

## **Exploratory data analysis in thermobarometry: An example from the Kiseynew sedimentary gneiss belt, Manitoba, Canada**

**TERENCE M. GORDON**

Department of Geology and Geophysics, University of Calgary, Calgary, Alberta T2N 1N4, Canada

**LEONID YA. ARANOVICH, VALENTIN V. FED'KIN**

Institute of Experimental Mineralogy, Russian Academy of Sciences, Chernogolovka, Moscow District 142432, Russia

### **ABSTRACT**

Generalized thermobarometry uses measured partial molar free-energy functions computed from standard state properties, activity models, and measured compositions, along with the linear dependence of the solutions to these functions at equilibrium to estimate pressure, temperature, and partial molar free energies of individual mineral end-members. Such problems are often inconsistent, i.e., there are no values of pressure and temperature that simultaneously satisfy all the equations. These inconsistencies result from errors in the measured partial molar free-energy functions and may be classified into random errors (those that can be modeled as arising from a statistical population) and systematic errors (those that cannot). Because the systematic error may dominate, exploratory data analysis is an essential step in the interpretation of thermobarometry results. Several complementary techniques may be used to provide insight into the sensitivity of estimated pressures and temperatures to systematic errors in the measured partial molar free-energy functions for individual end-members.

The number of equations is small, and there are usually only one to three fewer unknowns than equations; hence, statistical methods designed to investigate inconsistencies in large data sets cannot be readily applied. However, several procedures can be used to guide interpretation of the results of thermobarometric calculations. (1) Examination of residuals provides a direct indication of relative misfit of the measured partial molar free-energy functions to the theoretical values of chemical potential at equilibrium. The intrinsic correlation of residuals in small data sets limits the utility of this approach. (2) Contours of the sum of squares of residuals illustrate the covariance and overall reliability of the least-squares solution to the problem. Assigning actual confidence levels to contours requires that the errors have statistical properties that may not exist in practice. (3) Case deletion studies require that the equations be solved repeatedly, eliminating the equation for each individual end-member in turn. Diagrams illustrating these results serve to confirm or refute the existence of end-members that are highly influential in determining best-fit pressures and temperatures. (4) Perturbation analysis also requires the repeated solution of the problem, but, instead of deleting individual end-members, their measured free-energy estimates are perturbed by fixed amounts, and a new pressure and temperature are determined. Displays of this kind provide quantitative estimates of the change in estimated pressures and temperatures resulting from arbitrary perturbations in thermodynamic constants and measured compositions.

### **INTRODUCTION**

The application of thermobarometry to high-grade metamorphic terranes is often complicated by retrogression or polymetamorphism, which makes it difficult to be sure that measured mineral compositions correspond to equilibrium conditions. The problem may be compounded by systematic errors in the activity models and thermodynamic constants used to compute pressures and temperatures. In such cases, exploratory data analysis is an important first step in the interpretation of the results of the calculations. This paper illustrates extensions to Gordon's (1992) formulation of the inverse chemical

equilibrium problem to permit examination of the sensitivity of pressure and temperature determinations to systematic errors in the measured partial molar free-energy functions for individual end-members. The example uses mineral analyses from assemblages in high-grade gneisses of the Kiseynew sedimentary gneiss belt in the Canadian shield.

### **GEOLOGICAL SETTING**

The Kiseynew gneiss belt is a 150 × 300 km metamorphic complex in northern Manitoba and Saskatchewan that forms part of the Early Proterozoic Trans-Hud-

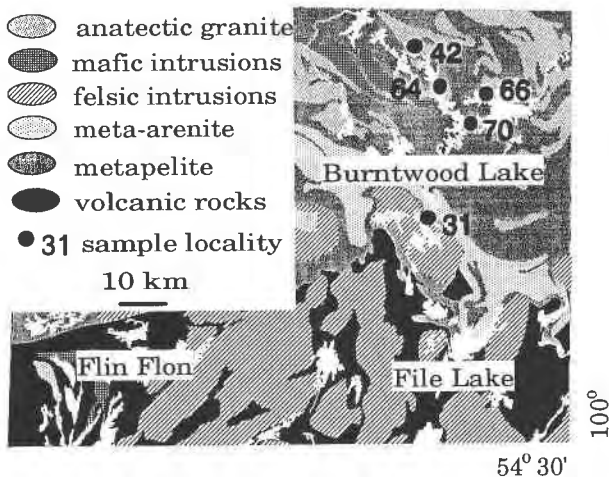


Fig. 1. Geological map and sample localities.

son orogen (Lewry and Stauffer, 1990). The rocks are dominantly garnet + cordierite and potassium feldspar + sillimanite migmatites derived from greywacke and shale turbidites and arenites, with subordinate anatectic granitoids. Samples used in this study were collected as part of a 70-km north-trending transect that extends from the File Lake area in the Flin Flon belt to the central part of the gneiss belt (Fig. 1). The study area has been mapped in detail (Baldwin et al., 1979); a regional compilation map is available (Manitoba Energy and Mines, 1988); and the regional metamorphic history is reviewed in Gordon (1989).

The five samples discussed in this paper were selected from a larger suite for which the following observations can be made: Most of the samples examined contain as major phases variable amounts of quartz, plagioclase, garnet, biotite, cordierite, potassium feldspar, and sillimanite, which is typical for upper amphibolite facies metapelites. Primary muscovite has been found only in one sample (not included in this study) from the southern part of the area. Although morphologically distinct generations of garnet, cordierite, and biotite occur in many samples, this study is concerned with the earliest gener-

TABLE 1. Plagioclase electron microprobe analyses and formulae\*

Sample no.	31	70	66	64	42
SiO <sub>2</sub>	62.10	61.15	64.09	62.01	62.01
Al <sub>2</sub> O <sub>3</sub>	23.90	25.03	22.64	24.19	23.92
CaO	3.89	6.31	5.25	5.83	3.99
Na <sub>2</sub> O	9.76	8.04	8.02	7.46	9.76
K <sub>2</sub> O	0.13	0.17	0.23	0.31	0.13
Total	99.78	100.70	100.23	99.80	99.81
Si	2.757	2.700	2.821	2.749	2.754
Al	1.251	1.302	1.175	1.264	1.252
Ca	0.185	0.298	0.248	0.277	0.190
Na	0.840	0.688	0.684	0.641	0.840
K	0.007	0.010	0.013	0.018	0.007

\* Computed on basis of eight O atoms.

