$$\frac{dx_{10}}{dt} = -2x_{10} + 2x_9^2 \quad \text{(III-10)}$$

$$\frac{dx_{11}}{dt} = -3x_{11} + 2x_3 \quad \text{(III-11)}$$

$$\frac{dx_{12}}{dt} = -2x_{12} + 4x_6 \quad \text{(III-12)}$$

$$\frac{dx_{13}}{dt} = -2x_{13} + 4x_6 \quad \text{(III-13)}$$

$$\frac{dx_{14}}{dt} = -3x_{14} + 2x_{10} \quad \text{(III-14)}$$

$$\frac{dx_{15}}{dt} = -4x_{15}x_{22} + 2x_{10} \quad \text{(III-15)}$$

$$\frac{dx_{16}}{dt} = -2k_{16} + 4x_{18} \quad \text{(III-16)}$$

$$\frac{dx_{17}}{dt} = -3x_{17} + 2x_{12} \quad \text{(III-17)}$$

$$\frac{dx_{18}}{dt} = -4x_{18} + 3x_{26} \quad \text{(III-18)}$$

$$\frac{dx_{19}}{dt} = -2x_{19} + 3x_{20} \quad \text{(III-19)}$$

$$\frac{dx_{20}}{dt} = -3x_{20} + 3x_{14} \quad \text{(III-20)}$$

$$\frac{dx_{21}}{dt} = -4x_{21} + 2x_{25} \quad \text{(III-21)}$$

$$\frac{dx_{22}}{dt} = -4x_{15}x_{22} + 4x_{21} \quad \text{(III-22)}$$

$$\frac{dx_{23}}{dt} = -3x_{23} - 4x_{23} + 3x_{11} + 3x_{17} \quad \text{(III-23)}$$

$$\frac{dx_{24}}{dt} = -4x_{24} + 4x_{18} \quad \text{(III-24)}$$

$$\frac{dx_{25}}{dt} = -2x_{25} + 2x_{16} + 2x_{19} \quad \text{(III-25)}$$

$$\frac{dx_{26}}{dt} = -3x_{26} - 2x_{31} \quad \text{(III-26)}$$

$$\frac{dx_{27}}{dt} = -4x_{27}x_{32} + 3x_{31} \quad \text{(III-27)}$$

$$\frac{dx_{28}}{dt} = -2x_{28} + 3x_{23} \quad \text{(III-28)}$$

$$\frac{dx_{29}}{dt} = -3x_{29} + 2x_{28} \quad \text{(III-29)}$$

$$\frac{dx_{30}}{dt} = -4x_{30} + 4x_{24} + 3x_{29} \quad \text{(III-30)}$$

$$\frac{dx_{31}}{dt} = -2x_{31} - 3x_{31} + 4x_{30} \quad \text{(III-31)}$$

$$\frac{dx_{32}}{dt} = -4x_{27}x_{32} + 4x_{23} \quad \text{(III-32)}$$

The system was first partitioned, giving two small loops and one large loop. LOOP 1 consisted of Eqs. III-9, III-10, III-13, III-14 and III-20 in which Eqs. III-9 and III-10 are nonlinear. LOOP 2 was made up of Eqs. III-18, III-24, III-26, III-30 and III-31 all of which are linear. Once these blocks were isolated they were collapsed into single nodes leaving one large cycle, LOOP 3, consisting of all of the remaining equations.

The constant k_2 was varied to make the system stiff. Solutions were obtained for $k_2 = 3$ (non-stiff) and for $k_2 = 50$ and $k_2 = 1000$ which made LOOP 3 stiff. Eigenvalues were not calculated because stiffness was obtained (and proven) in this manner for Example I.

Reduction of Thermodynamic Data by Means of the Multiresponse Maximum Likelihood Principle

The statistical behavior of some common parameter estimation methods is investigated by simulation. The methods based on the maximum likelihood principle give the best estimates of parameters in a given model. An analysis of the possibilities for detecting systematic errors by means of the maximum likelihood method showed that this method gives at best the same information as other well known tests.

S. KEMENY and J. MANCZINGER,
Department of Chemical Engineering, Technical University of Budapest, Budapest, Hungary

S. SKJOLD-JØRGENSEN,
Instituttet for Kemiteknik, Danmarks Tekniske Højskole, Lyngby, Denmark

K. TOTH
G. Richter Pharmaceutical Works, Budapest, Hungary

SCOPE

In recent years several methods have been proposed for the reduction of binary vapor-liquid equilibrium (VLE) information.

A VLE data point gives information either on the full set of variables (x,T,y,P) or on a subset of the variables (e.g., x,T,P). According to the Gibbs' phase rule there are two independent variables (e.g., x,T) in each binary VLE data point. Two vari-

Correspondence concerning this paper should be addressed to S. Kemény.
0001-1541-82-5024-0020-$2.00

ables are therefore dependent variables (e.g., y,P) when all four variables are measured in each data point. In case of measurements of a subset of the information, only one of the variables is a dependent variable (e.g., P).

Some of the methods which have been proposed for the reduction of VLE information are simple while others are more sophisticated. The methods based on the maximum likelihood principle belong to the latter group.

The maximum likelihood principle as applied by several investigators is based on the following assumptions:

1. Experimental errors are small.
2. Experimental errors are independent of one another, both within a given data point and from point to point.
3. The fitting equation (mathematical model) is able to represent the true values with deviations of an order less than the order of the experimental errors.
4. Experimental errors display a normal (Gaussian) distribution with zero mean from the true values.

A number of apparently different methods have been derived using the maximum likelihood principle. Some of these methods include only one response, i.e., one of the two dependent variables (e.g., y) while others include two responses (e.g., y and P).

It has previously not been clear if the use of the more sophisticated methods is only a question of taste or whether we lose some important information when not applying them. Furthermore, it has not been clear whether the apparently different methods based on the maximum likelihood principle and using the same amount of experimental information are identical.

In a recent paper Kemény and Manczinger (1978) present a comparison of several methods for reducing data for subsets. It was shown that any method derived from the maximum likelihood principle leads to the same set of estimated parameters.

This work contains a review of methods based on the multiresponse maximum likelihood principle, i.e., the maximum likelihood principle used for problems with more than one response. A derivation is given of a method which uses the error propagation law. This type of method is compared to an alternative method by Anderson et al. (1978) which does not use the error propagation law. The latter method estimates the true values of the variables as well as the model parameters.

Eight different, widely used methods are compared through simulation studies in order to illustrate their statistical properties and thereby decide which type of method to be recommended. The eight methods include methods with or without weighting and methods for reduction fo full data sets (x,T,y,P) as well as methods for reduction of subsets.

The method by Anderson et al. (1978) gives estimates of the true values of the variables and the parameters. The differences between the estimated true values of the variables and their experimental values might give some information on the sources of systematic errors due to systematic experimental inaccuracies or errors in the model used to reduce the data. We have therefore carried out two simulation studies with this estimation method applying the two above-mentioned types of systematic errors.

CONCLUSIONS AND SIGNIFICANCE

It is shown that any parameter estimation method properly derived from the multiresponse maximum likelihood principle and using the same amount of information will lead to identical estimates of the parameters.

A survey on the computational aspects indicates that methods with error propagation and methods without error propagation, but with estimation of the true values of the variables as well as the parameters, are of comparable complexity from a computational point of view.

A simulation study on several VLE systems shows that a method with proper weighting according to the maximum likelihood principle and the error propagation law is superior to similar methods with improper weighting or no weighting at all. This conclusion holds for methods using subsets of the information as well as for methods using the full information in each data point. Furthermore, methods which use the full information and weights derived from the multiresponse maximum likelihood principle are superior to any other investigated method. These methods are therefore to be recommended since they assure that maximum information is stored in the estimated parameters.

Methods which estimate the true values of the variables and the parameters give poor information on the source of systematic measurement errors. If the model is inherently able to represent the true values of the variables it is possible to obtain an indication of whether systematic measurement errors are present. This information could also be obtained from an ordinary point to point consistency test of the type recommended by Van Ness et al. (1973).

