives the Y-intercept uncertainty for the data array, or, in terms of Equations 8a and 8b, the value of s_B. As $\mathrm{Var}_{\hat{Y}_k}$ is a function of X_k, so too is s_B; the "resolved" uncertainty in ΔH (in 8a) and ΔS (in 8b) will vary as a function of T and $1/T$, respectively. The uncertainty is least at $T_k = \bar{T}$ or $(1/T)_k = \overline{(1/T)}$ and increases away from these values. As the thermodynamic uncertainties derived in this fashion are propagated into uncertainties in thermobarometric estimates of P-T, the effects are minimized when the calculated temperature of the sample coincides with the mean T of the experimental data, and the effects increase geometrically with increasing difference between the sample temperature and the experimental mean temperature.

TABLE 3. Bracketing data

T (K)	Hays (1966) P (GPa)	Hariya and Kennedy (1968) P (GPa)	Goldsmith (1980) P (GPa)	Bohlen et al. (1983) P (GPa)	Ferry and Spear (1978) (ln K)/3
1673	3.15(08)		3.05(05)		
1623		3.03(08)	2.85(05)		
1573	2.88(03)		2.80(10)		
1523	2.88(08)		2.70(10)		
1473	2.53(18)		2.60(10)		
1423		2.58(08)	2.50(10)		
1373			2.30(10)	1.59(02)	
1323				1.53(01)	
1273				1.46(01)	
1223		2.10(10)			
1173				1.32(02)	
1123				1.25(01)	
1073				1.20(01)	−1.21
1023				1.16(03)	−1.33
1011					−1.27
971					−1.35
924					−1.48
871					−1.68
823					−1.78
	Nominal bracketing uncertainties				
$2.8s_T$	17	11	7	7	
$2.8s_P$	0.14	0.07	0.07	0.07	
$2s_{1/T}$					5×10^{-6}
$2s_{\ln K}$					0.01

Figure 4 illustrates the confidence limits for experimental calibrations of R1 through R3 using the data of (1) Ferry and Spear (1978) for R1; (2) Hays (1966), Hariya and Kennedy (1968), and Goldsmith (1980) for R2; and (3) Bohlen et al. (1983) for R3. The pertinent data from these papers are summarized in Table 3. To these data, we assigned nominal 2s temperature and pressure uncertainties of 5 K and 0.5×10^8 Pa or the uncertainties given in the source if they were larger. The computed values for the input parameters of Equations 12 through 14 are given in Table 4. We derived covariance coefficients for R1 through R3 using a Monte Carlo approach. In this technique, we use each datum in the experimental calibration set and its corresponding uncertainty to derive a 200-element, normally distributed array. One value is chosen arbitrarily from each array, and a simple linear-regression analysis is performed on the resulting set of points. This procedure is repeated 1000 times in order to produce a suite of values for $\bar{Y}$, $\bar{X}$, and M. These are then regressed against each other, and the resulting correlation coefficients are used as the covariance coefficients in Equation 13. The values for these coefficients are listed in Table 4. For the three calibrations that we considered, only one covariance term ($\rho_{\bar{Y}M}$ in R3) is large enough to have a significant effect on $\mathrm{Var}_{\hat{Y}_k}$.

The most noticeable feature of Figure 4 is the (rather

TABLE 4. Regression results

Term	R1	R2	R3
n	7	14	7
t_{95}	2.57	2.18	2.57
M	-2.08×10^{3} K^{-1}	2.29×10^{6} Pa/K	1.28×10^{6} Pa/K
B	0.753	-7.47×10^{8} Pa	-1.67×10^{8} Pa
$\mathrm{Var}_{\bar{Y}}$	6.62×10^{-3}	1.12×10^{16} Pa^2	5.30×10^{15} Pa^2
$s^2_{\bar{Y}}$	4.61×10^{-2}	8.79×10^{16} Pa^2	2.85×10^{16} Pa^2
$\bar{Y}$	-1.44	2.71×10^{9} Pa	1.36×10^{9} Pa
Var_{M}	7.13×10^{3} K^{-2}	2.52×10^{10} (Pa/K)^2	1.19×10^{10} (Pa/K)^2
$\mathrm{Var}_{\bar{X}}$	1.50×10^{-9} K^{-2}	1.17×10^{3} K^2	2.48×10^{3} K^2
$s^2_{\bar{X}}$	1.05×10^{-8} K^{-2}	1.58×10^{4} K^2	1.74×10^{4} K^2
$\bar{X}$	1.05×10^{-3} K^{-1}	1.51×10^{3} K	1.19×10^{3} K
$\rho_{\bar{Y}M}$	0.16	0.05	-0.61
$\rho_{\bar{X}M}$	0.11	0.12	0.02
$\rho_{\bar{X}\bar{Y}}$	-0.011	-0.004	-0.006