TABLE 2. Garnet electron microprobe analyses and formulae\*

Sample no.	31	70	66	64	42
SiO <sub>2</sub>	37.31	38.42	38.80	38.63	37.00
TiO <sub>2</sub>	—	—	0.11	0.02	—
Al <sub>2</sub> O <sub>3</sub>	21.81	21.85	20.76	21.86	22.24
FeO	35.20	31.33	30.49	31.76	34.48
MgO	4.87	6.38	7.91	6.96	5.35
MnO	—	0.60	0.44	0.23	0.13
CaO	0.77	1.12	0.90	0.89	0.79
Total	99.96	99.70	99.41	100.35	99.99
Si	2.971	3.015	3.041	3.010	2.939
Ti	—	—	0.006	0.001	—
Al	2.047	2.021	1.918	2.007	2.082
Fe	2.344	2.057	1.999	2.069	2.290
Mg	0.578	0.746	0.924	0.808	0.633
Mn	—	0.040	0.029	0.015	0.009
Ca	0.060	0.094	0.076	0.074	0.067

\* Computed on basis of 12 O atoms.

ation of these minerals, Grt1, Crd1, and Bt1. All these samples contain quartz and sillimanite in addition to grains of Grt1, Crd1, and plagioclase, from which electron microprobe analyses were obtained. Three of the samples also contain Bt1. Compositional data for these phases are presented in Tables 1–4.

The assumptions required to apply equilibrium thermodynamics to these samples are that (1) Qtz + Pl + Sil + Bt1 + Grt1 + Crd1 formed an equilibrium assemblage; (2) the present compositions of the cores of Bt1, Grt1, and Crd1 grains were at equilibrium; and (3) the present plagioclase composition is the same as that at equilibrium with Bt1 + Crd1 + Grt1. Although similar assumptions are necessary in all thermobarometric studies, the complex textures in these rocks make the assumptions less reliable than for lower-grade terranes. Because of that, it is important to investigate the effects of the various assumptions on the resulting *P-T* estimates.

#### ERRORS AND EXPLORATORY DATA ANALYSIS

Two philosophies of error analysis are in common use (Colclough, 1987). In the first approach, all errors are considered to be random—that is, they arise from a single stable statistical population for which a mean and variance can be estimated from measurements. If that philosophy is adopted, then statistical theory provides weighting criteria for determining unbiased and maximum-likelihood solutions to least-squares problems.

In the second approach, errors are considered to arise from both random sources and systematic sources. Systematic sources of uncertainty cannot be treated as samples from a statistical population, and an error covariance matrix cannot be estimated for them. Sources of systematic error in thermobarometric studies have been identified by Kohn and Spear (1991a, 1991b) and Holdaway and Mukhopadhyay (1993), who elaborated on the kinds of error for which statistical analysis is not appropriate. These include systematic errors in thermochemical constants and activity models, as well as operator error in

**TABLE 3.** Biotite electron microprobe analyses and formulae\*

Sample no.	70	66	42
SiO <sub>2</sub>	35.68	36.82	35.98
TiO <sub>2</sub>	5.34	3.68	3.45
Al <sub>2</sub> O <sub>3</sub>	16.00	16.32	17.95
FeO	16.47	15.76	19.11
MgO	10.43	12.05	9.78
Na <sub>2</sub> O	—	0.01	0.30
K <sub>2</sub> O	9.54	9.75	8.24
Total	93.46	94.39	94.81
Si	2.742	2.782	2.726
Ti	0.309	0.209	0.197
Al	1.449	1.453	1.603
Fe	1.059	0.996	1.211
Mg	1.195	1.357	1.104
Na	—	0.001	0.045
K	0.935	0.940	0.796

\* Computed on basis of 11 O atoms.

the choice of phases and portions of mineral grains modeled as an equilibrium system.

In recognition of the second philosophy, this study is intended to illustrate the advantages and pitfalls of several exploratory data-analysis techniques applied to thermobarometry problems to provide information on the possible effects of systematic error. These are direct methods that allow the graphical analysis of the sensitivity of the solutions to perturbations in the input measurements. This approach allows geological experience, judgment, and knowledge of details of the particular thermobarometric study to be applied to the results and their interpretation. In this sense, the approach follows Hoaglin et al. (1983, p. 1): "Good statistical practitioners have always looked in detail at the data before producing summary statistics and tests of hypotheses. . . . The emphasis is on flexible probing of the data, often before comparing them to any probabilistic model." Gordon's (1992) methodology for generalized thermobarometry provides a particularly convenient formulation for such exploratory analysis.

### LEAST SQUARES

In Gordon's (1992) formulation, measured partial molar free-energy functions of the end-members of analyzed phases are expressed as a set of equations in the unknowns pressure, temperature, and the theoretical equilibrium chemical potentials of a set of compositionally independent end-members. A particular problem is underdetermined, exact, or overdetermined, depending on the total number of end-members for which there are measurements and the number of these that are compositionally independent. Least-squares solutions to overdetermined problems provide both optimal  $P$ - $T$  determinations and sets of residuals that provide a measure of the misfit of the measured partial molar free-energy functions to the equilibrium model. Appendix 1 provides details to the equations used in this study.

The least-squares procedure always provides a best solution in a geometric sense, but it produces a minimum-variance unbiased estimate of pressure and temperature

**TABLE 4.** Cordierite electron microprobe analyses and formulae\*

Sample no.	31	70	66	64	42
SiO <sub>2</sub>	50.13	50.64	50.03	49.56	48.85
TiO <sub>2</sub>	—	0.14	0.04	—	—
Al <sub>2</sub> O <sub>3</sub>	34.39	33.86	33.38	33.10	34.25
FeO	6.80	6.55	6.61	6.97	8.68
MgO	8.25	9.64	9.70	9.16	8.16
MnO	0.10	0.01	—	—	—
CaO	0.20	0.05	—	0.13	—
Na <sub>2</sub> O	—	0.06	0.81	—	—
K <sub>2</sub> O	0.13	—	—	0.07	0.06
Total	100.00	100.95	100.57	98.99	100.00
Si	5.029	5.0270	5.005	5.031	4.948
Ti	—	0.0100	0.003	—	—
Al	4.066	3.9620	3.936	3.961	4.089
Fe	0.571	0.5440	0.553	0.592	0.735
Mg	1.234	1.4270	1.447	1.386	1.232
Mn	0.008	0.0010	—	—	—
Ca	0.021	0.0050	—	0.014	—
Na	—	0.0120	0.157	—	—
K	0.017	—	—	0.009	0.008

\* Computed on basis of 18 O atoms.

only if the problem is weighted by a known covariance matrix of errors (Hamilton, 1964). Because this study is exploratory in nature, and because systematic errors may be more significant than random errors, no attempt was made to estimate error covariance matrices, and the least-squares calculations were carried out without weighting. The results are presented in Table 5. They include the best-fit temperature and pressure, values of the measured partial-molar free-energy functions at the best-fit  $P$  and  $T$ , the theoretical chemical potentials, and the residuals (differences between measured and theoretical chemical potentials).

