

Error propagation for barometers: 1. Accuracy and precision of experimentally located end-member reactions

MATTHEW J. KOHN, FRANK S. SPEAR

Department of Geology, Rensselaer Polytechnic Institute, Troy, New York 12180, U.S.A.

ABSTRACT

A new procedure for calculating the precision and accuracy of end-member reactions is presented that accounts for the kinetics of reaction progress and possible displacements between the experimental and true *P-T* reference frames. Application to recently collected experimental data indicates that the estimated precision (1σ) of the position of an end-member reaction at metamorphic temperatures (~ 900 K) using a single set of experimental data may be ± 100 to ± 310 bars and that a minimum estimate of accuracy (1σ) may be on the order of ± 270 to ± 410 bars, depending on how precisely the reaction is bracketed and assuming 1σ potential systematic errors in pressure and temperature of ± 250 bars and ± 5 °C, respectively. The accuracy of the grossular-anorthite-kyanite-quartz (GASP) reaction position using five different studies is estimated to be between ± 350 and ± 380 bars (1σ) at 900 K, depending on whether the individual studies were calibrated using independent standards.

INTRODUCTION

The accuracy and the precision of end-member reactions are of fundamental importance for understanding the uncertainty of both barometer calibrations and pressure estimates made on natural rocks. Although several methods have been presented that purport to estimate the precision of reaction positions (e.g., Demarest and Haselton, 1981; Hodges and McKenna, 1987; Powell and Holland, 1985; Berman et al., 1986), none of these methods simultaneously accounts for experimental bracket size, the reproducibility of experimental conditions, and the kinetics of disequilibrium reaction progress. Similarly, although Hodges and McKenna (1987) and McKenna and Hodges (1988) have described a method whereby the accuracy of an end-member reaction may be estimated, their results are inconsistent with the magnitudes of the potential systematic errors that have been reported for experimental devices.

Generally there are three types of errors that must be considered: (1) the precision with which a reaction has been bracketed, (2) systematic errors that result from uncertainty in the position of the standard reactions used for calibration, and (3) interlaboratory, procedural, and starting material variability. Each one of these error propagates to substantial pressure uncertainties.

The purposes of this paper are (1) to present a new method for calculating the precision of an experimentally located reaction constrained by a single set of data that is more in accord with bracket sizes and kinetic theory, (2) to describe the minimum uncertainty of an end-member reaction constrained by only a single set of data by propagating systematic experimental uncertainties reported in the literature, and (3) to present a technique for determining a best-fit position of a reaction and assigning

realistic uncertainties to that position when there are several sets of experimental data representing different laboratories and starting materials.

In this paper, we have followed the general terminology of Hodges and McKenna (1987), and the term "precision" will refer to reproducibility or randomly distributed errors, as contrasted to "systematic errors." The term "accuracy" will be reserved to describe the combination of precision and systematic uncertainties.

PRECISION OF REACTIONS: SINGLE DATA SETS

Kinetic theory

Two different approaches have been typically used to estimate the precision of reaction positions: linear (or mathematical) programming and linear regression. Linear programming relies on the assumption that for any single experimental bracket, the position of the reaction is equally likely at any point within the bracket. That is, experiments only contain information about the direction of equilibrium, not its distance from the experimental conditions. Use of linear regression implies that the most likely position for a reaction is in the middle of a bracket.

Transition state theory (e.g., Fisher and Lasaga, 1981; Lasaga, 1981b) provides a useful theoretical framework within which to evaluate the assumptions inherent in the two techniques for estimating reaction precisions. An important aspect of this theory revolves around the mathematical description of the means by which reactant is transformed to product. For many reactions, this transformation occurs via an activated complex, which has a higher Gibbs free energy than either the reactant or product. The difference between the energy of the activated complex and that of the reactant is referred to as the activation energy.

Because we know of no studies that have investigated quantitatively experimental reaction kinetics for pressure sensitive equilibria, we will assume that the activated complex for the forward and reverse reactions is the same. If this assumption is made, then transition state theory predicts that the reaction rate will have the form:

$$R_{\text{net}} = R_+ + R_- = R_+ \left[1 - \exp\left(\frac{nA}{RT}\right) \right] \quad (1a)$$

where R_{net} is the net reaction rate in units of concentration per time, R_+ and R_- are the absolute reaction rates in the forward and reverse directions, RT is the gas constant times temperature in kelvins, A is the reaction affinity, or the absolute Gibbs free energy change between reactant and product that drives the reaction, and n is a constant of order one (Fisher and Lasaga, 1981; Lasaga, 1986). We note that R_{net} may have a different functional dependence on A if defects dominate the kinetics (e.g., Lasaga, 1986), but that data collected by A. Koziol (personal communication) for the reaction grossular-kyanite-quartz-plagioclase (GASP) favor Equation 1a. In the case of the bracketing of pressure sensitive equilibria (i.e., at a specific temperature), A equals the absolute value of $\Delta P\Delta V$, where ΔV is the volume difference between reactants and products and ΔP is the difference between the experimental and equilibrium pressures:

$$R_{\text{net}} = R_+ \left[1 - \exp\left(\frac{n\Delta P\Delta V}{RT}\right) \right] \quad (1b)$$

R_+ is generally a complicated function of concentration and temperature, so that R_{net} depends on the duration of the experiment. This means that the magnitude of A and hence of the difference between experimental and equilibrium P - T conditions is best judged from the initial R_{net} , which is usually estimated using many experiments conducted at the same P - T conditions for different lengths of time. Fortunately, several experimental studies (e.g., Schramke et al., 1987; Heinrich et al., 1989) indicate that, even for relatively large changes in the number of moles of the phase used to monitor reaction progress, R_+ is nearly constant; that is, the extrapolated initial R_{net} is often equal to the measured R_{net} . R_+ will be assumed to be essentially constant for the small to moderate amounts of reaction considered below.

We know of no theoretical or empirical considerations that suggest that pressure has a significant effect on R_+ ; therefore, for a given T and with the assumptions above regarding the degree of reaction progress, the magnitude of n , and the nature of the activated complex, the magnitude of R_{net} is simply dependent on the amount of overstepping ($\Delta P\Delta V$) and will be symmetric about the true reaction position for brackets of pressure-sensitive equilibria. This means that the most likely position for a reaction that is bracketed by two experiments that (1) show equal but opposite degrees of reaction progress and (2) were conducted for the same amount of time, will be at the midpoint of the bracket.

Although few brackets meet these qualifications, Equation 1b also predicts that any two experiments conducted at different pressures with reaction progress less than 100% should uniquely define the position of the equilibrium. The true imprecision in the position of a reaction based on such experiments then depends on how well reaction progress can be measured, and as experimentalists become better at differentiating between small amounts of reaction progress, reaction positions should become increasingly more precise. It is worth noting, however, that (1) if one of the reactants or products has been entirely consumed during the experiment, then no unique solution for the reaction position is obtainable, and (2) because R_+ begins to deviate from constancy for large degrees of reaction progress (>40% to 75%) at a given pressure, temperature, and composition, estimating the initial R_{net} in such a case may require extra experiments for different amounts of time.