Systematic errors introduced by model deficiencies are detectable by means of the estimated true values of the variables. This detection assumes that the data contains no systematic measurement errors since the two types of systematic errors, i.e. model deficiencies and systematic measurement errors, cannot be separated. Detection of model errors is especially useful for model comparison. For this purpose the value of the objective function is as informative as the estimated true values of the variables.

The use of a method which leads to estimation of the true values of the variables gives no information which could not be obtained by other well known tests of the data or the considered models. This type of method gives merely a possibility of performing consistency tests, model evaluation, and parameter estimations using the same parameter estimation method.

The multiresponse maximum likelihood methods presented in this paper can readily be extended to other types of information. Although this work has focused on problems with two independent and two dependent variables the principles presented in the derivation of the methods are general and applicable to any number of dependent as well as independent variables.

The sophisticated treatment of experimental data, which is here exemplified by the principle of maximum likelihood, does not change the fact that there is no way of making good parameters out of low quality data.

REVIEW OF SOME POSSIBLE MAXIMUM LIKELIHOOD METHODS

In thermodynamics a number of papers have during the last decade dealt with parameter estimation based on the maximum likelihood principle. As a result of these works a variety of expressions has appeared stating which functions to be minimized

in order to compute the optimal values of the parameters in question. Throughout this paper such expressions are called criteria. Below, we present a few well known criteria and we give a derivation of another possible criterion.

Schubert's method

Schubert (1974) was the first to propose a maximum likelihood criterion for the reduction of all four variables of a vapor-liquid equilibrium measurement. His work was based on results of Box and Draper (1965), and Hunter (1967). Schubert used P, G^E/R_G, $ln(\gamma_1/\gamma_2)$, and y as dependent variables. His determinant criterion does not contain information on the measurement precision and on the mutual dependence of the variables (the covariance matrix). His criterion is based on the assumption that the covariance matrix is constant. This is not true, even if the variances of the direct measurements (x,T,y and P) are constant (Van Ness et al., 1973; Ulrichson and Stevenson, 1972.) Furthermore, the Schubert criterion contains four dependent variables instead of two. A binary VLE system has two degrees of freedom, and two variables are therefore independent while the remaining two must be dependent variables.

Method of Kemény and Manczinger

In a recent paper Kemény and Manczinger (1978) heuristically extended the methods of Box and Draper (1965), Hunter (1967), and Schubert (1974) taking into account the variances of both dependent and independent variables and allowing them to vary from measurement to measurement. Their criterion is,

$$\sum_{i=1}^{n} \frac{\left(\ln \frac{\gamma_{1i}}{\gamma_{2i}} - \ln \frac{\hat{\gamma}_{1i}}{\hat{\gamma}_{2i}}\right)^2}{\sigma^2_{(\Delta \ln \gamma_1/\gamma_2)i}} + \sum_{i=1}^{n} \frac{(P_i - \hat{P}_i)^2}{\sigma^2_{(\Delta P)i}} = \min \tag{1}$$

The σ^2_Δ variances are to be calculated by the error propagation law,

$$\sigma^2_{(\Delta Y_l)i} = \sum_{j=1}^{4} \left(\frac{\partial(\Delta \hat{Y}_l)}{\partial Z_j}\right)^2_i \sigma^2_{(Z_j)i} \tag{2}$$

where:

$\Delta \hat{Y}_l$ = difference between the experimental value Y_l and the estimated value $\hat{Y}_l$ of the l'th dependent variable. This difference is called a residual.

Z_j = directly measured variable (x,T,y, or P) with index j.

Equation 1 neglects the slight interdependence of the dependent variables due to errors in the independent variables, as represented by the covariance term

$$\sigma_{(\Delta \ln \gamma_1/\gamma_2, \Delta P)i}.$$

Method of Fabries and Renon

Bard and Lapidus (1968) present the following criterion,

$$\sum_{i=1}^{n} (\underline{\Delta \hat{Y}})_i^T \, \underline{\underline{\mathrm{Cov}}}_i^{-1} (\underline{\Delta \hat{Y}})_i$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{r} \sum_{k=1}^{r} \mathrm{Cov}_i^{jk} (\Delta \hat{Y}_j)_i (\Delta \hat{Y}_k)_i = \min \tag{3}$$

where

$\underline{\underline{\mathrm{Cov}}}_i$ = covariance matrix of the $(\underline{\Delta \hat{Y}})_i$ residuals. Cov_i^{jk} is the j,k'th element of $\underline{\underline{\mathrm{Cov}}}_i^{-1}$

It is essential for the derivation of Eq. 3 that the experimental errors of the dependent variables are independent of each other and that they follow normal distributions. Fabries and Renon (1975) take the two activity coefficients as dependent variables (Y_{1i} = γ_{1i},Y_{2i} = γ_{2i}). The activity coefficients contain products and other nonlinear combinations of the normally distributed variables, x, T, y and P. It is assumed that the experimental errors of the activity coefficients follow the normal distribution law and that the errors are independent of each other. Criterion 3 is heuristically extended to the case of independent variables subject to errors. The elements of the $\underline{\underline{\mathrm{Cov}}}_i$ covariance matrix are expressed from the error propagation law. The diagonal elements are calculated from Eq. 2 while the off-diagonal elements are calculated from

$$\sigma_{(\Delta Y_l, \Delta Y_k)i} = \sum_{j=1}^{4} \left[\frac{\partial(\Delta \hat{Y}_l)}{\partial Z_j}\right]_i \left[\frac{\partial(\Delta \hat{Y}_k)}{\partial Z_j}\right]_i \sigma^2_{(Z_j)i} \tag{4}$$

Assuming that the errors on the measured values of x, T, y, and P follow the normal distribution law the resulting errors on γ_1 and γ_2 (which cannot be presented by a linear combination of x, T, y, and P) do not follow the normal distribution law, and they are not independent of each other. Nevertheless, the resulting criterion 5 of Fabries and Renon represents a correct combination of the principle of maximum likelihood with error propagation.

$$\sum_{i=1}^{n} \frac{(\Delta \hat{Y}_1)_i^2 \sigma^2_{(\Delta \hat{Y}_2)i} + (\Delta \hat{Y}_2)_i^2 \sigma^2_{(\Delta \hat{Y}_1)i} - 2(\Delta \hat{Y}_1)_i (\Delta \hat{Y}_2)_i \sigma_{(\Delta \hat{Y}_1, \Delta \hat{Y}_2)i}}{\sigma^2_{(\Delta \hat{Y}_1)i} \sigma^2_{(\Delta \hat{Y}_2)i} - (\sigma_{(\Delta \hat{Y}_1, \Delta \hat{Y}_2)i})^2} = \min \tag{5}$$

Based on the estimated parameters Fabries and Renon estimate the true values of the measured variables minimizing the following sum of squares in each measurement point,

$$SSQ = \sum_{j=1}^{4} \frac{(Z_j - \hat{Z}_j)^2}{\sigma^2_{Z_j}} \tag{6}$$

with two equality constraints due to the equilibrium conditions. (See Eqs. 8 and 9.)

Method of Anderson et al.

Anderson et al. (1978) use the maximum likelihood criterion for variables taken from normal distributions with known variances,

$$\sum_{i=1}^{n} \sum_{j=1}^{r} \frac{(\Delta \hat{Y}_j)_i^2}{\sigma^2_{(Y_j)i}} + \sum_{i=1}^{n} \sum_{k=1}^{s} \frac{(\Delta \hat{X}_k)_i^2}{\sigma^2_{(X_k)i}} = \min \tag{7}$$

where Y designates dependent variables (j = 1,2, . . . , r), and X designates independent variables (k = 1,2, . . . , s). For VLE data x and T are chosen as independent variables and y and P as dependent variables. The minimization of Eq. 7 with respect to the parameters and the variables is carried out taking into consideration the two constraints,

$$P = x_1 \gamma_1 P_1^0 \frac{\phi_1^0}{\phi_1} + x_2 \gamma_2 P_2^0 \frac{\phi_2^0}{\phi_2} \tag{8}$$

$$y_1 = \frac{1}{1 + \phi_1 \phi_2^0 x_2 \gamma_2 P_2^0 / \phi_2 \phi_1^0 x_1 \gamma_1 P_1^0} \tag{9}$$

where the Poynting correction terms have been neglected. The criterion of Anderson et al. (1978) leads to estimates of the true values of the variables as well as the parameters. Equation 7 may be extended to other types of information related to the model under consideration (e.g., H^E, LLE, γ^∞ etc. when a G^E model is considered). The additional terms in the criterion are defined by the additional measured variables plus constraints.