The sum of the squared residual (error) terms provides a single measure of the degree of misfit of the data to the equilibrium model. The smaller this term, the closer the solution is to being exact. In the examples studied, sample 64 has an extremely small error sum of squares, indicating that the measured partial molar free-energy func-

**TABLE 5.** Results of the least-squares solution of the thermobarometry problem for assemblages and compositions in Tables 1–4

Sample	31	70	66	64	42
Error sum of squares (kJ <sup>2</sup> )	3.913	3.360	2.190	0.0048	5.591
$T$ (°C)	580	773	794	834	718
$P$ (bars)	4308	5436	5607	5594	4764
<b>Residuals <math>g_i(P,T) - \Sigma\mu_i(\text{kJ})</math></b>					
qtz	0.014	-0.260	-0.093	0.001	-0.301
sil	-0.267	0.104	0.270	0.009	0.457
an	0.694	-0.781	-0.860	-0.020	-1.744
alm	0.761	0.134	-0.627	-0.027	-0.847
prp	-0.663	-0.343	0.474	0.024	0.494
grs	-0.231	0.260	0.287	0.007	0.581
Fe-crd	-1.141	-0.854	-0.418	0.040	0.321
Mg-crd	0.994	1.166	0.646	-0.036	0.209
ann	—	0.435	0.196	—	0.634
phl	—	-0.435	-0.196	—	-0.634

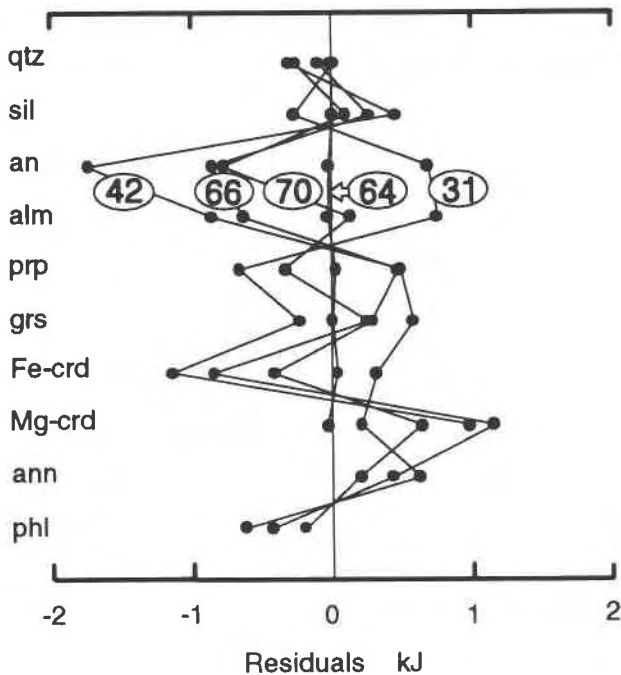


Fig. 2. Residuals in least-squares solutions. Sample nos. are circled.

tions for each end-member provide values very close to the theoretical equilibrium chemical potentials at the optimal  $P$  and  $T$ . This sample, however, gives an anomalously high temperature estimate. The lowest temperature estimate comes from sample 31, which has the largest sum of squares at 3.9 kJ<sup>2</sup>. The exploratory data procedures discussed below provide important additional insights for the interpretation of these results.

#### EXAMINATION OF RESIDUALS

As might be expected for rocks with complex mineralogical histories, the  $P$ - $T$  results in Table 5 show a considerable range. However, all the residuals in partial molar free energy are <1.8 kJ, and all but three are <1 kJ. They are comparable in magnitude with uncertainties in tabulated thermochemical constants and activities based on compositional measurement. For example, the published uncertainty in the standard state Gibbs free energy of formation of quartz is 1.1 kJ (Robie et al., 1979). Similarly, because  $d[RT \ln(a)] = RT da/a$ , a 1-kJ error in measured partial molar free energy is approximately equivalent to a 12% relative error in the activity of any end-member at 740 °C. For ideal mixing on a single site with multiplicity,  $n$ , a 1-kJ error in partial molar free energy thus corresponds to a relative error in mole fraction of  $12/n\%$  at 740 °C.

Exploratory data analysis places a heavy emphasis on examination of residuals to identify outliers, influential observations, and high leverage points, e.g., the texts by Belsley et al. (1980) and Cook and Weisberg (1982) and the review papers by Chatterjee and Hadi (1986) and

Rousseuw and van Zomeren (1990), along with the appended commentaries. However, the utility of most statistical diagnostic tests in thermobarometry problems is reduced by the relatively small number of equations to be solved. Free-energy estimates can be determined only for end-members for which there are thermochemical data and activity models; hence, current thermobarometry problems often have 12 or fewer equations and, more importantly, are seldom overdetermined by more than three equations.

Examination of residuals can be done by inspection (Table 5) but more commonly by graphic display (Fig. 2). The purpose of such plots is to reveal outliers and any dependence of the variance of the residuals on the independent variables. An important caveat is that residuals in all least-squares problems are correlated, and, although the effects become small as the number of observations approaches infinity, they cannot be eliminated. In small data sets, this intrinsic correlation can be overwhelming. This phenomenon makes it difficult to interpret residuals from small data sets simply by comparing their signs and magnitudes. For example, correlated residuals can be detected in Table 5 and Figure 2. In particular, the residuals for annite and phlogopite are of opposite sign and identical magnitude. This effect depends on all equations in the problem, not just the equations for which correlation can be identified. For example, in contrast to the assemblages studied in this paper, the example problem in Gordon (1992) gives residuals for annite and phlogopite with identical signs but different magnitudes.