The transition state theory as applied to experimental reaction progress has profound implications on whether linear programming or linear regression is better suited for estimating the precision of a reaction position from experimental brackets. If experimental brackets represent 100% reaction progress, then no best estimate of the reaction position is obtainable based on kinetic theory, and a linear programming approach is probably better. On the other hand, if at least two experiments that represent partial reaction progress are performed at the same temperature but at different pressures, then an estimate of the most likely reaction position is possible from kinetic theory, and linear regression is probably better.

Because experimental data collected in the last few years for pressure-sensitive equilibria increasingly represent several experiments at a single temperature that exhibit only partial reaction progress, it seems likely that linear regression analysis will become the method of choice for calculating individual reaction positions and their precisions. However, to justify this approach it is important for experimentalists to report kinetic data, including masses and average grain sizes of all phases in the starting mixture and products, along with the conditions and duration of the experiments. Furthermore, investigations of reaction rate as a function of pressure overstepping are required either to verify that transition state theory is an adequate model or to develop a better kinetic model for pressure-sensitive equilibria. In the remainder of this paper, we will rely on linear regression for estimating reaction positions and their uncertainties. For simplicity we restrict discussion to experimental reversal data, in which each reversal is composed of two experiments conducted at the same temperature that show opposite reaction directions.

An example using transition state theory to estimate reaction position

Suppose we have collected the hypothetical experimental data for the GASP reaction as listed in Table 1 and wish to calculate the most likely reaction position and its

TABLE 1. Hypothetical experimental data for GASP reaction

T (K)	P (kbar)	σ_p (kbar)	r_{final}	$\sigma_{r_{\text{final}}}$	R_{net} (mole/cm ² ·h)	$\sigma_{R_{\text{net}}}$ (mole/cm ² ·h)
1223	21.00	0.1	1.3	0.1	1.3×10^{-7}	4×10^{-8}
1223	20.00	0.1	0.75	0.06	-1.4×10^{-7}	4×10^{-8}

Note: The following assumptions were made: duration of each experiment was 20 h, initial moles of grossular was 2×10^{-4} (90 mg), the initial peak height intensity ratio was 1.0, and the surface area of grossular was 10 cm² (average grain radius of 75 μm). The term r_{final} is the measured peak height intensity ratio (Grs/An) of the product, and $\sigma_{r_{\text{final}}}$ is the standard deviation of the measurement of r_{final} . R_{net} is calculated from Equation 3b; $\sigma_{R_{\text{net}}}$ calculated from propagation of $\sigma_{r_{\text{final}}}$ through Equation 3b.

uncertainty. For the initial molar quantities of reactant and product ($m_{\text{rci,init}}$ and $m_{\text{prod,init}}$), the rate of change in the number of moles of the phase j may be calculated from the equation:

net rate of production or consumption of product

$$= \frac{(r_{\text{final}} - r_{\text{init}})\nu_{\text{prod}}/t}{(\nu_{\text{rci}}r_{\text{init}}/m_{\text{rci,init}}) + (\nu_{\text{prod}}r_{\text{final}}/m_{\text{prod,init}})} \quad (2)$$

where ν_{rci} and ν_{prod} are the absolute values of the stoichiometric coefficients of the reactant and product in the reaction of interest (i.e., 3 and 1 for anorthite and grossular in the GASP reaction), t is the duration of the experiment, and r_{final} and r_{init} are the peak height intensity ratios (product/reactant) of the final product and the initial starting material. We note here that although we have assumed in our example that reaction progress is monitored using X-ray diffraction methods, the ratios in Equation 2 are perfectly general, and other techniques of measuring reaction progress may also be used (e.g., SEM imaging).

Equation 2 simplifies significantly if the starting materials are mixed in stoichiometric proportions (i.e., $m_{\text{rci,init}} = \nu_{\text{rci}}m_{\text{prod,init}}/\nu_{\text{prod}}$):

$$\text{net rate of production or consumption of product} = \frac{(r_{\text{final}} - r_{\text{init}})m_{\text{prod,init}}/t}{r_{\text{init}} + r_{\text{final}}} \quad (3a)$$

Net reaction rates in terms of changes of concentration per time may be calculated by normalizing Equation 3a to an extensive property of the system. For heterogeneous reactions (Lasaga, 1981a) such as the GASP reaction, the important parameter is the surface area (S) of one of the phases in the experiment, usually one with a low solubility:

$$\frac{\text{net rate of production or consumption of product}}{S} = \frac{(r_{\text{final}} - r_{\text{init}})m_{\text{prod,init}}/S}{(r_{\text{init}} + r_{\text{final}})t} \quad (3b)$$

If (1) only the reaction position and its uncertainty are desired, and (2) the composition of the starting mix and duration are the same for all the experiments, and noting that surface areas are linearly proportional to the amount

of starting material, then the terms $m_{\text{prod,init}}$, t , and S are not required, as they do not affect the magnitude of A .

An example of the application of Equation 3b to a set of hypothetical data is given in Table 1. We first assumed that there were stoichiometric proportions of grossular and anorthite in the starting material, the initial moles of grossular ($m_{\text{prod,init}}$) was 2×10^{-4} (90 mg), the initial peak height intensity ratio (r_{init}) was 1.0, the durations (t) were 20 h, and the surface area of grossular (S) was 10 cm² (average grain radius of 75 μm). Using these values and the hypothetical data in Table 1 for T , P , r_{final} , and $\sigma_{r_{\text{final}}}$, the measured reaction rates and their uncertainties may be calculated from Equation 3b, and are tabulated in Table 1. From the measured reaction rates, Equation 1b can be solved using a Newton-Raphson procedure to obtain the equilibrium pressure, which is 20527 bars. The best-fit reaction position is at a slightly higher pressure than the midpoint of the hypothetical bracket because the reaction rate was faster for the lower pressure reaction; therefore, this fictive experiment must have been farther from the equilibrium position than the higher pressure experiment.

Uncertainties in the reaction position may be estimated using a Monte Carlo approach (Anderson, 1976) or numerical error analysis (Roddick, 1987). For the data in Table 1, either of these procedures results in an estimated uncertainty of ± 145 bars (1σ), of which about ± 70 bars represents the contribution of σ_p , while approximately ± 125 bars represents the effect of $\sigma_{r_{\text{final}}}$ (Table 1).

The calculated uncertainty is much smaller than that estimated using the method described by Demarest and Haselton (1981), which considers the reaction position uncertainty to be a function of the full bracket width as well as the experimental pressure precision. For the data in Table 1, the approach of Demarest and Haselton results in an estimated reaction position of 20500 bars with an uncertainty of ± 305 bars (1σ). Thus, transition state theory can allow a much more precise estimate of the reaction position than would be suggested by Demarest and Haselton (1981).