A Derivation of a Criterion with Error Propagation (cf., Kemény, 1980)

Box (1970) gives the following maximum likelihood criterion,

$$\sum_{i=1}^{n} \sum_{j=1}^{r} \sum_{k=1}^{r} \sigma_i^{jk} (\Delta \hat{Y}_j)_i (\Delta \hat{Y}_k)_i = \min \tag{10}$$

which is a generalized version of the criterion 7 used by Anderson et al. (1978).

The errors in the dependent variables are defined by,

$$\epsilon_{ji} = Y_{ji} - Y^*_{ji} \tag{11}$$

where Y^*_{ji} stands for the true value. The expected values (E) of the errors are assumed to be given by,

$$E(\epsilon_{ji}) = 0; \quad E(\epsilon_{ji}\epsilon_{ki}) = 0 \text{ if } j \neq k; \quad E(\epsilon_{ji}\epsilon_{jl}) = 0 \text{ if } i \neq l$$

Equation 10 is transformed in such a way that one minimizes not the residuals of the directly measured variables $(\Delta\hat{Y}_j)_i$, but the residuals of some functions of them $(\Delta\hat{F}_j)_i$. The extended criterion takes into account the errors in the independent as well as the dependent variables. For the derivation of Eq. 12 see Appendix 1.

$$\sum_{i=1}^{n}\sum_{j=1}^{r}\sum_{k=1}^{r}\sigma_i^{jk}\,(\Delta\hat{F}_j)_i(\Delta\hat{F}_k)_i = \min \tag{12}$$

where $\hat{\sigma}_i^{jk}$ is an element of the inverse of the covariance matrix $\underline{\underline{\text{Cov}}}_i(\underline{\Delta F})$. The general element of the covariance matrix is defined by,

$$(\sigma_{jk})_i = E[(\Delta F_j)_i(\Delta F_k)_i] \tag{13}$$

The resultant errors $(\Delta F_j)_i$ have two sources: the errors in the Y_{ji} dependent variables (ϵ_{ji}) and the errors in the X_{li} independent variables (δ_{li}) in point i which have the same expected values as ϵ_{ji} (Eq. 11). Using the error propagation law one gets,

$$(\Delta F_j)_i = \sum_{k=1}^{r}\left[\frac{\partial(\Delta F_j)}{\partial Y_k}\right]_i \epsilon_{ki} + \sum_{l=1}^{s}\left[\frac{\partial(\Delta F_j)}{\partial X_l}\right]_i \delta_{li} \tag{14}$$

To express the expected values defined by Eq. 13 let us rewrite formula 14,

$$(\Delta F_j)_i = \sum_{k=1}^{r+s} C_{jki}w_{ki} \tag{15}$$

where

$$C_{jki} = \left[\frac{\partial(\Delta F_j)}{\partial Y_k}\right]_i; \quad w_{ki} = \epsilon_{ki} \quad \text{for} \quad k = 1, \ldots, r$$

$$C_{jki} = \left[\frac{\partial(\Delta F_j)}{\partial X_{k-r}}\right]_i; \quad w_{ki} = \delta_{(k-r)i} \quad \text{for} \quad k = 1 + r, \ldots, r + s$$

The elements of the covariance matrix are,

$$(\sigma_{jk})_i = \sum_{l=1}^{r+s} C_{jli}C_{kli}\sigma_{li}^2$$

with

$$\sigma_{li}^2 = E(w_{li}^2)$$

In the general case $(\Delta F_j)_i$ is a function of all the dependent and independent variables,

$$(\Delta F_j)_i = g_j(Y_{1i}, Y_{2i}, \ldots, Y_{ri}; \quad X_{1i}, X_{2i}, \ldots, X_{si}) \tag{17}$$

If the error $(\Delta F_j)_i$ is a function of only one of the dependent variables and of the independent variables of point i, that is,

$$(\Delta F_j)_i = g_j(Y_{ji}, X_{1i}, X_{2i}, \ldots, X_{si}) \tag{18}$$

the diagonal elements of the covariance will be given by,

$$(\sigma_{jj})_i = \left[\frac{\partial(\Delta F_j)}{\partial Y_j}\right]_i^2 \sigma^2_{(Y_j)i} + \sum_{l=1}^{s}\left[\frac{\partial(\Delta F_j)}{\partial X_l}\right]_i^2 \sigma^2_{(Xl)i} \tag{19}$$

The off-diagonal elements are given by,

$$(\sigma_{jk})_i = \sum_{l=1}^{s}\left[\frac{\partial(\Delta F_j)}{\partial X_l}\right]_i\left[\frac{\partial(\Delta F_k)}{\partial X_l}\right]_i \sigma^2_{(Xl)i} \tag{20}$$

since

$$\left[\frac{\partial(\Delta F_j)}{\partial Y_k}\right]_i = 0 \text{ if } j \neq k$$

and according to Eq. 18

$$C_{jli}C_{kli} = 0 \text{ if } j \neq k,\ j \leq r, \text{ and } k \leq r$$

In the more general case (Eq. 17) where $(\Delta F_j)_i$ contains several dependent variables the expression (Eq. 20) must be extended by,

$$\sum_{l=1}^{r}\left[\frac{\partial(\Delta F_j)}{\partial Y_l}\right]_i\left[\frac{\partial(\Delta F_k)}{\partial Y_l}\right]_i \sigma^2_{(Yl)i} \tag{21}$$

whereas the first term on the right hand side of Eq. 19 is substituted by,

$$\sum_{m=1}^{r}\left[\frac{\partial(\Delta F_j)}{\partial Y_m}\right]_i^2 \sigma^2_{(Ym)i} \tag{22}$$

This correct variant of the estimation criterion (Eqs. 12, 17, and 19–22) was used by Péneloux et al. (1975) without presenting the derivation of their formulas. Their F transforms of the dependent variables are, $F_1 = \ln\gamma_1/\gamma_2$, $F_2 = G^E/R_GT$. Fabries and Renon's (1975) criterion corresponds to the transforms, $F_1 = \gamma_1$, $F_2 = \gamma_2$. We make the following choice of the transforms,

$$F_1 = \ln\,\gamma_1/\gamma_2; \quad F_2 = P \tag{23}$$

which is advantageous because F_1 contains only the first dependent variable (y), and F_2 contains only the second one (P).
The resultant criterion is,

$$\sum_{i=1}^{n}\frac{\sigma^2_{(\Delta P)i}(\Delta\ln\,\gamma_1/\gamma_2)_i^2}{D} + \sum_{i=1}^{n}\frac{\sigma^2_{(\Delta\ln\gamma 1)/\gamma 2)i}(\Delta P)_i^2}{D} - 2\sum_{i=1}^{n}\frac{\sigma_{(\Delta\ln\gamma 1/\gamma 2,\Delta P)i}(\Delta\ln\,\gamma_1/\gamma_2)_i(\Delta P)_i}{D} = \min \tag{24}$$

where

$$D = \underline{\underline{\text{Cov}}}_i = \sigma^2_{(\Delta\ln\gamma 1/\gamma 2)i}\sigma^2_{(\Delta P)i} - (\sigma_{(\Delta\ln\gamma 1/\gamma 2,\Delta P)i})^2 \tag{25}$$

since

$$\underline{\underline{\text{Cov}}}_i = \begin{pmatrix}\sigma^2_{(\Delta\ln\gamma 1/\gamma 2)i} & \sigma_{(\Delta\ln\gamma 1/\gamma 2,\Delta P)i}\\ \sigma_{(\Delta\ln\,\gamma 1/\gamma 2,\Delta P)i} & \sigma^2_{(\Delta P)i}\end{pmatrix} \tag{26}$$