For this reason, although Table 5 and Figure 2 can be used to identify particularly large residuals (Fe-crd in sample 31, Mg-crd in sample 70, an in sample 42), these extreme values do not necessarily indicate that there is a systematic error in the free-energy estimates for these particular end-members. The case deletion and sensitivity analysis procedures described below provide more information for assessing the influence of errors for individual species.

#### CONTOURS OF THE SUM OF SQUARES

The least-squares algorithm finds an optimal pressure and temperature at the minimum of the sum of the squared error terms of the equations given in Appendix 1. If the problem is repeatedly solved with pressure and temperature fixed at other values, the resulting sum of squares surface can be contoured. Such diagrams are shown in Figure 3. The contours have been drawn at 5, 10, 15, and 20 kJ<sup>2</sup> above the minimum sum of squares at each optimal solution. These plots provide significant information for the interpretation of the least-squares results. First, the relative areas indicate the overall robustness of the solution. Large areas such as those for samples 31 and 64 indicate solutions that are more sensitive to perturbation than the solutions for samples 70, 66, and 42. Second, the shapes of the contours illustrate the relative sensitivity and correlation of the  $P$ - $T$  determinations. The diagrams for samples 31 and 64 show a higher

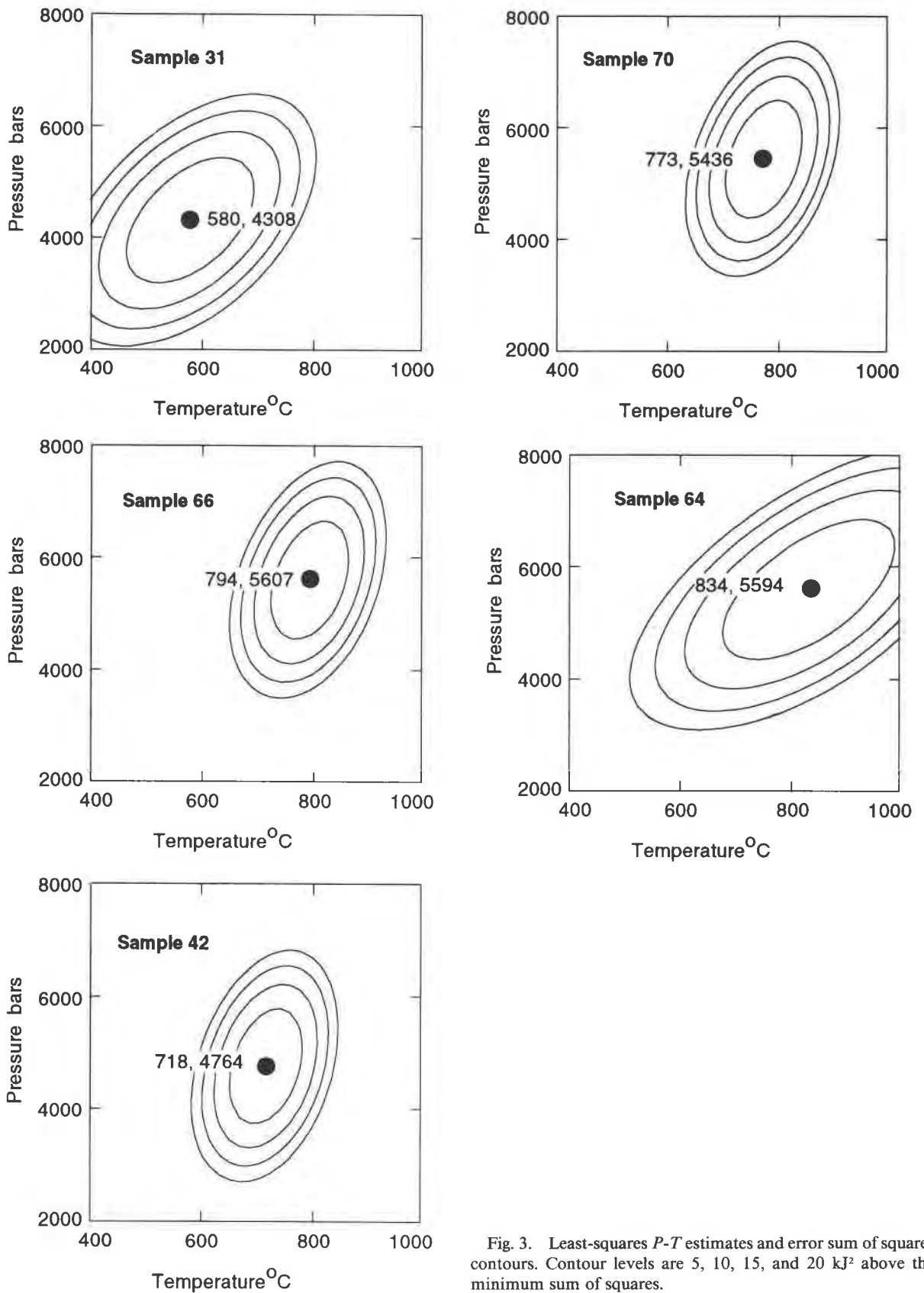


Fig. 3. Least-squares  $P$ - $T$  estimates and error sum of squares contours. Contour levels are 5, 10, 15, and 20  $\text{kJ}^2$  above the minimum sum of squares.

degree of correlation between pressure and temperature than the remaining samples and also show that temperature is relatively poorly constrained by these biotite-free samples. The large differences in the sum of squares for samples 31 and 64 and the extreme values of temperature that they produce can thus be understood as a result of this effect.

If all the errors could be assumed to be random, confidence regions for the fitted  $P$  and  $T$  could be represented by contours of constant sum of squares. Confidence levels can be assigned if the distribution function of the errors is known or if the number of equations greatly exceeds the number of parameters to be determined. Because these requirements cannot be met in this study, meaningful confidence levels cannot be assigned.

### CASE DELETION

Many procedures in regression diagnostics rely on case deletion methods, in which various sample statistics are calculated with one or more cases (equations) deleted from the original data set. Large changes in the problem's solution identify highly influential data points. Most thermobarometry problems are overdetermined by only one to three equations; hence, the problem may become exact when cases are deleted. This precludes the use of statistical measures of influence that require estimates of the variance of the errors. In problems that remain overdetermined, the already small difference between the numbers of unknowns and parameters to be estimated is further reduced. For these reasons, sophisticated methods of data analysis are not generally applicable to thermobarometric work. The approach taken here is a simple graphical examination of the sensitivity of the  $P$ - $T$  estimates to deletion of the equation for each end-member in turn. Such a procedure is useful if an investigator has some reason to suspect that inconsistency in the equilibrium model is due entirely to a systematic error in the partial molar free-energy equation for a single end-member.