A physical depiction of the results above is shown in Figure 1 for the hypothetical data. This figure shows the data and uncertainties in the the pressure and reaction rate plotted on a normalized reaction rate vs. pressure diagram. The three curves show the best-fit and bounding functions predicted for the data by transition state theory. The 2σ confidence limits for the equilibrium pressure are calculated from the intersection of the bounding curves and the horizontal line defined by $R_{\text{net}}/R_+ = 0$. The 2σ confidence limits calculated from the technique of Demarest and Haselton (1981) are also presented and are clearly much larger than those estimated using transition state theory.

Regression analysis: an example

Regression analysis of the precision of bracketed equilibria has been described by several different workers (e.g., Demarest and Haselton, 1981; Powell and Holland, 1985; Hodges and McKenna, 1987). Many workers choose the

midpoints of brackets when regressing experimental data and then weight each point according to the size of the bracket. The differences among various regression approaches described in the literature often lies in how uncertainties are assigned to a single bracket (e.g., Demarest and Haselton, 1981; Powell and Holland, 1985) and whether or not uncertainties in temperature are accounted for (e.g., Hodges and McKenna, 1987). Other differences include the choice of regression type, some workers preferring a weighted least squares method (e.g., Powell and Holland, 1985), while others have used a Monte Carlo approach (e.g., Hodges and Crowley, 1985). As shown below, the variations in the assignment of uncertainty to each datum have a nontrivial effect on the calculated reaction position.

We believe it is quite important that for many experimental data sets, the bracket midpoints are much more colinear than is consistent with the error assigned to each bracket according to the bracket width and experimental reproducibility. In such cases, estimates of the reaction position uncertainty based on the scatter of the data about the best-fit line will be erroneously small, and should be calculated from the uncertainties assigned to the data themselves (see Draper and Smith, 1981, p. 83). This approach has been used for several of the data sets considered below.

As an exercise in regression analysis, we considered the effects of various regression and weighting approaches on the calculated slope, intercept, and their uncertainties for the data of Koziol and Newton (1988). Specifically, we compared the results of an ordinary weighted least squares regression using the uncertainties calculated from transition state theory (TST, see below) with those obtained using the weights suggested by Demarest and Haselton (1981: DH) and by Powell and Holland (1985: PH), as well as with those of a York (1969: YK) regression recommended by Hodges and McKenna (1987) (Table 2).

Transition state theory may be applied to the data of Koziol and Newton (1988) for the GASP reaction to provide a set of nominal reaction positions and their uncertainties on which regressions can be performed. We first assumed that (1) grossular and anorthite were initially mixed in stoichiometric proportions (Koziol and Newton, 1988), so that Equations 3a and 3b could be applied, and (2) any dissolution of grossular and anorthite into the melt phase present in the experiments did not affect their relative molar abundances. Based on grossular and anorthite X-ray data for the experiments conducted at two temperatures and general descriptions of all experiments (A. Koziol, personal communication), we concluded that most bracketing experiments represented between 10% and 40% reaction progress. The assumption of stoichiometric proportions of the starting material does not strongly affect the reaction progress calculations, and the minimum acceptable change in X-ray peak intensity ratios (Koziol and Newton, 1988) implies reaction progress of at least 13%. Finally, we assumed that the pressures of the experiments could be reproduced within ± 100 bars; a poorer reproducibility is difficult to reconcile with the

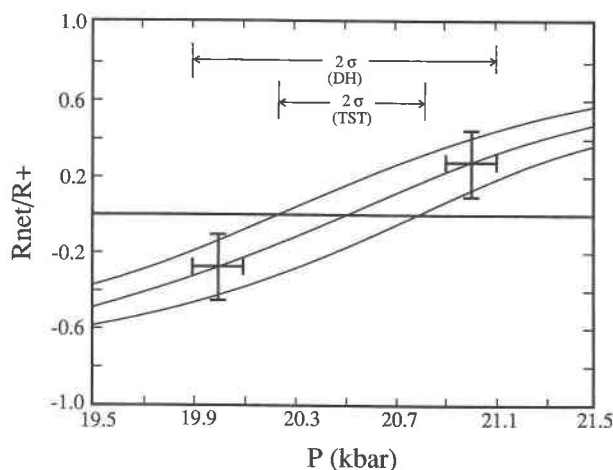


Fig. 1. Plot of R_{net}/R_+ vs. pressure for the hypothetical data in Table 1. The curves show the functions that were fit to the midpoints of the data (middle curve) and the 2σ bounds on those data using transition state theory (TST). Also shown are the uncertainties calculated using TST and the technique of Demarest and Haselton (1981) (DH).

internal consistency of the data. A nominal reaction position and its uncertainty were then calculated for each bracket as was done above for the hypothetical data.

Powell and Holland (1985) suggest assigning the bracket midpoint as the nominal reaction position and then assigning an uncertainty (1σ) equal to the bracket quarterwidth, although it should be understood that they were only attempting to obtain the reaction intercept position, not its slope; Hodges and McKenna (1987) suggest that temperature imprecisions may be important when estimating slope and intercept uncertainties. Because Hodges and McKenna (1987) did not state their preferred weighting scheme for pressures, we simply used the uncertainties calculated from transition state theory; this allows a direct evaluation of the importance of accounting for experimental temperature imprecisions for this data set. Temperature imprecisions (1σ) were assumed to be ± 5 °C (A. Koziol, personal communication); this value is twice that suggested by Hodges and McKenna (1987). An ordinary weighted least squares method was used for the PH and DH regressions.

As can be seen in Table 2, the TST, PH, and YK standard errors for slope and intercept are quite similar, and each regression indicates similar uncertainties in the best-fit reaction position. Ignoring temperature imprecisions (TST vs. YK) causes an underestimation of the reaction position uncertainty of less than ± 30 bars over mid to lower crustal temperatures; if we had assumed the same 2.5 °C uncertainty as suggested by Hodges and McKenna (1987), then the YK regression would actually have indicated smaller uncertainties than the TST. Thus, we see little practical advantage in accounting for temperature imprecisions with data collected on modern experimental devices for pressure sensitive reactions (slopes less than about 25 bars/°C), although for later analysis, we accept

TABLE 2. Selected regression results for GASP data sets

	Intercept $\pm 1\sigma$	Slope $\pm 1\sigma$	Correlation coefficient
Koziol and Newton (1988)			
TST:	-7253 \pm 817	22.70 \pm 0.59	-0.9977
YK:	-7222 \pm 889	22.68 \pm 0.65	-0.9977
DH:	-6128 \pm 1215	21.94 \pm 0.88	-0.9973
PH:	-5873 \pm 826	21.76 \pm 0.60	-0.9980
Gasparik (1985)			
DH:	-3617 \pm 3038	19.97 \pm 1.98	-0.9987
YK:	-3641 \pm 3256	19.98 \pm 2.11	-0.9987
Gasparik (1984)			
DH:	-5689 \pm 1535	21.31 \pm 0.99	-0.9979
YK:	-5743 \pm 1710	21.35 \pm 1.11	-0.9979
Goldsmith (1980)			
DH:	-7373 \pm 2845	22.43 \pm 1.80	-0.9980
YK:	-7438 \pm 2835	22.48 \pm 1.80	-0.9980
Hays (1966)			
DH:	-7481 \pm 11496	23.18 \pm 7.28	-0.9996
YK:	-7964 \pm 11664	23.50 \pm 7.40	-0.9996