The elements of the $\underline{\underline{\text{Cov}}}_i$ covariance matrix are given by,

$$\sigma^2_{(\Delta\ln\gamma 1/\gamma 2)i} = \left\{\left[\frac{\partial(\ln\gamma_1/\gamma_2)}{\partial x}\right]_i - \left[\frac{\partial(\ln\hat{\gamma}_1/\hat{\gamma}_2)}{\partial x}\right]_i\right\}^2\sigma^2_{xi} + \left\{\left[\frac{\partial(\ln\gamma_1/\gamma_2)}{\partial T}\right]_i - \left[\frac{\partial(\ln\hat{\gamma}_1/\hat{\gamma}_2)}{\partial T}\right]_i\right\}^2\sigma^2_{Ti} + \left[\frac{\partial(\ln\gamma_1/\gamma_2)}{\partial y}\right]_i^2\sigma^2_{yi} \tag{27}$$

$$\sigma^2_{(\Delta P)i} = \left(\frac{\partial\hat{P}}{\partial x}\right)_i^2\sigma^2_{xi} + \left(\frac{\partial\hat{P}}{\partial T}\right)_i^2\sigma^2_{Ti} + \sigma^2_{\hat{P}i} \tag{28}$$

$$\sigma_{(\Delta\ln\gamma 1/\gamma 2,\Delta P)i} = \left\{\left[\frac{\partial(\ln\gamma_1/\gamma_2)}{\partial x}\right]_i - \left[\frac{\partial(\ln\hat{\gamma}_1/\hat{\gamma}_2)}{\partial x}\right]_i\right\}\left(\frac{\partial\hat{P}}{\partial x}\right)_i\sigma^2_{xi} + \left\{\left[\frac{\partial(\ln\gamma_1/\gamma_2)}{\partial T}\right]_i - \left[\frac{\partial(\ln\hat{\gamma}_1/\hat{\gamma}_2)}{\partial T}\right]_i\right\}\left(\frac{\partial\hat{P}}{\partial T}\right)_i\sigma^2_{Ti} \tag{29}$$

The proposed criterion (Eq. 24) differs from the criteria of Fabries and Renon (1975) and Peneloux et al. (1975) by the choice of F transforms, and their common basis has a strict mathematical derivation. (See Appendix 1.) The criterion is also identical with that of Anderson et al. (1978) as indicated in Appendix 1.

Criterion 24 can be extended to other types of information related to the considered model. The necessary additional terms in the extended criterion must be derived through proper use of the error propagation law.

COMPUTATIONAL ASPECTS

As indicated in the previous section and in Appendix 1 any correctly derived multiresponse maximum likelihood criterion based on a given amount of information will lead to identical model parameters. However, each of the possible criteria will possess advantages and disadvantages when combined with a specific optimization algorithm.

We will therefore give a short survey of the computational aspects of two alternative criteria, namely the criterion by Anderson et al. (1978) and the criterion developed in this paper using the error propagation law (Eq. 24).

Anderson et al. Criterion

The criterion by Anderson et al. (Eqs. 7–9) leads to minimization of a sum of squares with two implicit constraints in each data point.

The constraints are implicit since the fugacity coefficients depend on y and P in addition to T. A common solution to constrained optimization problems is the use of Lagrange multipliers (Kowalik and Osborne, 1968). Each constraint is introduced into the criterion by means of an additional parameter, the Lagrange multiplier. The resultant minimization problem is solved without any discrimination between dependent and independent variables.

However, constrained optimization is in general very cumbersome when compared to unconstrained optimization and a transformation of the problem is therefore desirable. Anderson et al. (1978) note that the implicitness of the constraints can be eliminated through the following assumption,

$$\phi(\hat{y},\hat{T},\hat{P})_i \cong \phi(y,T,P)_i \tag{30}$$

Equation 30 states that the fugacity coefficient at the estimated true values of the variables is about equal to the fugacity coefficient at the experimental values of the same quantities. The constraints (Eqs. 8 and 9) are now explicit equations of the following form,

$$\hat{Y}_{ji} = \hat{Y}_{ji}(\hat{X}_i,\hat{\theta}) \quad j = 1,\ldots,r \tag{31}$$

where $\hat{\theta}$ is the estimated parameter vector.

The constraints can now be substituted into the criterion 7, and the minimization problem is henceforth unconstrained. (See Eq. 32.) To solve the minimization problem Anderson et al. (1978) linearize the constraints (Eq. 31) by expanding them into Taylor series and inserting the resultant constraints into the criterion. The resultant criterion is differentiated with respect to the elements of $\hat{\theta}$ and $\hat{X}$ and the derivatives are set equal to zero. This set of linear equations is solved for $\hat{\theta}$ and $\hat{X}$, and the solution is substituted into the linearized constraints and so on. This iteration procedure is of the Newton type, and the convergence is therefore fast from good initial guesses while it may fail starting from poorer initial guesses.

Skjold-Jørgensen (1980) combines the effective and reliable algorithm for unconstrained optimization of sums of squares by Levenberg and Marquardt (Kowalik and Osborne, 1968) with the unconstrained transform of the criterion by Anderson et al.

$$\sum_{i=1}^{n}\sum_{j=1}^{r}\frac{(\Delta\hat{Y}_j(\hat{\underline{X}}_i,\hat{\underline{\theta}}))^2}{\sigma^2_{(Y_j)i}} + \sum_{i=1}^{n}\sum_{k=1}^{s}\frac{(\Delta\hat{X}_k)_i^2}{\sigma^2_{(X_k)i}} = \min \tag{32}$$

The optimization algorithm requires first derivatives of criterion 32 with respect to $\hat{X}$ and $\hat{\theta}$. The resultant Jacobian matrix is sparse since the variables in a specific data point are independent of the variables in all other data points. By taking this sparsity into consideration the required number of computer storage units is reduced by more than 90%. Furthermore, Skjold-Jørgensen divides the minimization into two steps.

In the first step the $\hat{\underline{\theta}}$ parameters are estimated with the independent variables $\underline{\underline{X}}$ fixed at their experimental values. Each iteration requires only knowledge of the partial derivatives with respect to the parameters $\hat{\underline{\theta}}$. The second step of the minimization procedure leads to the final estimates of the parameters $\hat{\underline{\theta}}$ as well as the independent variables $\hat{\underline{\underline{X}}}$. This step requires the partial derivatives with respect to the independent variables in addition to the partial derivatives with respect to the parameters. However, according to our experience the second step is finished within a few iterations (typically five iterations) due to the excellent initial values of the optimized quantities and to the linear end convergence of the Levenberg-Marquardt method. The described parameter estimation procedure converges even from poor initial estimates of the parameters $\hat{\underline{\theta}}$.

Criterion Derived in this Article

Criterion 24 is no sum of squares, and the special optimization algorithms for such problems (e.g., Levenberg-Marquardt) are therefore not applicable. The criterion is to be minimized only with respect to the parameters $\hat{\underline{\theta}}$. The true values of the independent variables $\underline{\underline{X}}$ are not estimated, and the complexity of the problem is thereby considerably reduced. Information about the influence of the independent variables is included in the criterion through the error propagation (Eqs. 27–29). The minimization of Eq. 24 leads inevitably to the calculation of the first partial derivatives of the dependent variables with respect to the independent variables.

Kemény (1980) uses a search method to minimize Eq. 24 (e.g., the Hooke and Jeeves method, 1961). No partial derivatives with respect to the estimated parameters $\hat{\underline{\theta}}$ are therefore required. The partial derivatives needed for the error propagation (Eqs. 27–29) depend on the estimated parameters, and they must therefore be updated after each step of the iterations, at least in principle. However, the optimization of $\hat{\underline{\theta}}$ is only moderately dependent of the error propagation. Hence the partial derivatives with respect to the independent variables are not updated in each iteration step.