For the assemblages in this study,  $P$ - $T$  estimates determined with the successive deletion of equations for individual end-members are illustrated in Figure 4. These diagrams have several interesting properties. First, the deletion of a particular equation does not always produce the same  $P$ - $T$  difference for all samples. For example, although the deletion of the equation for either biotite end-member in samples 70, 66, and 62 generally results in solutions with lower pressure and temperature, there is no similar correlation for cordierite or garnet end-members. Second, the estimated pressures are very sensitive to the inclusion of quartz and, to a lesser extent, sillimanite in the model equilibrium assemblage. This is not a problem in these rocks, but it does show the importance of these phases to the solutions. Most importantly, a comparison with the large residuals noted above (Fe-crd in 31, Mg-crd in 70, and an in 42) shows that the deletion of the equations for these suspicious end-members would cause little change in the estimated pressures and temperatures for those samples. This reinforces the

observation that interpreting the least-squares residuals for small data sets must be done with caution.

In these high-grade rocks, Fe-Mg exchange after peak metamorphism might have affected even the core compositions used in the calculations. The effect of discarding the equations for individual Fe or Mg end-members is readily apparent in Figure 4. For any of the biotite-bearing samples (70, 66, and 42), the elimination of either phl or ann results in significantly lower temperature estimates, whereas the elimination of Fe and Mg garnet end-members produces changes that are dependent on the specific assemblage. In spite of the reputation of cordierite for causing difficulties in thermobarometry, the  $P$ - $T$  estimates are insensitive to the equations for the cordierite end-members.

### SENSITIVITY TO PERTURBATIONS

Case deletion effectively ignores the data for individual measured partial molar free energies. A more precise measure of sensitivity can be obtained by perturbing each partial molar free-energy equation by a constant amount and noting the effect this has on the solution. Plots of the resulting set of perturbation vectors permit the visual identification of end-members for which small changes in measured thermodynamic properties or composition produce large shifts in estimated pressures and temperatures. These results can then be interpreted on the basis of the details of each particular problem. Because apparent free energies of solids are nearly linear over limited pressure and temperature ranges, such displays can be used to assess quantitatively the sensitivity of  $P$ - $T$  determinations to measurement errors.

For the end-member in each assemblage, new least-squares solutions were obtained by adding and subtracting a constant number of kilojoules to the measured partial molar free-energy functions for the individual end-members. For purposes of illustration, a  $\pm 2$ -kJ perturbation was chosen for all end-members, although these values could be chosen on an individual basis. The results are shown in Figure 5. Over a range of 200 K in temperature and 2 kbar in pressure, the nonlinearity of the free-energy equations produces no significant effect, so the new pressure and temperature estimates fall on straight lines and positive and negative perturbations of the same magnitude produce equal and opposite effects. Changing the value of the free-energy perturbation for any end-member produces an equivalent change in the length of the perturbation vector.

An immediately obvious feature of Figure 5 is that the  $P$ - $T$  estimates for assemblages that are overdetermined by only one equation (31 and 64) are far more sensitive to perturbations than those overdetermined by two equations (70, 66, and 42). Within each of these two groups of samples, the free-energy perturbations produce nearly identical effects. Note that for samples 70, 66, and 42, the temperature estimates are relatively insensitive to 2-kJ perturbations in free energies of any end-member, whereas 2-kJ perturbations in qtz and sil free energies produce