Note: TST, DH, and PH refer to regressions weighted according to transition state theory, Demarest and Haselton (1981), and Powell and Holland (1985) respectively. YK refers to regression using the technique of York (1969).

the YK results in Table 2 as the best estimate of the uncertainty in reaction position. The DH standard errors are approximately 1.5 times larger than those calculated using the other approaches as a result of the larger uncertainties assigned to each datum. Because the PH and DH slopes and intercepts are somewhat different from the results of the other two regressions, predicted pressures differ by about 100–300 bars at the temperatures of the experiments, and by about 350–650 bars at crustal temperatures; the predicted pressures using the PH and TST regression are the most different of those considered.

Precision of reaction position

The precision of the reaction position may be calculated by propagating the uncertainties in the regressed slope and intercept and constructing corresponding 2σ confidence limits. For a line defined by the equation:

$$Y = mX + b \quad (4)$$

where m is the slope and b is the intercept, the propagated uncertainty in Y may be calculated as:

$$\sigma_Y^2 = \sigma_b^2 + (X\sigma_m)^2 + 2\rho_{mb}\sigma_b X\sigma_m \quad (5)$$

where σ_b and σ_m are the uncertainties in the intercept and slope respectively, and ρ_{mb} is the correlation coefficient between m and b .

The 2σ confidence limits for the reaction position resulting from the propagated YK regression uncertainties and correlation coefficient are shown in Figure 2 for the data of Koziol and Newton (1988) (inner bounds) with the experimental brackets for reference. As can be seen, the limits are well within seven of the brackets and are slightly inconsistent (although within assigned experimental reproducibility) with only one very narrow bracket. The estimated precision of the reaction position using the YK regression is about ± 60 bars (1σ) at 1350 K, near the midpoint of the data and ± 310 bars (1σ) at 900 K.

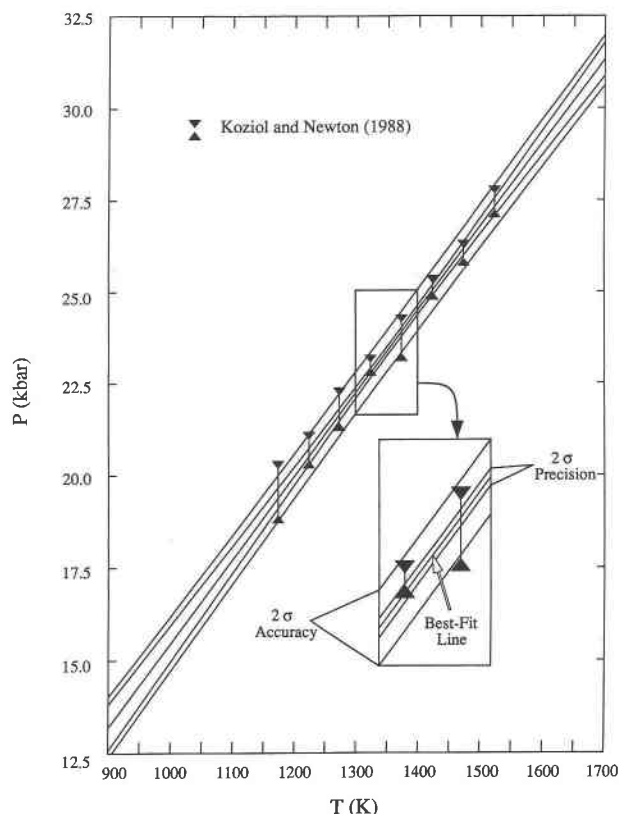


Fig. 2. P - T diagram showing the experimental reversals reported by Koziol and Newton (1988) for the GASP reaction, with the best-fit reaction position and 2σ precision (inner envelope) and accuracy (outer envelope) confidence limits. Inset shows two brackets with reaction position and confidence limits for clarity. Precision of the reaction position for each bracket was estimated using transition state theory.

As a final note on precision, we would like to point out that the kinetics of some reactions are fast enough so that very precise reaction positions are available from the bracket positions themselves. For example, the equilibria investigated by Bohlen et al. (1980, 1983a, 1983b) and Bohlen and Liotta (1986) are extremely tightly bracketed and have precisions of better than ± 100 to ± 200 bars over experimental and metamorphic temperature ranges. Thus, applying kinetic theory to some experimental data can substantially improve the precision of the reaction location (e.g., the Koziol and Newton data set), while other reactions are kinetically so rapid that transition state theory provides little additional constraint.

ACCURACY OF REACTIONS: SINGLE DATA SET

Method

It is an unfortunate fact that experimental devices are subject to systematic errors, and that these errors may be large compared to the precision to which a reaction is bracketed. As was the case with experimental reaction kinetics, the magnitude of systematic errors over P - T space for different experimental devices is not particularly well

known, but some estimates have been summarized by Hodges and McKenna (1987), and the interested reader is referred to that paper and to the references therein. Typical reported values for potential systematic pressure and temperature uncertainties (1σ) for piston cylinder devices using salt cell assemblies are ± 250 bars and ± 2.5 – 5 °C, respectively.

The effect of systematic errors on experimental data is quite different from that of imprecisions, as depicted in Figure 3, which shows the brackets for a hypothetical reaction as well as the 2σ confidence limits corresponding to the propagated precision for the reaction. The easiest way to deal with systematic errors in pressure is to assign them as an uncertainty in the reference pressure. That is, the pressure axis in Figure 3A has a systematic uncertainty ($\sigma_{P,\text{Syst}}$). Similarly, systematic uncertainties in temperature contribute to the error in pressure as the product of the slope of the reaction (m) and the uncertainty in temperature ($\sigma_{T,\text{Syst}}$; Fig. 3B). As noted by Hodges and McKenna (1987) the reaction slope is independent of the accuracy of the experimental device, and so is available from the regression of the experimental data for the precision of the bracketed reaction.

Once systematic errors in pressure and temperature have been assigned and the precision of the slope and intercept have been estimated from the regression of the experimental data, the accuracy of the reaction position may be calculated from the equation:

$$\sigma_p^2 = \sigma_{p,\text{Prec}}^2 + \sigma_{p,\text{Syst}}^2 + (m\sigma_{T,\text{Syst}})^2 \quad (6)$$

where $\sigma_{p,\text{Syst}}$ and $\sigma_{T,\text{Syst}}$ are the estimates of systematic pressure and temperature uncertainties and $\sigma_{p,\text{Prec}}$ is the precision of the reaction position as calculated using Equation 5.