disappointing) size of the uncertainty fields. The dashed lines indicate the much smaller uncertainty limits predicted by the standard York regression routine. The Var_M term in Equation 14 is within 5% of the York limits for all T_k for all three equilibria. At temperatures far removed from the mean temperatures of the data, this term accounts for between 30% (R2) and 12% (R1) of the uncertainty field. Near the mean temperatures, the effect of this term becomes negligible, and the other terms in Equation 14 combine to produce uncertainty limits that are somewhat greater than the 95% confidence limits of the data points themselves. The magnitude of the $2s$ uncertainty fields for these equilibria is striking when compared to the realm of crustal P-T space. The unfilled boxes in Figure 4 correspond to the P-T conditions represented in Figure 2. Table 5 gives the derived thermodynamic values, and their associated uncertainties, for R1, R2, and R3.

Accuracy of the Funeral Mountains P-T estimates

The thermodynamic uncertainties in Table 5 can be propagated into P-T uncertainties in much the same way that we propagated analytical uncertainties previously: ΔH and ΔS Monte Carlo arrays are created for each thermobarometer, and randomly selected elements are used to calculate an uncertainty ellipse for the P-T estimate. By combining both analytical and calibration uncertainties, we can produce an estimate of the minimum accuracy of a given thermobarometric calculation. Figure 5 shows the results for the Funeral Mountains sample; for R1-R3, $\omega_T = 138$ K and $\omega_P = 4.10 \times 10^8$ Pa, and for R1-R2, $\omega_T = 156$ K and $\omega_P = 5.16 \times 10^8$ Pa. Clearly, the accuracy limits in Figure 5 dominantly reflect calibration uncertainties.

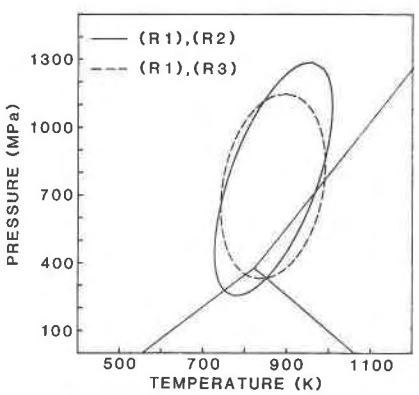

Fig. 5. Accuracy limits ($2s$) for the Funeral Mountains sample derived by combining analytical and calibration uncertainties.

Discussion

Thermobarometric studies are generally aimed at determining the actual equilibration temperature and depth of a sample (absolute thermobarometry), the relative equilibration temperatures and depths among a suite of samples (comparative thermobarometry), or both. For absolute thermobarometry, we are concerned with the accuracy of a P-T estimate, and Figure 5 is a sobering reflection on the state of the art. Equilibria R1, R2, and R3 are among the most commonly used and best-calibrated crustal thermobarometers. Nonetheless, at the 95% confidence level, these equilibria only broadly constrain the P-T conditions for the Funeral Mountains sample. If we are asked to estimate the equilibration depth of this

TABLE 5. Derived thermodynamic constants

Equilibrium R1

$\Delta V = 2.65 \times 10^{-6}$

$\Delta S = 18.8 \pm 64.1\{(1.31 \times 10^{-2}) + (7.13 \times 10^{3}) \cdot [(1/T) - (1.054 \times 10^{-3})]^2 + 3.70[(1/T) - (1.054 \times 10^{-3})]\}^{1/2}$