The parameters $\hat{\underline{\theta}}$ are estimated with fixed values of the resultant variances (Eqs. 27–29). New partial derivatives for the error propagation are calculated from the present estimate of the parameters, and the updated resultant variances are used for a new estimation of the parameters. This procedure is carried out until convergence is achieved. According to our experience the partial derivatives are recalculated three to five times for a typical problem. The above-mentioned moderate influence of the error propagations supports furthermore the use of numerical derivatives for the calculation of resultant variances.

The use of a relatively slow optimization algorithm of the direct search type is justified by a relatively low number of computations of partial derivatives. The search methods converge even from poor initial estimates of the parameters $\hat{\underline{\theta}}$.

As suggested in the derivation of criterion 24 and in Appendix 1 any criterion based on the multiresponse maximum likelihood principle will lead to identical estimates of the parameters in question. Table 1 illustrates numerically this identity between the criterion 24 derived in this paper and the criterion 7 by Anderson et al. (1978).

SIMULATION STUDIES

Reduction of VLE data was used to compare the quality of parameters estimated from different criteria. In these studies the parameters to be estimated are the interaction energies of a model for the excess Gibbs' energy (e.g., Wilson or UNIQUAC). If the models described phase equilibria perfectly it would be possible to compare the criteria by extrapolation, i.e., by predicting VLE information at conditions not included in the reduced data set or

TABLE 1. PARAMETERS ESTIMATED BY TWO CRITERIA BASED ON THE MULTIRESPONSE MAXIMUM LIKELIHOOD PRINCIPLE.

Assumed standard deviations of the measurements: $\sigma_x = \sigma_y = 0.0001$, $\sigma_T = 0.01°K$, $\sigma_P = 13.3$ Pa

Systems	Page in Gmehling and Onken§	UNIQUAC parameters [J/mol] Eq. 24 a_{12}	Eq. 24 a_{21}	Eq. 7 a_{12}	Eq. 7 a_{21}
Ethyl Iodide—Ethanol	318	3037.1	−112.8	3037.5	−112.9
Chloroform—Methanol	19	5287.0	−1115.9	5287.7	−1115.7
Acetone—Methanol	85	2875.5	−1015.0	2875.8	−1015.1

§ Experimental (x,T,y,P) data were taken from the VLE data collections by Gmehling and Onken (1977). The same source provided the r and q values for the UNIQUAC model and the Antoine constants.

by predicting properties which are related to VLE through thermodynamic relations (e.g., H^E). However, none of the existing models are error-free. The suggested possibility for comparison is therefore hypothetical since the extrapolations will inevitably contain systematic model errors which cannot be separated from the errors in the parameters.

The only real possibility is therefore a statistical investigation of the estimated parameters. If the probability density functions corresponding to the criteria were simple mathematical expressions it would be natural to compare the parameter estimators directly. However, since we deal with nonlinear models and relatively complicated relationships between the variables in question it would in practice be impossible. We have therefore two choices, a) Repetition of the same experiment several times. b) Monte Carlo simulation of the experiments. We chose alternative b).

Our simulation studies have the following two main purposes,

1) Comparison of the maximum likelihood criteria (Eqs. 5, 7 and 24) including all observed variables with other common criteria.
   a) Comparison of weighted and unweighted criteria.
   b) Examination of the improvement obtained by reduction of data for all observed variables compared to reduction of subsets.

   For this part of our simulation studies only random errors were added to the true values of the observed variables.
2) Investigation of the possibilities to detect systematic errors in the measured data or in the applied model.

The following procedure was used for our simulations (cf., Bard, 1974)

1) Define the system. Assign values of the parameters $\underline{\theta}^*$ and of the variances of the probability density functions of measurement errors (mean values are zero).
2) Generate the true values of variables in the i'th data point: $\underline{Z}_i^*$ $(i = 1, \ldots, n)$.
3) Generate pseudo random numbers of the prescribed probability density function and add them as errors (ϵ_i and δ_i) to the $\underline{Z}_i^*$ true values. The resultant values are the simulated "measurement data" $\underline{Z}_i$.
4) Estimate the p parameters from the $\underline{Z}_i$ data. This procedure gives the $\hat{\underline{\theta}}_j$ estimated parameter vector and eventually the estimated values of the variables $\hat{\underline{Z}}_i$ if criterion 7 is used.
5) Repeat steps 3 and 4 N times generating new random numbers in each cycle.
6) Calculate the characteristics of the empirical probability distribution of the estimates,

$$E(\hat{\underline{\theta}}) = \bar{\underline{\theta}} = 1/N \sum_{j=1}^{N} \hat{\underline{\theta}}_j \tag{33}$$

(the estimated expected value of the estimator)

$$\widehat{\underline{\underline{\mathrm{COV}}}}(\hat{\theta}) = \frac{1}{N-p} \sum_{j=1}^{N} (\hat{\underline{\theta}}_j - \hat{\underline{\theta}})(\hat{\underline{\theta}}_j - \hat{\underline{\theta}})^T \tag{34}$$

(the estimated covariance matrix of the estimator), where N is the total number of repetitions. We often used $N = 50$.
The estimated bias is,

$$b = \bar{\underline{\theta}} - \underline{\theta}^* \tag{35}$$

The mean square error is useful for characterizing the estimators because it combines the covariance and the bias of the estimators

$$MSE = E[(\hat{\underline{\theta}} - \underline{\theta}^*)^T(\hat{\underline{\theta}} - \underline{\theta}^*)] \tag{36}$$

An estimate of MSE is given by,

$$\widehat{MSE} = \frac{1}{N} \sum_{j=1}^{N} (\hat{\underline{\theta}}_j - \underline{\theta}^*)^T(\hat{\underline{\theta}}_j - \underline{\theta}^*) \tag{37}$$

Comparison of 8 Different Criteria with Datasets Subject to Random Errors

The criteria which we have investigated are summarized in Table 2. A detailed description of the criteria A-F is given by Kemény and Manczinger (1978).

The first two criteria of Table 2 use the x,T,P subset of the measured data while the criteria C and D use the x,T,y subset. Due to proper weighting according to the maximum likelihood principle with error propagation we recommend criteria B and D for subsets (Kemény and Manczinger, 1978). Criteria E and F were used by Sutton and MacGregor (1977). These criteria include only one response while they make use of data for all the four variables (x,T,y, and P). Criterion G was proposed by Kemény and Manczinger (1978). This criterion neglects the covariance of the system responses $\ln\gamma_1/\gamma_2$ and P due to errors in common independent variables.

The last criterion H, Eq. 24, is identical to several other criteria

TABLE 2. ESTIMATION CRITERIA FOR REDUCTION OF VLE DATA. THE CRITERIA A-D REDUCE SUBSETS WHILE CRITERIA E-H REDUCE THE FULL DATA SETS (x,T,y,P)

No.	Objective Function
A	$\sum_{i=1}^{n} (P_i - \hat{P}_i)^2$
B	$\sum_{i=1}^{n} (P_i - \hat{P}_i)^2/\sigma^2_{(\Delta P)i}$
C	$\sum_{i=1}^{n} (\ln\gamma_{1i}/\gamma_{2i} - \ln\hat{\gamma}_{1i}/\hat{\gamma}_{2i})^2$
D	$\sum_{i=1}^{n} (\ln\gamma_{1i}/\gamma_{2i} - \ln\hat{\gamma}_{1i}/\hat{\gamma}_{2i})^2/\sigma^2_{(\Delta\ln\gamma_1/\gamma_2)i}$
E	$\sum_{i=1}^{n} (\hat{y}_i - y_i)^{2*}$
F	$\sum_{i=1}^{n} (\hat{y}_i - y_i)^2/\sigma^2_{(\Delta y)i}{}^*$
G	$\sum_{i=1}^{n} (\ln\gamma_{1i}/\gamma_{2i} - \ln\hat{\gamma}_{1i}/\hat{\gamma}_{2i})^2/\sigma^2_{(\Delta\ln\gamma_1/\gamma_2)i} + \sum_{i=1}^{n} (P_i - \hat{P}_i)^2/\sigma^2_{(\Delta P)i}$
H	$\sum_{i=1}^{n} \frac{\sigma^2_{(\Delta P)i}(\ln\gamma_{1i}/\gamma_{2i} - \ln\hat{\gamma}_{1i}/\hat{\gamma}_{2i})^2 + \sigma^2_{(\Delta\ln\gamma_1/\gamma_2)i}(P_i - \hat{P}_i)^2}{\sigma^2_{(\Delta\ln\gamma_1/\gamma_2)i}\sigma^2_{(\Delta P)i} - [\sigma_{(\Delta\ln\gamma_1/\gamma_2,\Delta P)i}]^2} - \sum_{i=1}^{n} \frac{2\sigma_{(\Delta\ln\gamma_1/\gamma_2,\Delta P)i}(\ln\gamma_{1i}/\gamma_{2i} - \ln\hat{\gamma}_{1i}/\hat{\gamma}_{2i})(P_i - \hat{P}_i)}{\sigma^2_{(\Delta\ln\gamma_1/\gamma_2)i}\sigma^2_{(\Delta P)i} - [\sigma_{(\Delta\ln\gamma_1/\gamma_2,\Delta P)i}]^2}$