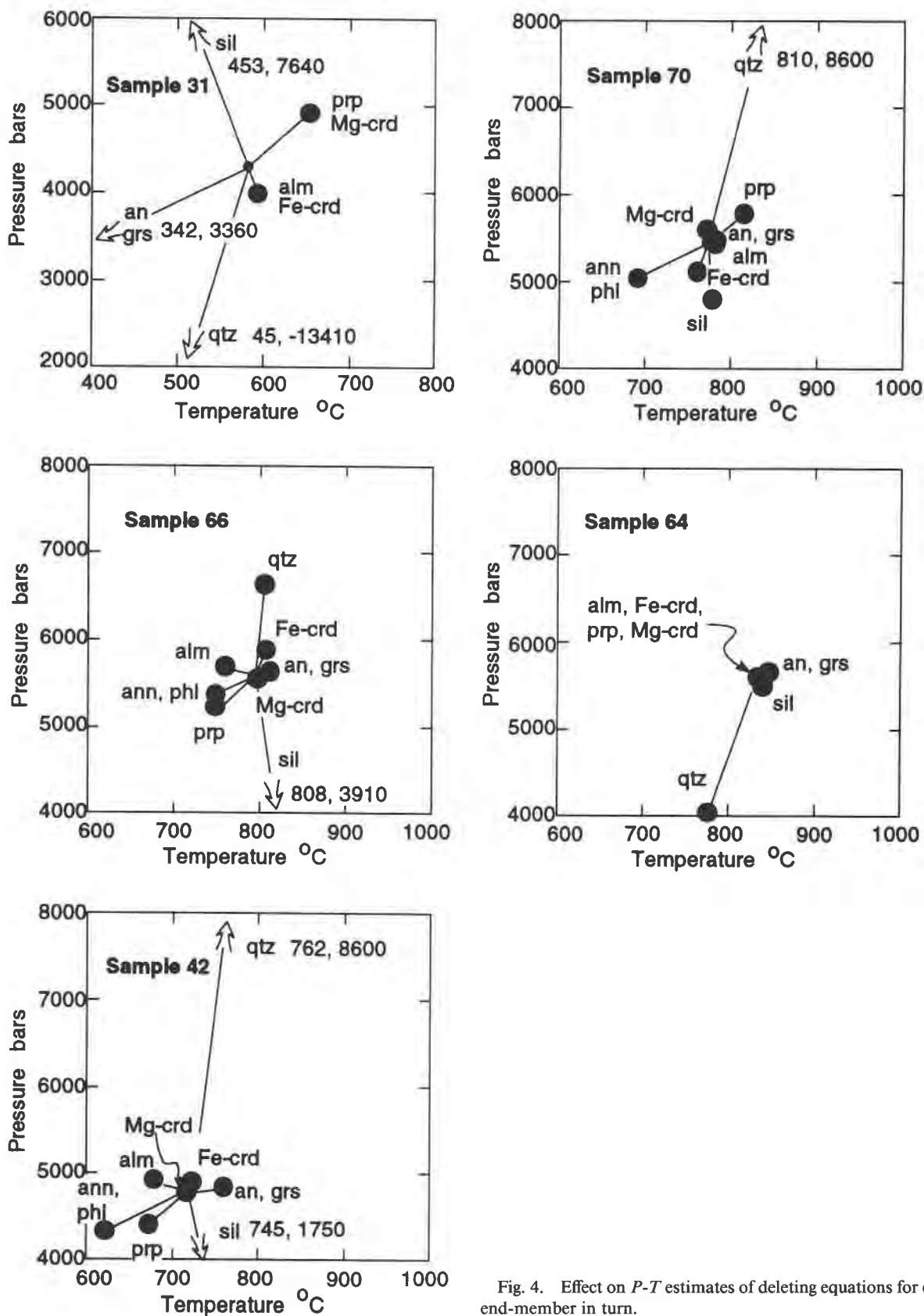


Fig. 4. Effect on *P-T* estimates of deleting equations for each end-member in turn.

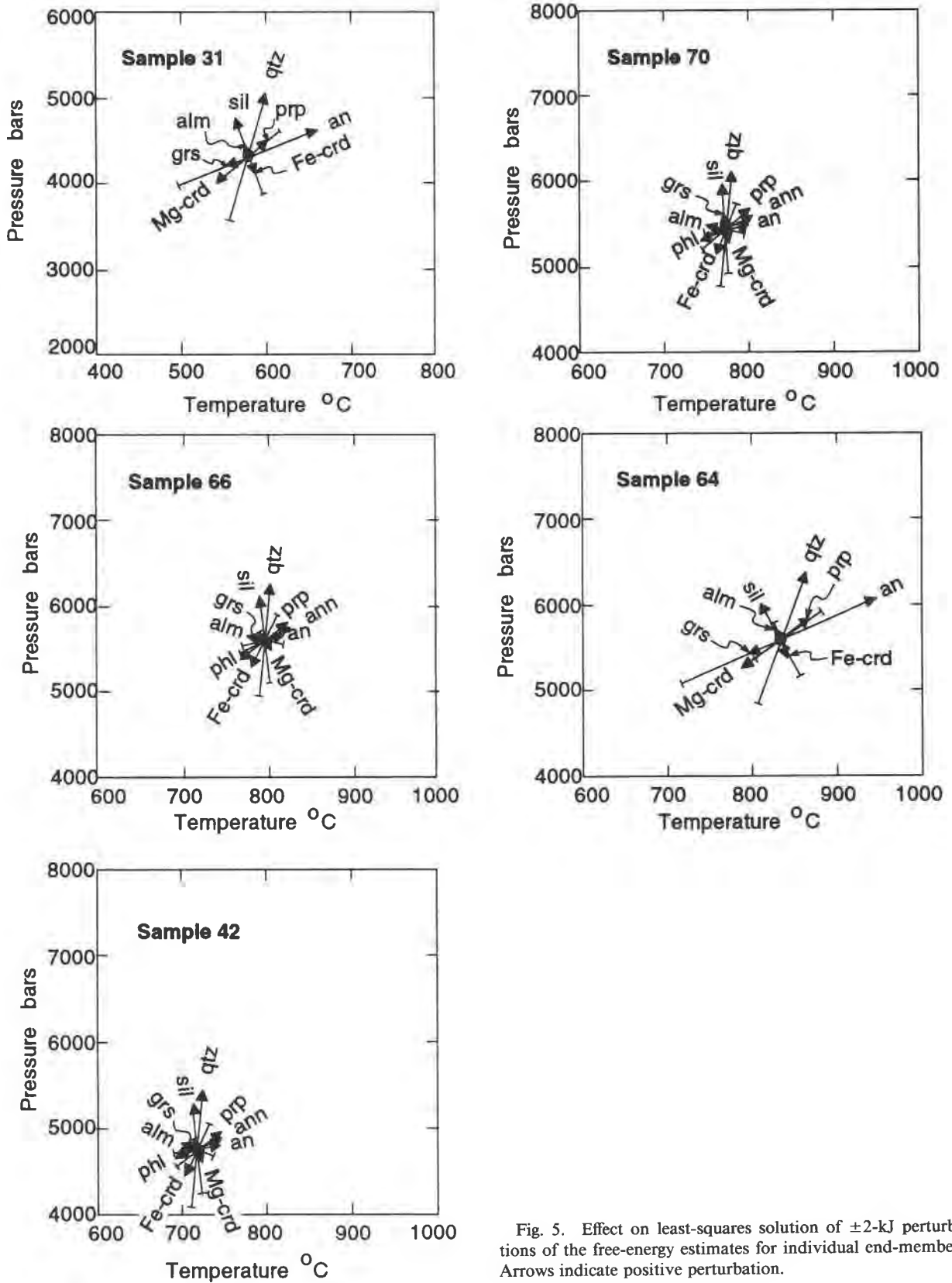


Fig. 5. Effect on least-squares solution of  $\pm 2$ -kJ perturbations of the free-energy estimates for individual end-members. Arrows indicate positive perturbation.



nearly isothermal 500-bar changes in estimated pressures.

For this study, the most telling features of Figure 5 are the diagrams for samples 31 and 64. These samples have two characteristics in common: (1) they have no biotite and hence are overdetermined by only one equation and (2) their unperturbed  $P$ - $T$  estimates are at opposite extremes in both  $P$ - $T$  and magnitudes of sums of squares. Although all the perturbation vectors are large, those for plagioclase are particularly pronounced. This indicates that for these two assemblages a relatively small change in the partial molar free-energy function for an can make a major change in the estimated pressure and temperature and the value of the sum of squares. The same thermochemical data were used in all calculations; hence, suspicion must fall on the compositions of plagioclase used in the equations. For these samples it was assumed that the present plagioclase composition was at equilibrium with the remaining phases. Figure 5 suggests that this assumption has a major influence on the estimated temperatures and may be the explanation for the anomalous temperatures estimated for these samples.