$\Delta H = (5.02 \times 10^{4}) \pm (0.54 \times 10^{4})$

Equilibrium R2

$\Delta V = 6.62 \times 10^{-5}$

$\Delta S = 150 \pm 23$

$\Delta H = (4.95 \times 10^{4}) \pm (1.44 \times 10^{-4})\{(1.74 \times 10^{16}) + (2.52 \times 10^{10})(T - 1510)^2 - (1.30 \times 10^{12})(T - 1510)\}^{1/2}$

Equilibrium R3

$\Delta V = -1.28 \times 10^{-5}$

$\Delta S = -12.3 \pm 5.2$

$\Delta H = (3.66 \times 10^{3}) \pm (4.79 \times 10^{-5})\{(9.43 \times 10^{15}) + (1.19 \times 10^{10})(T - 1190)^2 - (9.98 \times 10^{12})(T - 1190)\}^{1/2}$

Note: V, S, and H are in m^3/mol [= J/(Pa·mol)], J/(mol·K), and J/mol, respectively. Volume data from Lang and Rice (1985). All uncertainties represent 95% confidence limits.

sample, we are forced to report the estimate with a minimum $2s$ uncertainty of 10 to 20 km. Despite this unfortunate truth, it is noteworthy that these equilibria usually yield P-T estimates that are consistent with independent indicators of metamorphic conditions (e.g., the aluminosilicate stability fields), and it seems likely that these thermobarometers are better than statistical rigor leads us to believe.

In a sense, the uncertainty limits in Figure 4 make a stronger statement about calibration techniques than about the inherent quality of specific thermobarometers. Statistically meaningful absolute thermobarometry requires some modification of commonly accepted calibration procedures. The form of Equation 14 suggests three changes that could significantly improve thermobarometric accuracy: (1) increasing the number of runs in the calibration data set; (2) increasing the P-T range of the calibration data set, and (3) increasing the P and T precision of the experimental presses.

Because the $\mathrm{Var}_{\hat{x}}$ and $\mathrm{Var}_{\hat{y}}$ terms in Equation 14 scale as n^{-1}, the magnitude of these terms is reduced by increasing the number of statistically independent bracketing experiments in the calibration data set. By "statistically independent," we mean experiments run at separate times or in separate presses, even if the experiments duplicate P-T conditions of other experiments. Although duplicating the P-T conditions of runs will not improve the York regression statistics, increasing the number of bracketing runs at a given temperature increases the likelihood that the pressure bracket lies on the modeled regression line.

Extensions of the P-T range of calibration experiments affects uncertainties in two ways. First, the Var_M term in Equation 14 changes in direct proportion to the ratio between the data uncertainty and the temperature range of the experiments. Second, extending the calibration to include crustally accessible P-T space ensures that the univariant reaction has been fitted near the conditions at which it will be used, reducing extrapolation errors. Although low-temperature experiments can be extremely time-consuming, their practical value is significant.

It is also apparent from Figure 4 that no P determination can be better than the bracketing uncertainty of the calibration data. In the limit where both the number and P-T range of the calibration experiments become large, Equation 13 shows that

$$\lim_{n\to\infty}(\mathrm{Var}_{\hat{y}_k}) = (1.4s_P)^2 + M^2(1.4s_T)^2 - 2\rho_{\hat{x}\hat{y}}M(1.4s_P)(1.4s_T). \quad (15)$$

For many experimental calibrations, bracketing uncertainies are ~100 MPa and make a major contribution to the overall uncertainty of the thermobarometer. Clearly, a conscientious effort to reduce these uncertainties is one of the surest and most cost-effective ways to improve the accuracy of thermobarometry.

For comparative thermobarometry involving a single pair of equilibria, the systematic errors associated with the calibrations cancel, leaving only the effects of analytical uncertainties (Fig. 2). Carefully analyzed samples yield $2s$ precision limits for R1-R3 equilibria sufficient to distinguish between equilibration conditions differing by as little as 70 K or 70 MPa. Thus, comparative thermobarometry is a statistically valid means of studying tectonic processes on length scales exceeding 3 to 5 km.