Applications

The accuracy of the data of Koziol and Newton (1988) using the YK regression (Table 2) after assigning estimated systematic experimental pressure and temperature uncertainties of ± 250 bars and ± 5 °C is

$$\begin{aligned} \sigma_{p,2} = & (889)^2 + (0.65T)^2 - 2(0.9977)(889)(0.65T) \\ & + (250)^2 + [(5)(22.68)]^2 \end{aligned}$$

where the first three terms on the right side of the expression represent the precision of the reaction position as calculated using Equation 5. Because the slope of the reaction is rather shallow in P - T space (22.68 bars/K), the systematic uncertainty in temperature does not significantly affect the accuracy of the reaction position, as indicated by the small magnitude of the last term.

The resulting total 2σ confidence limits for the reaction are also plotted in Figure 2 (outer bounds). As can be seen, the relative contribution of the systematic errors to the uncertainty in the reaction position is greatest near the midpoint of the data and has diminishing importance as pressures are extrapolated outside the range of the data, and reaction precision becomes worse.

The accuracies represented by the bounds of Figure 2

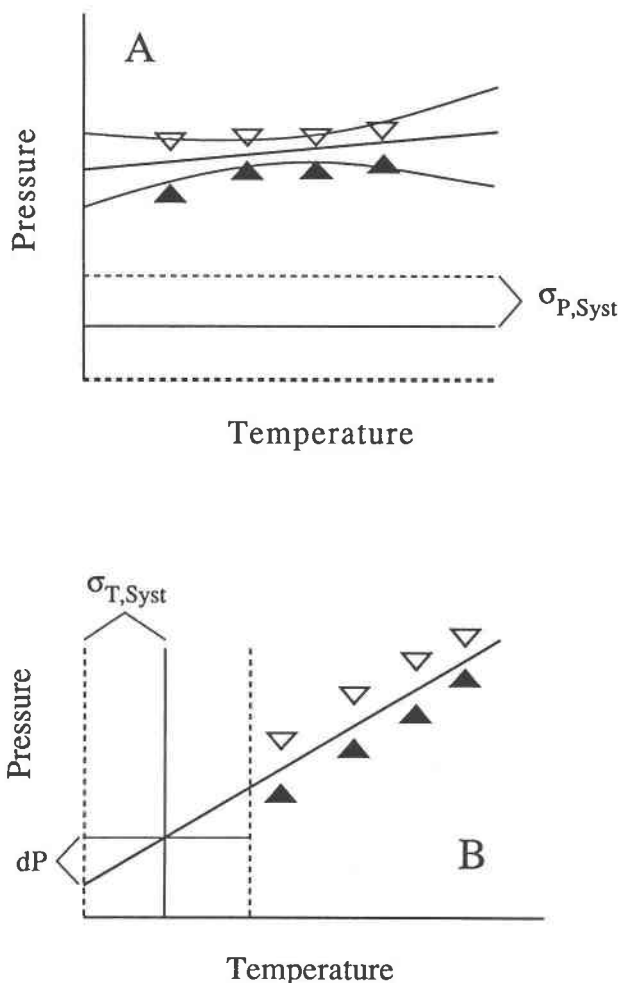


Fig. 3. Schematic P - T diagrams showing effects of systematic errors. (A) Systematic errors in pressure ($\sigma_{P,\text{Syst}}$) are equivalent to an uncertainty in the reference pressure. (B) Systematic errors in temperature ($\sigma_{T,\text{Syst}}$) introduce an uncertainty in pressure (dP) equal to the product of the slope (m) and $\sigma_{T,\text{Syst}}$.

are substantially smaller than those estimated by McKenna and Hodges (1988) for the data of Koziol and Newton (1988). For example, at 900 K, our propagated uncertainty (1σ) is about ± 410 bars while McKenna and Hodges (1988) estimate the uncertainty to be about ± 2400 bars (1σ) and virtually independent of temperature. In fact the discrepancy between our results and theirs is somewhat larger than these numbers might at first indicate because we chose to omit a theoretical bracket at 923 K when performing our regressions. If this bracket is included in the regression, then our method predicts an accuracy of about ± 375 bars (1σ) at 900 K. We are not able to explain how McKenna and Hodges derived such large uncertainties, but we note: (1) when we apply the Hodges and McKenna technique (Hodges and McKenna, 1987; McKenna, personal communication) assuming systematic uncertainties as described above, we merely reproduce our estimates of the accuracy, and (2) the accuracies described by McKenna and Hodges (1988)

for the GASP data are physically inconsistent with the discussion in Hodges and McKenna (1987) regarding experimental uncertainties. Possibly, McKenna and Hodges (1988) used different values or measures for systematic experimental errors.

For the data sets of Bohlen et al. (1980, 1983a, 1983b) and Bohlen and Liotta (1986), the accuracy of the end-member reactions at 900 K is estimated to be between ± 270 and ± 330 bars (1σ) for the same assumed systematic uncertainties in experimental pressures and temperatures.

We emphasize that any realistic error propagation analysis for single data sets requires a priori knowledge of potential systematic uncertainties in the pressure and temperature of the experiments, because any indication of their magnitudes is precluded using data collected on the same experimental device with the same standards and operating procedures. Only by comparing the results obtained on a single reaction using different types of experimental devices (e.g., Hays and Bell, 1973; Bohlen, 1984) or using different calibrations of the same type of apparatus (e.g., Johannes et al., 1971) can any statistical estimate of systematic experimental uncertainties and hence of reaction position accuracy be made.

MULTIPLE DATA SETS

While the description above is valid when propagating errors for single data sets, some reactions have been studied by several experimentalists. Therefore, it is desirable to develop a method whereby multiple data sets can be used together to obtain a best-fit reaction position and to assign uncertainties to that position. Furthermore, it is by comparing different data sets that estimates of the effects of interlaboratory and starting material variability can be made.

If different data sets are collected with different systematic errors, they will be displaced from each other when plotted together. The effect of the displacement of the experimental reference frames relative to each other leads to very real difficulties in estimating both the line that best fits experimental data collected on more than one apparatus and the errors corresponding to that best-fit line. For example, it is generally not justified to regress several different data sets together to estimate slope and intercept because of systematic pressure and temperature errors between different data sets and because the regressed best fit line will be most strongly controlled by the tightest reversals, even if those reversals occur in different data sets. Clearly, simultaneously regressing more than one set of experimental data can result in an erroneous slope and intercept.

Another important concern that must be addressed is the standards used in each study. If the same reference standards have been used, then discrepancies should reflect variability in reactant and product materials, or some interlaboratory variation not related to the uncertainty of the standard reactions; therefore, this variability constitutes precision, and additional terms, representing potential systematic errors would be required to determine ac-

curacy. If different standards have been used, then the variability in a bracketed reaction position between studies may be taken as a measure of interlaboratory variation (including starting materials and standard reactions), and so the variability may be simply taken as the accuracy.