* $\hat{y}_i = \frac{x_{1i}\hat{\gamma}_{1i}P^0_{1i}\phi^0_{1i}}{P_i\phi_{1i}}$

TABLE 3. SIMULATION RESULTS FOR AN ACETONE-BENZENE SYSTEM AT $1.013 \cdot 10^5$ PA. THE SIMULATED MEASUREMENTS CONTAIN ONLY RANDOM ERRORS. $N = 50$, $n = 13$.

Criterion	$\bar{l}_1$ (J/mol)	$\bar{l}_2$ (J/mol)	$\hat{\sigma}^2_{\bar{l}_1}$ $\times 10^{-3}$	$\hat{\sigma}^2_{\bar{l}_2}$ $\times 10^{-3}$	$\hat{\sigma}_{\bar{l}_1,\bar{l}_2}$ $\times 10^{-3}$	$\widehat{MSE}$ $\times 10^{-3}$
A	2506.7	−823.2	22.8	8.76	−14.0	31.8
B	2507.6	−823.4	23.2	8.91	−14.2	32.3
C	2492.2	−816.5	21.6	7.82	−12.5	29.4
D	2502.2	−822.5	4.24	1.55	−2.41	5.88
E	2505.8	−823.8	9.58	3.02	−5.27	12.8
F	2510.1	−825.8	9.53	2.98	−5.23	12.8
G	2503.5	−823.1	3.58	1.24	−2.02	4.94
H	2503.7	−823.3	3.54	1.25	−2.03	4.92

True Wilson Parameters: $l_1^* = \lambda_{12}^* - \lambda_{11}^* = 2493.2$ J/mole
$l_2^* = \lambda_{12}^* - \lambda_{22}^* = -818.5$ J/mole
Applied Random Errors: $\sigma_x = \sigma_y = 0.001$
$\sigma_T = 0.1$°K
$\sigma_P = 133.0$ Pa

based on the multiresponse maximum likelihood principle (e.g., Eqs. 5 and 7).

The first part of our investigations is comparable to the investigations by Sutton and MacGregor (1977) and by van Ness et al. (1978). The former authors compared only two estimation criteria. Our work can be regarded as an extension of their studies. We have not adopted the method of van Ness et al. for our simulation studies since their method do not guarantee statistical independence of the data sets from different cycles.

Table 3 shows the details of a simulation for an acetone-benzene isobaric system. All the estimates of $\bar{l}_1$ and $\bar{l}_2$ are unbiased (checked by t-tests) and strongly interrelated since the covariance term has a very large value. For this system weighting does not improve the properties of the estimated parameters when the subset x,T,P is used. This is due to the fact that the resultant variance $\sigma^2_{(\Delta P)}$ is almost constant (Figure 1). Figures 1 and 2 show the resultant variances of the transformed dependent variables of criterion H. The variations depend on the system in question as well as on the assumed variances of the measurement errors.

Results in terms of mean square errors for simulations with 3 different binaries and various source variances are shown in Table 4. For subsets the weighted criteria (B and D) give better results than the corresponding unweighted criteria (A and C). The improvements from weighting depend on the accuracy of the data ($\sigma_x,\sigma_T,\sigma_y,\sigma_P$) and on the system in question. Criterion H of Table 4 is the best for reduction of complete data sets (x,T,y,P). Criterion G is slightly inferior due to the missing covariance term. The weighted criterion F proposed by Sutton and MacGregor (1977) is in most cases better than the corresponding unweighted criterion E. However, in some cases the result is worse than the result obtained with the subset x,T,P, i.e., from criterion B. This is a consequence of the erroneous use of information on the vapor phase composition and the result may be worse than by neglecting this

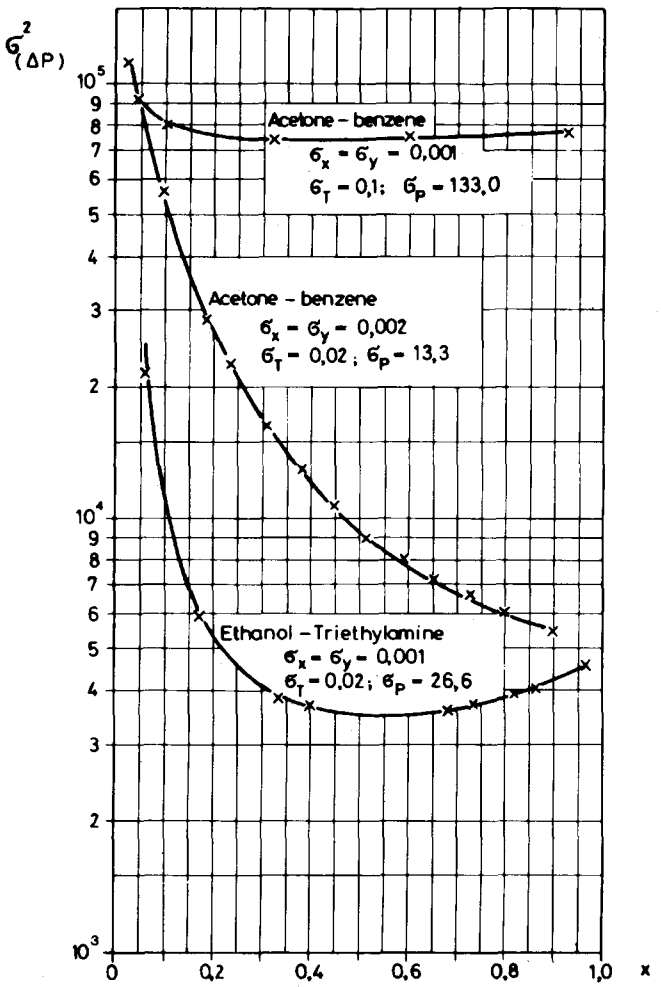

Figure 1. Resultant variances of ΔP for three binary VLE systems. Units: T [°K], P [Pa].