Acknowledgments

We would like to thank Steve Bohlen for useful discussions concerning the precision of experimental apparatus, Dave Gutzler for an invaluable review of the statistics, Peter Tilke for providing computational insights, and Jane Selverstone and Jim Munoz for reviews of the entire text. This work was supported by National Science Foundation Grants EAR 8319768 and EAR 8407730.

References

Anderson, G.M. (1976) Error propagation by the Monte Carlo method in geochemical calculations. Geochimica et Cosmochimica Acta, 40, 1533–1538.

——— (1977a) Uncertainties in calculations involving thermodynamic data. In H.J. Greenwood, Ed., Short course in application of thermodynamics to petrology and ore deposits, p. 199–215. Mineralogical Association of Canada.

——— (1977b) The accuracy and precision of calculated mineral dehydration equilibria. In D.G. Fraser, Ed., Thermodynamics in geology, p. 135–136. NATO Advanced Studies Institute, Series C, Reidel, Boston.

Bohlen, S.R., Wall, V.J., and Boettcher, A.L. (1983) Experimental investigations and geological applications of equilibria in the system FeO-TiO_2-Al_2O_3-SiO_2-H_2O. American Mineralogist, 68, 1049–1058.

Carpenter, M.A., and Putnis, A. (1986) Cation order and disorder during crystal growth: Some implications for natural mineral assemblages. In A.B. Thompson and D.C. Rubie, Eds., Metamorphic reactions, kinetics, textures and deformation, p. 1–26. Springer-Verlag, New York.

Draper, N.R., and Smith, H. (1966) Applied regression analysis. Wiley, New York.

Essene, E.J. (1982) Geologic thermometry and barometry. Mineralogical Society of America Reviews in Mineralogy, 10, 153–206.

Eugster, H.P., Albee, A.L., Bence, A.E., Thompson, J.B., and Waldbaum, D.R. (1972) The two phase region and excess mixing properties of paragonite-muscovite crystalline solutions. Journal of Petrology, 13, 147–179.

Ferry, J.M. (1980) A comparative study of geothermometers and geobarometers in pelitic schists from southern-central Maine. American Mineralogist, 65, 720–732.

Ferry, J.M., and Spear, F.S. (1978) Experimental calibration of the partitioning of Fe and Mg between biotite and garnet. Contributions to Mineralogy and Petrology, 66, 113–117.

Ganguly, J., and Kennedy, G.C. (1974) The energetics of natural garnet solid solution I. Mixing of the aluminosilicate end-members. Contributions to Mineralogy and Petrology, 48, 137–148.

Ganguly, J., and Saxena, S.K. (1984) Mixing properties of aluminosilicate garnets: Constraints from natural and experimental data, and applications to geothermo-barometry. American Mineralogist, 69, 88–97.

Ghent, E.D., Robbins, D.B., and Stout, M.Z. (1979) Geothermometry, geobarometry, and fluid compositions of metamorphosed calc-silicates and pelites, Mica Creek, British Columbia. American Mineralogist, 64, 874–885.

Ghent, E.D., Knitter, C.C., Raeside, R.P., and Stout, M.Z. (1982) Geothermometry and geobarometry of pelitic rocks, upper kyanite and sillimanite zones, Mica Creek area, British Columbia. Canadian Mineralogist, 20, 295–305.

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 Crowley, P. (1985) Error estimation and empirical geothermobarometry for pelitic systems. American Mineralogist, 70, 702–709.

Hodges, K.V., and Royden, L. (1984) Geologic thermobarometry of retrograded metamorphic rocks: An indication of the uplift trajectory of a portion of the northern Scandinavian Caledonides. Journal of Geophysical Research, 89, 7077–7090.

Hodges, K.V., and Spear, F.S. (1981) Geothermometry, geobarometry, garnet closure temperatures and the Al_2SiO_5 triple point, Mt. Moosilauke, New Hampshire (abs.). EOS, 62, 1060.