Realistically, it may be difficult or impossible to demonstrate the independence of the standards used in different experiments. Some standards, such as the albite-jadeite-quartz reaction and salt melting curves, are used almost globally for calibrating piston cylinder devices, and it is unlikely that many of the experiments conducted within the last few years are truly independent with respect to their standards. Consequently, it is probably more generally correct to calculate uncertainties in reaction position assuming the same standards for different data sets; as is shown below, this approach tends to maximize calculated uncertainties for the data on GASP.

One approach to calculating slopes and intercepts for multiple data sets that preserves individual internal consistency is to regress each set of data separately for slope and intercept pressure. All individual slopes and pressures may then be averaged, weighting each value in accordance with the uncertainty of each parameter. Following Young (1962), a mean value and a minimum limit on its uncertainty when individual data are of varying quality are given by the equations:

$$\mu = \left[\sum \left(\frac{x_i}{\sigma_{x_i}^2} \right) \right] / \left[\sum \left(\frac{1}{\sigma_{x_i}^2} \right) \right] \quad (7)$$

and

$$\sigma_{\mu}^2 = 1 / \sum \left(\frac{1}{\sigma_{x_i}^2} \right) \quad (8a)$$

where μ is the mean value and x_i and σ_{x_i} are the individual values and their uncertainties. Another useful quantity is the (weighted) standard deviation of the observations about the mean:

$$\sigma_{\mu}^2 = \sum \{ [(x_i - \mu) / \sigma_{x_i}]^2 \} / \sum [1 / \sigma_{x_i}^2] \quad (8b)$$

Whereas Equation 8a represents a lower limit on the uncertainty in the mean value, Equation 8b is a direct measurement of the variability of the observations about the mean. Note that the mean value (μ) and its uncertainty (σ_{μ}) must be explicitly evaluated at each value of x_i . For experimental data such as for the GASP reaction, this implies that the best-fit reaction position may not be a straight line.

Applications: accuracy of the GASP reaction position

Experimental determinations of the end-member GASP reaction at different temperatures have been determined in seven studies (Hays, 1966; Hariya and Kennedy, 1968; Schmid et al., 1978; Goldsmith, 1980; Gasparik, 1984, 1985; and Koziol and Newton, 1988). We have chosen to exclude the two studies of Hariya and Kennedy (1968) and Schmid et al. (1978) from consideration because of the difficulties of interpreting their data in terms of the

restrictive requirements for the technique developed above for single data sets. The remaining five studies may not be completely independent of each other because all have been calibrated in part to the albite-jadeite-quartz equilibrium and because the corrections to the data of Gasparik (1985) are partly based on the positions of reactions determined by his earlier study (Gasparik, 1984). However, the accepted positions of standard reactions have changed over the last 25 years, and not all studies used the same set of standards. We will first assume that the standardizations were truly independent, and the possibility of lack of standard independence will be discussed later.

Consistent estimates of the slope and its possible uncertainties for the combination of all five data sets may be calculated from the individual slopes and their uncertainties listed in Table 2 (YK), using Equations 7, 8a, and 8b; the results are independent of both *T* and assumptions regarding systematic errors, and are listed at the bottom of Table 3. Use of Equations 8b indicates that the scatter of the individual slopes about the weighted mean is approximately 1.5 times the statistical minimum calculated using Equation 8a.

Assuming standards are independent. The pressures of the position of an individual reaction at a given temperature may be calculated from individual regressed slopes and intercepts (Table 2 and Eq. 4). However, the propagation of uncertainties in pressure depends on whether the standards used by the different studies are independent. If they are independent, then the precisions and accuracies of pressure for each set of experimental data at the temperature of interest may be calculated by using Equation 5 and then Equation 6, and a consistent mean and its possible uncertainties may be estimated by using Equations 7, 8a, and 8b; that is, the calculations follow the sequence: Equations 5, 6, 7, and then 8. This procedure accounts for systematic uncertainties prior to the calculation of μ and σ_μ .

Assuming that the standardizations of the five GASP data sets are independent, the weighted mean pressure and its uncertainties for a variety of temperatures are the values listed in columns 2, 3, and 4 of Table 3. These were calculated by using the YK regression results of Table 2 and by assuming systematic pressure and temperature uncertainties of ± 250 bars and ± 5 °C for each of the four most recent data sets, and ± 500 bars and ± 10 °C for the data of Hays (1966). The uncertainty in pressure calculated using Equation 8b (column 4) is larger than that calculated using Equation 8a (column 3), except in the range 800–1300 K. This is because the two most precise data sets (Gasparik, 1984; Koziol and Newton, 1988) converge in this temperature range. The larger of the two uncertainties tabulated in the third and fourth columns of Table 3 ranges from about ± 200 bars at the experimental temperatures to about ± 350 bars at crustal temperatures.

Assuming standards are the same. Although it might be argued that if the standards are the same, then the data from the different studies can be regressed together,

TABLE 3. Average pressures, slope, and uncertainties for GASP reaction

Average pressure and its uncertainty at different temperatures							
T (K) (1)	Assuming different standards:				Assuming same standards:		
	P (bars)		Uncertainties (bars)		P (bars)		Uncertainties (bars)
	Eq. 7 (2)	Eq. 8a (3)	Eq. 8b (4)	Eq. 7 (5)	Eq. 8a (6)	Eq. 8b (7)	Eq. 6 (8)
0	-6756	764	957	-6780	739	937	976
750	9941	402	408	9910	348	365	456
800	11055	380	374	11025	322	328	427
850	12169	359	342	12140	296	293	403
900	13284	338	310	13256	270	258	384
950	14398	317	280	14373	244	224	367
1000	15512	298	252	15491	217	192	349
1100	17738	261	205	17731	165	135	319
1200	19954	229	178	19980	114	97	296
1300	22152	201	179	22240	70	97	290
1400	24320	176	209	24473	59	140	307
1500	26469	155	270	26515	69	262	379
1600	28646	157	367	28635	80	353	447
2000	37554	288	646	37505	327	658	713

Average slope and its uncertainty (independent of T)		
Slope (bars/K)	Uncertainties (bars/K)	
Eq. 7	Eq. 8a	Eq. 8b
22.22	0.52	0.78

Note: Average slope and its uncertainties are independent of interpretation of experimental calibration standards and temperature. Accuracies (column 8) were calculated using Equation 6 from the maximum of the Equation 8a and 8b pressure uncertainties (columns 6 and 7) and from assumed ± 250 bars and ± 5 °C systematic pressure and temperature uncertainties.

we believe that this approach is only justifiable in the extremely unlikely event that exactly the same starting materials and laboratory procedures were used. Because of the variability of fluxes used in the GASP studies, the fact that most of the studies were conducted in different laboratories, and the long period of time over which the different studies were conducted, we do not believe that a simultaneous regression of all GASP data is warranted, even if the standard reactions were identical. Consequently, we favor application of Equations 7, 8a, and 8b.