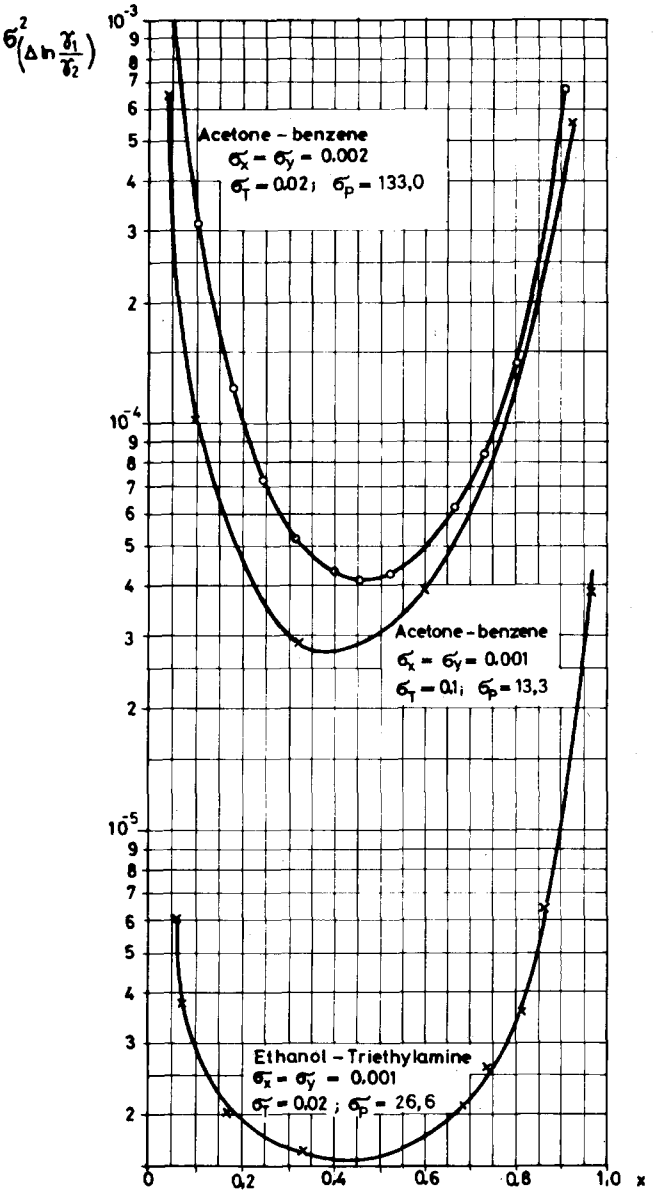

Figure 2. Resultant variances of $\Delta \ln \gamma_1/\gamma_2$ for three binary VLE systems, Units: T [°K], P [Pa].

TABLE 4. SIMULATION RESULTS FOR THREE DIFFERENT BINARIES. THE MEASUREMENTS ARE SUBJECT TO RANDOM ERRORS. $N = 50$

Data set / Criterion	$\widehat{MSE} \times 10^{-3}$ §					
	Acetone-Benzene			Ethanol-Triethylamine		Ethanol-Isooctane
	1	2	3	4	5	6
A	31.8	15.1	93.6	97.5	64.7	185.0
B	32.3	8.64	40.1	47.8	19.7	39.2
C	29.4	120.0	788.0	6650000.	9910000.0	674.0
D	5.88	22.8	469.0	2930000.	135.0	61.9
E	12.8	32.4	208.0	348.0	248.0	577.0
F	12.8	26.2	162.0	176.0	125.0	52.6
G	4.94	7.86	42.3	56.7	21.3	30.0
H	4.92	4.82	23.0	43.5	18.3	24.2
$\lambda^*_{12} - \lambda^*_{11}$ (J/mole)	2493.2	2493.2	2493.2	−6070.9	−6070.9	9633.4
$\lambda^*_{12} - \lambda^*_{22}$ (J/mole)	−818.5	−818.5	−818.5	2679.5	2679.5	1089.8
P (kPa)	101.3	101.3	101.3	101.3	101.3	—
T (°K)	—	—	—	—	—	349.15
n	13	13	13	5	13	10
$\sigma_x = \sigma_y$	0.001	0.002	0.005	0.001	0.001	0.005
σ_t (°K)	0.1	0.02	0.02	0.02	0.02	0.1
σ_P (Pa)	133.0	13.3	13.3	26.6	26.6	70.0

§ Mean square error of the estimated parameters.

piece of information. For each simulation a correct multiresponse maximum likelihood criterion (criterion H of Table 4) which takes into account the existence of two dependent variables will lead to the smallest MSE value. This means that the probability of estimating parameters strongly deviating from the true values is the smallest with this method.

Prospects for Detecting Systematic Errors from the Estimated True Values of the Variables

The parameters obtained by means of criterion 7 by Anderson et al. (1978) are identical to the parameters obtained by means of any other multiresponse maximum likelihood method. Yet, it yields some additional information, namely the estimated true values of the measured variables. In the case of small normally distributed errors and a perfect model, these estimated values give no new information when compared to other criteria (e.g., Eqs. 5 and 24). This is due to the fact that the variances calculated from the estimated residuals will be asymptotically identical to the assumed variances of the measurement errors as the number of experimental points increases.

A simulation study was carried out with data containing systematic measurement errors or errors due to model imperfections. The systematic measurement errors were superimposed on the measurement data $\underline{Z}_i$ (step 3 of the simulation). We have investigated the influence of systematic errors $(\Delta\tilde{y},\Delta\tilde{T},\Delta\tilde{P})$ on y, T, and P given by the following expressions,

$$\begin{aligned} \Delta\tilde{y}_i &= \xi_y(x_i - y_i) \\ \Delta\tilde{T}_i &= \xi_T \\ \Delta\tilde{P}_i &= \xi_P \end{aligned} \tag{38}$$

The relative systematic error on y was applied due to the fact that errors in the vapor phase composition are caused by condensation. It is therefore reasonable to assume that these errors depend on the difference between the compositions of the two phases.

The systematic model errors were introduced by the use of two different models for the generation respectively the reduction step of the simulation. We used the Wilson equation for the generation step while the UNIQUAC equation was used for the data reduction.

The Anderson criterion was used throughout the studies on systematic errors to enable the detection of possible valuable information gained from the estimation of true values of the variables.

A test similar to Abbe's test (Linnik, 1961) was used for the detection of systematic trends in the resulting residuals of the variables. The test quantity R_j is defined by

$$R_j = \frac{A_j^2}{S_j^2} \tag{39}$$

where

$$A_j^2 = \frac{1}{2(n-1)} \sum_{i=1}^{n-1} ((\Delta\hat{Z}_j)_{i+1} - (\Delta\hat{Z}_j)_i)^2$$

$$S_j^2 = \frac{1}{n-1} \sum_{i=1}^{n} ((\Delta\hat{Z}_j)_i - (\overline{\Delta Z}_j))^2$$

The test has two critical regions. If R_j exceeds unity significantly, the oscillations of the residuals are faster than the expected oscillations of random errors. The upper critical value of R_j is 1.42 for $n = 13$ on a 95% confidence level. If the residuals exhibit a systematic trend, either monotonous or slowly oscillating, the successive differences of A_j^2 will be relatively small. In this case R_j will decrease to values significantly smaller than unity. The lower critical value of R_j is 0.58 for $n = 13$ on a 95% confidence level.

In order to check the residuals of the variables for systematic off-sets from zero mean value a t-test was carried out. The test quantity is

$$\rho_j = \frac{\Delta\bar{\hat{Z}}_j}{S_j/\sqrt{n}}$$

where

$$\Delta\bar{\hat{Z}}_j = \frac{1}{n} \sum_{i=1}^{n} (\Delta\hat{Z}_j)_i$$

The critical region is given by

$$\rho_j < t(n-1)_{\alpha/2} \vee \rho_j > t(n-1)_{1-\alpha/2}$$

on an α confidence level.

Table 5 shows the results from simulations with an acetone-benzene system at 1.013 bar applying random as well as systematic measurement errors. The UNIQUAC equation was used for both data generation and data reduction. No model errors were thereby introduced. Abbe's test for randomness rejects systems with systematic errors which are significantly larger than the random errors. However, the estimated true values of the variables give little or no information on the source of the systematic errors. This is due to the fact that the maximum likelihood principle is based on the assumption of random errors. The systematic errors are distributed to all the estimated variables according to the assumed variances of the random errors. In some cases (systematic errors on y) the t-tests for systematic off-sets rejected systems which the Abbe's test accepted. This demonstrates that these two tests supplement each

TABLE 5. SIMULATION WITH AN ACETONE-BENZENE SYSTEM AT $1.013 \cdot 10^5$ PA. THE MEASUREMENTS ARE SUBJECT TO SYSTEMATIC AS WELL AS RANDOM ERRORS. $N = 50$. $n = 13$.