## DISCUSSION

In thermobarometric problems, the number of equations is small and very nearly equal to the number of unknowns. In addition, the sources of error are commonly believed to include systematic effects. These two characteristics preclude the use of most statistical techniques designed for large data sets. In such cases, the scientific judgment of the investigator is necessary to interpret the results. The exploratory procedures described here permit the visual examination of the effect of errors in individual free-energy estimates on  $P$ - $T$  determinations. This permits the recognition of particularly influential end-members as well as providing a quantitative estimate of the sensitivity of the solution to errors in thermodynamic constants and chemical analyses.

For the samples in this study, exploratory data analysis points to plagioclase composition as highly influential in determining the results of the thermobarometric calculations. Conversely, the influence of cordierite end-members on the results is shown to be small. These conclusions can be used to guide further study of the thermobarometry of these complex rocks.

## ACKNOWLEDGMENTS

This research benefited from discussions with Jim Nicholls, Ed Ghent, and Dave Pattison. Constructive reviews by Matthew Kohn and Howard Day led to substantial improvements to the manuscript. This is Lithoprobe report 567.

## REFERENCES CITED

- Baldwin, D.A., Frohlinger, T.G., Kendrick, G., McRitchie, W.D., and Zwanig, H.V. (1979) Geology of the Nelson House-Pukatawagan Region (Burntwood Project). Manitoba Mineral Resources Division, Geological Maps 78-3-1 to 78-3-22.
- Belsley, D.A., Kuh, E., and Welsch, R.E. (1980) Regression diagnostics: Identifying influential data and sources of collinearity, 292 p. Wiley, New York.
- Berman, R.G. (1988) Internally-consistent thermodynamic data for minerals in the system  $\text{Na}_2\text{O-K}_2\text{O-CaO-MgO-FeO-Fe}_2\text{O}_3\text{-Al}_2\text{O}_3\text{-SiO}_2\text{-TiO}_2\text{-H}_2\text{O-CO}_2$ . *Journal of Petrology*, 29, 445-522.
- (1990) Mixing properties of Ca-Mg-Fe-Mn garnets. *American Mineralogist*, 75, 328-344.
- Berman, R.G., and Brown, T.H. (1984) A thermodynamic model for multicomponent melts, with application to the system  $\text{CaO-Al}_2\text{O}_3\text{-SiO}_2$ . *Geochimica et Cosmochimica Acta*, 48, 661-678.
- Chatterjee, S., and Hadi, A.S. (1986) Influential observations, high leverage points, and outliers in regression. *Statistical Science*, 1, 379-416.
- Colclough, A.R. (1987) Two theories of experimental error. *Journal of Research of the National Bureau of Standards*, 92, 167-185.
- Cook, R.D., and Weisberg, S. (1982) Residuals and influence in regression, 230 p. Chapman and Hall, New York.
- Fuhrman, M.L., and Lindsley, D.H. (1988) Ternary-feldspar modeling and thermometry. *American Mineralogist*, 73, 201-215.
- Golub, G.H., and Van Loan, C.F. (1989) Matrix computations (2nd edition), 642 p. Johns Hopkins University Press, Baltimore, Maryland.
- Gordon, T.M. (1989) Thermal evolution of the Kiseynew sedimentary gneiss belt, Manitoba: Metamorphism at an Early Proterozoic accretionary margin. In J.S. Daly, R.A. Cliff, and B.W.D. Yardley, Eds., *Evolution of metamorphic belts: Special publication of the Geological Society*, vol. 43, p. 233-243. The Geological Society, London.
- (1992) Generalized thermobarometry: Solution of the inverse geochemical problem using data for individual species. *Geochimica et Cosmochimica Acta*, 56, 1793-1800.
- Grace, A. (1990) Optimization toolbox for use with MATLAB™, 75 p. The MathWorks, Natick, Massachusetts.
- Hamilton, W.C. (1964) Statistics in physical science, 230 p. Roland, New York.
- Hoaglin, D.C., Mosteller, F., and Tukey, J.W. (1983) Understanding robust and exploratory data analysis, 447 p. Wiley, New York.
- Holdaway, M.J., and Mukhopadhyay, B. (1993) Geothermobarometry in pelitic schists: A rapidly evolving field. *American Mineralogist*, 78, 681-693.
- Kohn, M.J., and Spear, F.S. (1991a) Error propagation for barometers. I. Accuracy and precision of experimentally located end-member reactions. *American Mineralogist*, 76, 128-137.
- (1991b) Error propagation for barometers. II. Application to rocks. *American Mineralogist*, 76, 138-147.
- Kretz, R. (1983) Symbols for rock-forming minerals. *American Mineralogist*, 68, 277-279.
- Lewry, J.F., and Stauffer, M.R. (1990) The early Proterozoic Trans-Hudson Orogen of North America. *Geological Association of Canada Special Paper*, 37, 505 p.
- Manitoba Energy and Mines (1988) Bedrock compilation map series (preliminary edition), Kississing, NTS 63N, 1:250000.
- McMullin, D.W.A., Berman, R.G., and Greenwood, H.J. (1991) Calibration of the SGAM thermobarometer for pelitic rocks using data from phase-equilibrium experiments and natural assemblages. *Canadian Mineralogist*, 29, 889-908.
- Robie, R.A., Hemingway, B.S., and Fisher, J.R. (1979) Thermodynamic properties of minerals and related substances at 298.15 K and 1 bar ( $10^5$  pascals) pressure and at higher temperatures (reprinted with corrections 1979). *Geological Survey Bulletin*, 1452, 456 p.
- Rousseeuw, P.J., and van Zomeren, B.C. (1990) Unmasking outliers and leverage points. *Journal of the American Statistical Association*, 85, 633-651.

MANUSCRIPT RECEIVED SEPTEMBER 29, 1993

MANUSCRIPT ACCEPTED MAY 27, 1994

## APPENDIX 1. CALCULATIONS

In the samples studied, the equilibrium assemblage was taken to be quartz + plagioclase + sillimanite + garnet + cordierite ± biotite. For these phases, thermodynamic constants and measured or assumed compositions are available for the ten end-members shown in the composition matrix in Appendix Table 1.

APPENDIX TABLE 1. End-member composition matrix

	qtz	sil	an	alm	prp	grs	Fe-crd	Mg-crd	ann	phl
SiO <sub>2</sub>	1	1	2	3	3	3	5	5	3	3
Al <sub>2</sub> O <sub>3</sub>	0	1	1	1	1	1	2	2	0.5	0.5
FeO	0	0	0	3	0	0	2	0	3	0
MgO	0	0	0	0	3	0	0	2	0	3
CaO	0	0	1	0	0	3	0	0	0	0
K <sub>2</sub> O	0	0	0	0	0	0	0	0	0.5	0.5
H <sub>2</sub> O	0	0	0	0	0	0	0	0	1	1

Note: following Kretz (1983), the abbreviations for end-members begin with lower-case letters.