Assuming the same reference reactions for the different studies, uncertainties in the position of an individual reaction may be calculated from Equation 5. With these σ_x , new values of μ and its uncertainties may be calculated by using Equations 7, 8a, and 8b. These calculations are reported in Table 3 for the GASP data in the fifth, sixth, and seventh columns. The accuracy of the reaction position may then be estimated from Equation 6 by assuming values for the systematic uncertainties in pressure and temperature and by noting that $\sigma_{P,prec}$ is the uncertainty in μ at the temperature of interest; that is, the equations are applied in the sequence 5, 7, 8, and then 6, and so μ and its precision are calculated before accounting for systematic uncertainties in pressure and temperature. For the GASP data, we assumed systematic pressure and tem-

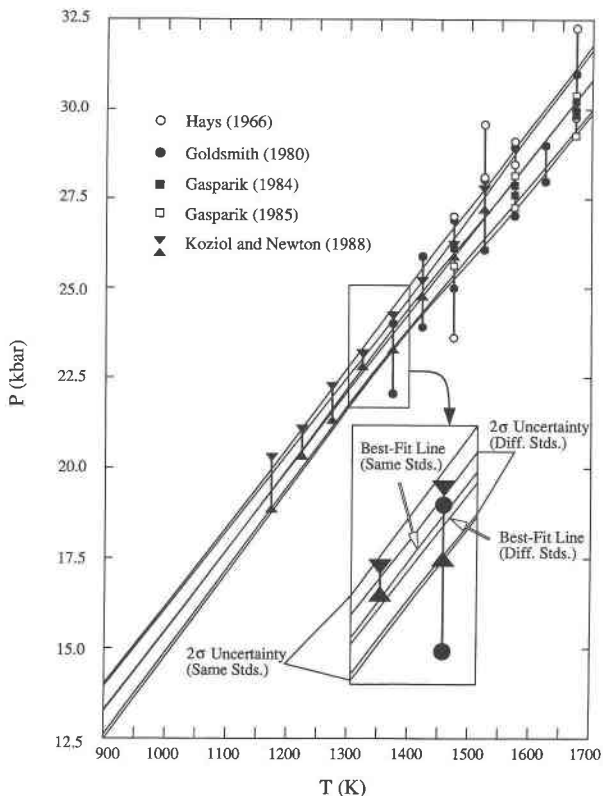


Fig. 4. P - T diagram showing the experimental reversals reported for the GASP reaction in five studies (Hays, 1966; Goldsmith, 1980; Gasparik, 1984, 1985; and Koziol and Newton, 1988), with the best-fit reaction positions and 2σ confidence limits. Inset shows brackets, reaction positions (central two curves; lower assumes that different experimental calibration standards were used for each experimental data set, upper assumes same standards), confidence limits assuming different calibration standards for each experimental data set (intermediate two curves), and confidence limits assuming the same standards for each data set (outer two curves).

perature uncertainties of ± 250 bars and ± 5 °C, respectively, and applied Equation 6 using the maximum of precisions for multiple data sets reported in Table 3 (columns 6 and 7); the results are tabulated in column 8 of Table 3. For this case, the accuracy of the GASP reaction position at 900 K is estimated to be approximately ± 380 bars (1σ).

Figure 4 shows the best-fit reaction position (fourth line from top), the individual brackets, and the 2σ confidence limits (second and fifth lines from top) for the assumption that the standards were different, and the best-fit reaction position (third line from top) and 2σ confidence limits (outer envelope) corresponding to the assumption that the standards were the same for all data sets. For the latter assumption the data of Koziol and Newton (1988) are much more heavily weighted than the other sets of data, and so the best-fit reaction position and confidence limits are bowed to higher pressure near the weighted mean temperature of these data. Because very similar standards

were used to calibrate the precise data of both Gasparik (1984) and Koziol and Newton (1988) (e.g., the albite-jadeite-quartz reaction location of Holland, 1980), we favor the assumption that the standards were the same for each data set and believe that the outer bounds of Figure 4 best represent the accuracy of the GASP reaction position based on the five data sets. Furthermore, we believe it is likely that the disparity of pressures at the experimental conditions (about ± 200 bars) represents differences in starting materials or unspecified interlaboratory variability. That is, a measure of interlaboratory, procedural, and starting material variability using the different GASP determinations may be approximately ± 200 bars (1σ).

As with the results for the single data set of Koziol and Newton (1988), our estimate of the accuracy is much smaller than that of McKenna and Hodges (1988). Using a data set similar to that shown in Figure 4, McKenna and Hodges (1988) calculated a 1σ accuracy at 900 K of about ± 1150 bars. As demonstrated in our study, the GASP data sets are simply much more consistent with each other than the uncertainties calculated by McKenna and Hodges (1988) would imply.

DISCUSSION

The small errors indicated for the precision and accuracy of the position of the GASP reaction should not be assumed to apply to all experimentally determined equilibria. The small propagated errors are a result of the consistency of several high-quality experimental determinations of the reaction. Equilibria that have only been studied in a single set of experiments usually cannot fulfill these conditions and may be correspondingly less accurate.

On the other hand, it should be recognized that a single set of extremely precise data (e.g., GRAIL) may be just as accurate as several sets of only moderately precise data (e.g., the collective GASP data), especially if the moderately precise data sets are calibrated to the same standards. For example, our analysis of the extremely tightly reversed single data sets presented by Bohlen et al. (1980, 1983a, 1983b) and Bohlen and Liotta (1986) all indicate precisions of better than ± 100 to ± 200 bars (1σ) at metamorphic temperatures. If potential systematic experimental pressure and temperature errors are assumed to be ± 250 bars and ± 5 °C, and if interlaboratory and starting material variabilities are assumed to contribute ± 200 bars of uncertainty (from the calculations above for the GASP reaction), then the accuracy of each reaction position is about ± 340 to ± 380 bars (1σ) at metamorphic temperatures. That is, the reactions studied by Bohlen et al. (1980, 1983a, 1983b) and Bohlen and Liotta (1986) may be just as accurately located as the GASP reaction, even though the GASP reaction has been more intensely studied.

Although it may be argued that the systematic experimental uncertainties we have assumed are not correct, it is relatively straightforward to adjust our estimates for

different values of systematic errors. We leave it to those who are skeptical (or at least uncertain) to perform any recalculations according to his or her belief as to the magnitude of such errors. Furthermore, we note that our analysis only accounts for those errors about which we know (e.g., bracket widths) or can make assumptions (e.g., systematic errors). There may, of course, be other unknown sources of error that bias reaction positions, and we caution that our estimates of uncertainty must be minima.

Finally, we emphasize that the preceding analysis does not represent the total uncertainty in the pressure of equilibration of a natural sample. Additional sources of error such as precision and accuracy of electron microprobe analyses, uncertainty in activity coefficients, and interpretation of compositional heterogeneity in natural samples must be considered, as is discussed in Part 2 (Kohn and Spear, 1991).