——— (1982) Geothermometry, geobarometry and the Al_2SiO_5 triple point at Mt. Moosilauke, New Hampshire. American Mineralogist, 67, 1118–1134.

Holdaway, M.J. (1971) Stability of andalusite and the aluminosilicate phase diagram. American Journal of Science, 271, 97–131.

Hollister, L.S. (1979) Metamorphism and crustal displacements: New insights. Episodes, 3, 3–8.

Johannes, W. (1978) Pressure comparing experiments with NaCl, AgCl, talc and pyrophyllite assemblies in a piston cylinder apparatus. Neues Jahrbuch für Mineralogie Monatshefte, 2, 84–92.

Johannes, W., Bell, P.M., Mao, H.K., Boettcher, A.L., Chipman, D.W., Hays, J.F., Newton, R.C., and Seifert, F. (1971) An interlaboratory comparison of piston-cylinder pressure calibration using the albite breakdown reaction. Contributions to Mineralogy and Petrology, 32, 24–38.

Labotka, T.C. (1980) Petrology of a medium-pressure regional metamorphic terrane, Funeral Mountains, California. American Mineralogist, 65, 670–689.

Lang, H.M., and Rice, J.M. (1985) Geothermometry, geobarometry and T-X (Mg-Fe) relations in metapelites, Snow Peak, Northern Idaho. Journal of Petrology, 26, 889–924.

Larsen, R.J., and Marx, M.L. (1981) An introduction to mathematical statistics and its applications. Prentice-Hall, Englewood Cliffs, New Jersey.

Loomis, T.P. (1976) Irreversible reactions in high-grade metamorphic rocks. Journal of Petrology, 17, 559–588.

——— (1983) Compositional zoning of crystals: A record of growth and reaction history. In S.K. Saxena, Ed., Kinetics and equilibrium in mineral reactions, p. 1–60. Springer-Verlag, New York.

Newton, R.C., and Haselton, H.T. (1981) Thermodynamics of the garnet-plagioclase-Al_2SiO_5-quartz geobarometer. In R.C. Newton et al., Eds., Thermodynamics of minerals and melts, p. 131–147. Springer-Verlag, New York.

Newton, R.C., Charlu, T.V., and Kleppa, O.J. (1977) Thermochemistry of high pressure garnets and clinopyroxenes in the system CaO-MgO-Al_2O_3-SiO_2. Geochimica et Cosmochimica Acta, 41, 369–377.

Powell, R. (1985) Geothermometry and geobarometry: A discussion. Geological Society of London Journal 142, 29–38.

Robie, R.A., Hemingway, B.S., and Fisher, J.R. (1978) Thermodynamic properties of minerals and related substances at 298.13 K and 1 bar (10^5 pascals) pressure and at higher temperatures. U.S. Geological Survey Bulletin 1452.

Schmid, R., Cressey, G., and Wood, B.J. (1978) Experimental determination of univariant equilibria using divariant solid solution and assemblages. American Mineralogist, 63, 511–515.

Selverstone, J. (1985) Petrologic constraints on imbrication, metamorphism, and uplift in the SW Alps, Tauern Window, eastern Alps. Tectonics, 4, 686–704.

Smith, D.G.W., Ed. (1976) Short course in microbeam techniques. Mineralogical Association of Canada, 186 p.

Steltenpohl, M., and Bartley, J.M. (1984) Statistical evaluation of P-T trajectories based on elemental partitioning geothermometers and geobarometers. Geologic Society of America Abstracts with Programs, 16, 667.

Tracy, R.J., and Robinson, P. (1980) Evolution of metamorphic belts: Information from detailed petrologic studies. In D.J. Wones, Ed., The Caledonides in the USA, 189–195. Virginia Polytechnic Institute and State University, Memoir 2.

Tracy, R.J., Robinson, P., and Thompson, A.B. (1976) Garnet composition and zoning in the determination of temperature and pressure of metamorphism, central Massachusetts. American Mineralogist, 61, 762–775.

York, D. (1969) Least-squares fitting of a straight line with correlated errors. Earth and Planetary Science Letters, 5, 320–324.

Manuscript received September 29, 1986
Manuscript accepted April 3, 1987