The matrix in Appendix Table 1 has a rank of six; hence, any six compositionally independent end-members form a basis, i.e., are sufficient to describe the compositions of the remaining four. In biotite-free assemblages, the total number of end-members is eight, and a basis will contain only five end-members. In the ten end-member problem, if sil, alm, grs, Fe-crd, ann, and phl are chosen as a basis, the compositions of the remaining four end-members are expressed as

$$\begin{aligned} \text{qtz} &= -4/5 \text{ sil} - 2/5 \text{ alm} + 3/5 \text{ Fe-crd} \\ \text{an} &= 2/5 \text{ sil} - 2/15 \text{ prp} + 1/3 \text{ grs} + 1/5 \text{ Mg-crd} \\ \text{prp} &= 3/5 \text{ Mg-crd} - 2/5 \text{ prp} - 4/5 \text{ sil} \\ \text{Mg-crd} &= -2/3 \text{ alm} + 2/3 \text{ prp} + \text{Fe-crd}. \end{aligned}$$

The stoichiometric coefficients in these equations also define the relationships among the theoretical chemical potentials of all the species at equilibrium; hence, although there are ten mineral end-members in total, at equilibrium the chemical potentials of only six end-members are sufficient to define all chemical potentials in the system. In these problems, the basis set of end-members was chosen using the algorithm described in Golub and Van Loan (1989, p. 572) as QR with column pivoting.

For each end-member, there are measurements of thermodynamic constants and mineral compositions such that the apparent partial molar free energy of formation [ $g_i(P, T) = g_i^0(P, T) + RT \ln a_i(P, T)$ ] is known as a nonlinear function of the unknowns  $P$  and  $T$ . At equilibrium, each function should equate to the unknown theoretical chemical potential of the respective end-member,  $\mu_i$ .

APPENDIX TABLE 2. Equilibrium equations

$g_{\text{sil}}(P, T)$	$-\mu_{\text{sil}}$									= 0
$g_{\text{alm}}(P, T)$			$-\mu_{\text{alm}}$							= 0
$g_{\text{grs}}(P, T)$					$-\mu_{\text{grs}}$					= 0
$g_{\text{Fe-crd}}(P, T)$						$-\mu_{\text{Fe-crd}}$				= 0
$g_{\text{ann}}(P, T)$							$-\mu_{\text{ann}}$			= 0
$g_{\text{prp}}(P, T)$								$-\mu_{\text{prp}}$		= 0
$g_{\text{qtz}}(P, T)$	$+1/5\mu_{\text{sil}}$		$+2/5\mu_{\text{alm}}$			$-3/5\mu_{\text{Fe-crd}}$			$-\mu_{\text{phl}}$	= 0
$g_{\text{an}}(P, T)$	$-2/5\mu_{\text{sil}}$		$+2/15\mu_{\text{alm}}$		$-1/3\mu_{\text{grs}}$	$-1/5\mu_{\text{Fe-crd}}$				= 0
$g_{\text{prp}}(P, T)$	$+4/5\mu_{\text{sil}}$		$+2/5\mu_{\text{alm}}$			$-3/5\mu_{\text{Fe-crd}}$				= 0
$g_{\text{Mg-crd}}(P, T)$						$-\mu_{\text{Fe-crd}}$	$+2/5\mu_{\text{ann}}$		$-2/5\mu_{\text{prp}}$	= 0

Standard state apparent molar free energies ( $g_i^0$ ) were computed using the apparent free-energy convention of Berman (1988) given by the equation

$$\begin{aligned} g_i^0 &= \Delta_r h_i^0 - TS_i^0 \\ &+ k_0[(T - T_r) - T(\ln T - \ln T_r)] \\ &+ 2k_1[(T^{0.5} - T_r^{0.5}) + T(T^{-0.5} - T_r^{-0.5})] \\ &- k_2[(T^{-1} - T_r^{-1}) - (T/2)(T^{-2} - T_r^{-2})] \\ &- k_3[(T^{-2} - T_r^{-2}) - (T/3)(T^{-3} - T_r^{-3})] \\ &+ \nu_i^0\{(\nu_2/3)(P^3 - P_r^3) + (\nu_1/2 - \nu_2)(P^2 - P_r^2) \\ &+ [1 - \nu_1 + \nu_2 + \nu_3(T - T_r) + \nu_4(T - T_r)^2 \\ &+ (1 - P_r)(P\nu_2 + \nu_1 - \nu_2)](P - P_r)\}. \end{aligned}$$

The required thermochemical data for end-members  $\beta$  quartz, sillimanite, anorthite, grossular, pyrope, almandine, iron-cordierite (anhydrous), magnesium-cordierite (anhydrous), annite, and phlogopite were obtained from the June 1992 version of the data base of Berman (1988).

The activity models used follow Berman and Brown (1984) and hence have the general form  $RT \ln [a_i(P, T)] = A_i + B_i T + C_i P$  where the constants  $A_i$ ,  $B_i$ , and  $C_i$  were computed from the measured compositions given in Tables 1-4. The following activity models were used: for plagioclase, Fuhrman and Lindsley (1988); for biotite, McMullin et al. (1991); for cordierite, ideal mixing; and for garnet, Berman (1990). The partial molar free-energy function for each end-member is the sum of the respective standard state function and activity function.

These relationships give rise to the ten nonlinear equations in eight unknowns shown in Appendix Table 2. Because of systematic and random errors, the computed chemical potentials ( $\mu_i$ ) are not exactly equal to the measured apparent partial molar free energies of formation [ $g_i(P, T) = g_i^0(P, T) + RT \ln a_i$ ] at any  $P$  and  $T$ . An ordinary least-squares solution to these equations can be determined by minimizing the sum of squares of the discrepancies for each equation [ $g_i(P, T) - \Sigma \mu_j$ ]. These quantities are known as residuals or errors. The calculations were carried out using the Broyden-Fletcher-Goldfarb-Shanno quasi-Newton method in the FMINS code in the Matlab Optimization Toolbox (Grace, 1990).