ACKNOWLEDGMENTS

This study is part of the M.S. work of M.J.K. We thank S. Bohlen, T. Gasparik, K. Hodges, A. Koziol, L. McKenna, E.B. Watson, and E. Young for reviewing early version of this paper, S. Roecker and A. Koziol for discussing error propagation and experimental techniques, and T. Hoisch, R. Powell, and J.C. Roddick for formal reviews; we particularly appreciated R. Powell's correcting some "wrong-headed reasoning" on our part. We thank L. McKenna for supplying us with a printout of his York regression program and for spurring our interest in the York-type regression approach. We are very grateful to A. Koziol for supplying us with unpublished X-ray data of the GASP experiments; however, any misinterpretation or misuse of the data is entirely our responsibility. This work was funded by NSF grants EAR-8708609 and EAR-8803785 to F.S.S. and a National Science Foundation graduate fellowship to M.J.K.

REFERENCES CITED

- Anderson, G.M. (1976) Error propagation by the Monte Carlo method in geochemical calculations. *Geochimica et Cosmochimica Acta*, 42, 1408-1414.
- Berman, R.G., Engi, M., Greenwood, H.J., and Brown, T.H. (1986) Derivation of internally-consistent thermodynamic data by the technique of mathematical programming, a review with application to the system MgO-SiO₂-H₂O. *Journal of Petrology*, 27, 1331-1364.
- Bohlen, S.R. (1984) Equilibria for precise pressure calibration and a frictionless furnace assembly for the piston-cylinder apparatus. *Neues Jahrbuch für Mineralogie Monatshefte*, 9, 404-412.
- Bohlen, S.R., and Liotta, J.J. (1986) A barometer for garnet amphibolites and garnet granulites. *Journal of Petrology*, 27, 1025-1034.
- Bohlen, S.R., Essene, E.J., and Boettcher, A.L. (1980) Reinvestigation and application of olivine-quartz-orthopyroxene barometry. *Earth and Planetary Science Letters*, 47, 1-10.
- Bohlen, S.R., Wall, V.J., and Boettcher, A.L. (1983a) Experimental investigation and application of garnet granulite equilibria. *Contributions to Mineralogy and Petrology*, 83, 52-61.
- (1983b) Experimental investigations and geological applications of equilibria in the system FeO-TiO₂-Al₂O₃-SiO₂-H₂O. *American Mineralogist*, 68, 1049-1058.
- Demarest, H.H., Jr., and Haselton, H.T., Jr. (1981) Error analysis for bracketed phase equilibrium data. *Geochimica et Cosmochimica Acta*, 45, 217-224.
- Draper, N., and Smith, H. (1981) *Applied regression analysis* (2nd edition). Wiley, New York.
- Fisher, G.W., and Lasaga, A.C. (1981) Irreversible thermodynamics in petrology. In A.C. Lasaga and R.J. Kirkpatrick, Eds., *Kinetics of geochemical processes*. Mineralogical Society of America Reviews in Mineralogy, 8, 171-185.
- 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.
- (1985) Experimental study of subsolidus phase relations and mixing properties of pyroxene and plagioclase in the system Na₂O-CaO-Al₂O₃-SiO₂. *Contributions to Mineralogy and Petrology*, 89, 346-357.
- 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 pressure. *American Journal of Science*, 266, 193-203.
- Hays, J.F. (1966) Lime-alumina-silica. *Carnegie Institution of Washington Year Book*, 65, 234-239.
- Hays, J.F., and Bell, P.M. (1973) Albite-jadeite-quartz equilibrium: A hydrostatic determination. *Eos*, 54, 482.
- Heinrich, W., Metz, P., and Gottschalk, M. (1989) Experimental investigation of the kinetics of the reaction 1 tremolite + 11 dolomite = 8 forsterite + 13 calcite + 9CO₂ + 1 H₂O. *Contributions to Mineralogy and Petrology*, 102, 163-173.
- Hodges, K.V., and Crowley, P.D. (1985) Error estimation and empirical geothermobarometry for pelitic systems. *American Mineralogist*, 70, 702-709.
- Hodges, K.V., and McKenna, L.W. (1987) Realistic propagation of uncertainties in geologic thermobarometry. *American Mineralogist*, 72, 671-680.
- Holland, T.J.B. (1980) The reaction albite = jadeite + quartz determined experimentally in the range 600-1200 °C. *American Mineralogist*, 65, 129-134.
- 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.
- Kohn, M.J., and Spear, F.S. (1991) Error propagation for barometers: 2. Application to rocks. *American Mineralogist*, 76, 138-147.
- Koziol, A.M., and Newton, R.C. (1988) Redetermination of the anorthite breakdown reaction and improvement of the plagioclase-garnet-Al₂SiO₄-quartz geobarometer. *American Mineralogist*, 73, 216-223.
- Lasaga, A.C. (1981a) Rate laws of chemical reactions. In A.C. Lasaga and R.J. Kirkpatrick, Eds., *Kinetics of geochemical processes*. Mineralogical Society of America Reviews in Mineralogy, 8, 1-68.
- (1981b) Transition state theory. In A.C. Lasaga and R.J. Kirkpatrick, Eds., *Kinetics of geochemical processes*. Mineralogical Society of America Reviews in Mineralogy, 8, 135-170.
- (1986) Metamorphic reaction rate laws and development of isograds. *Mineralogical Magazine*, 50, 359-373.
- McKenna, L.W., and Hodges, K.V. (1988) Accuracy versus precision in locating reaction boundaries: Implications for the garnet-plagioclase-aluminum silicate-quartz geobarometer. *American Mineralogist*, 73, 1205-1208.
- Powell, R., and Holland, T.J.B. (1985) An internally consistent thermodynamic dataset with uncertainties and correlations: 1. Methods and a worked example. *Journal of Metamorphic Geology*, 3, 327-342.
- Roddick, J.C. (1987) Generalized numerical error analysis with applications to geochronology and thermodynamics. *Geochimica et Cosmochimica Acta*, 51, 2129-2135.
- Schmid, R., Cressey, G., and Wood, B.J. (1978) Experimental determination of univariant equilibria using divariant solid-solution assemblages. *American Mineralogist*, 63, 511-513.
- Schramke, J.A., Kerrick, D.M., and Lasaga, A.C. (1987) The reaction muscovite + quartz = andalusite + K-feldspar + water. *American Journal of Science*, 287, 517-559.
- York, D. (1969) Least squares fitting of a straight line with correlated errors. *Earth and Planetary Science Letters*, 5, 320-324.
- Young, H.D. (1962) *Statistical treatment of experimental data*, 172 p. McGraw-Hill, New York.

MANUSCRIPT RECEIVED JANUARY 29, 1990

MANUSCRIPT ACCEPTED NOVEMBER 19, 1990