Random Errors	Systematic Errors	Abbe's Test: Rejections of Randomness Hypothesis, 90% Confidence Level x	T	y	P	Any
$\sigma_x = \sigma_y = 0.001$	$\zeta_T = 0.05$°K	34%	36%	12%	34%	58%
$\sigma_T = 0.01$°K	$\zeta_P = 67.0$ Pa	10%	4%	8%	4%	20%
$\sigma_P = 13.3$ Pa	$\zeta_y = 0.025$*	28%	16%	36%	16%	58%
$\sigma_x = \sigma_y = 0.0001$	$\zeta_T = 0.05$°K	100%	100%	100%	100%	100%
$\sigma_T = 0.001$°K	$\zeta_P = 67.0$ Pa	100%	100%	98%	100%	100%
$\sigma_P = 1.33$ Pa	$\zeta_y = 0.025$*	100%	100%	100%	100%	100%

* t-tests on the residuals of x and y showed a systematic offset from zero mean value in all cases. Confidence level is 90%.

other. The maximum information which can be extracted from the estimated true values of the variables is an indication of whether the data contains systematic errors. Furthermore, this is only true if the model used in the data reduction is inherently able to represent the true values of the variables. If this is not the case, further confusion will be introduced by model errors. Estimation of the true values of the variables gives therefore at best the same information on systematic measurement errors as a point to point consistency test of the type recommended by a Van Ness et al. (1973). Models for G^E introduce model errors for most systems. It is often impossible to extract information on systematic measurement errors from the estimated values of the variables calculated by means of such models.

Table 6 shows the mean square errors (MSE) for simulations with an acetone-benzene system. It may be seen that the estimation of true values of the variables and of the parameters is affected by the presence of systematic measurement errors. The MSE values reflect the biases and the variances of the estimators. The estimation method is not able to arrive at the true values of the variables or of the parameters from data subject to systematic measurement errors. This indicates the limited possibilities of methods which are based on the assumption of random errors.

Table 7 shows the results from simulations where the data reduction is carried out with one model (UNIQUAC) while the data generation is carried out by means of another (Wilson). For the acetone-benzene system with realistic random errors the randomness hypothesis was accepted in most cases. This means that the model used for data reduction represents the generated data introducing no severe systematic errors. When very small random errors are applied, even modest model errors are detected by the modified Abbe's test. For the hexane-methanol system the randomness hypothesis is rejected in all cases. The UNIQUAC equation is not able to represent the generated experimental data without significant systematic errors.

Abbe's test gives information about randomness of the residuals. However, for the comparison of the ability of different models to represent a given data set the value of the objective function is sufficient since it directly reflects the magnitude of the residuals. The two types of model tests are to some extent overlapping since it is most likely that severe model errors will cause systematic errors in the estimated variables.

The two simulation studies on Eq. 7 lead to the conclusion that the estimation of true values of the variables gives no essential information which could not be obtained from other well known tests of the data or models. However, a combination of a perfect fitting equation with Eq. 7 could be used for consistency tests. It is

TABLE 6. SIMULATIONS WITH AN ACETONE-BENZENE SYSTEM AT 1.013×10^5 PA. THE MEASUREMENTS ARE SUBJECT TO SYSTEMATIC AS WELL AS RANDOM ERRORS. $N = 50$, $n = 13$.

Random Errors					$\widehat{\text{MSE}}$					
σ_x	σ_y	σ_T °K	σ_P Pa	Systematic Errors	x ($\times 10^6$)	T ($\times 10^3$) °K^2	y ($\times 10^6$)	P ($\times 10^{-4}$) Pa^2	a_{12} (J/mol)^2	a_{21} (J/mol)^2
0.001	0.001	0.01	13.3	None	2.90	1.05	2.65	0.217	10.9	32.3
				$\zeta_T = 0.05$°K	9.86	30.5	5.58	0.226	29.2	55.7
				$\zeta_P = 67.0$ Pa	4.49	1.10	3.16	5.95	13.4	34.2
				$\zeta_y = 0.025$	37.5	1.16	50.7	0.220	54.9	208.0
0.0001	0.0001	0.001	1.33	None	0.039	0.012	0.031	0.0021	0.137	0.376
				$\zeta_T = 0.05$°K	7.41	30.0	2.82	0.0112	17.6	21.3
				$\zeta_P = 67.0$ Pa	1.46	0.074	0.59	5.70	2.94	3.19
				$\zeta_y = 0.025$	35.0	0.129	49.1	0.0057	44.0	181.0

TABLE 7. SIMULATIONS WITH 2 DIFFERENT BINARIES AT $1.013 \cdot 10^5$ PA. THE MEASUREMENTS ARE SUBJECT TO RANDOM ERRORS AND THEY WERE GENERATED BY MEANS OF THE WILSON EQUATION. THE UNIQUAC EQUATION WAS USED FOR DATA REDUCTION. $N = 50$, $n = 13$.

Systems	Random Errors*	Abbe's Test: Rejections of Randomness Hypothesis, 90% Confidence Level x	T	y	P	Any
Acetone-Benzene	$\sigma_x = \sigma_y = 0.001$, $\sigma_T = 0.01$, $\sigma_P = 13.3$	6%	14%	10%	14%	28%
	$\sigma_x = \sigma_y = 0.0001$, $\sigma_T = 0.001$, $\sigma_P = 1.33$	100%	100%	66%	100%	100%
Hexane-Methanol	$\sigma_x = \sigma_y = 0.001$, $\sigma_T = 0.01$, $\sigma_P = 13.3$	100%	100%	100%	100%	100%
	$\sigma_x = \sigma_y = 0.0001$, $\sigma_T = 0.001$, $\sigma_P = 1.33$	100%	100%	100%	100%	100%

* Temperature in °K, pressure in Pa.

therefore possible to perform consistency tests and estimations of model parameters using the same Eq. 7.

ACKNOWLEDGMENT

Part of this work was done during Sándor Kemény's stay at Instituttet for Kemiteknik, Danmarks Tekniske Højskole. He is very grateful for having had the opportunity to use the facilities of this university and for the kindness of the staff. The authors appreciate the many inspiring discussions with Peter Rasmussen.

Appendix 1

Eq. 10 by Box (1970) can be derived using results of Clutton-Brock (1967) for one dependent variable.

The conditional normal probability density function (pdf) for the j'th dependent variable (Y_j) can be written as

$$\phi_s(Y_j|\underline{X}^*) = \frac{1}{\sqrt{2\pi}\sigma_{Y_j}} \exp\left(-\tfrac{1}{2}\frac{(Y_j - f_j(\underline{X}^*))^2}{\sigma^2_{Y_j}}\right) \quad \text{(A1)}$$

where the true value Y_j^* of Y_j is given by

$$\underline{Y}_j^* = f_j(\underline{X}^*)$$

and $\underline{X}^*$ is the true values of the independent variables. Let us define the random variable ΔY_j by,

$$\Delta Y_j \equiv Y_j - f_j(\underline{X}) = (Y_j - Y_j^*) + f_j(\underline{X}^*) - f_j(\underline{X})$$
$$\simeq (Y_j - Y_j^*) + \sum_{k=1}^{s} (X_k^* - X_k) f_j'(X_k^*) \quad \text{(A2)}$$

For our purpose it can be supposed that

$$\mathrm{Cov}(X_j, X_k) = \mathrm{Cov}(X_j, Y_k) = \mathrm{Cov}(Y_j, Y_k) = 0, \quad j \neq k \quad \text{(A3)}$$

ΔY_j is a sum of two random variables following Gaussian distributions with zero expected value. ΔY_j is therefore itself a Gaussian distributed random variable with zero expected value. The variance of ΔY_j is given by,

$$\mathrm{Var}[(Y_j - f_j(\underline{X}))|\underline{X}^*] \simeq \mathrm{Var}(Y_j - Y_j^*)$$
$$- \sum_{k=1}^{s} (f_j'(X_k))^2 \mathrm{Var}(X_k - X_k^*) = \sigma^2_{Y_j} + \sum_{k=1}^{s} (f_j'(X_k))^2 \sigma^2_{X_k}$$
$$= \sigma^2_{(\Delta Y_j)} \quad \text{(A4)}$$

The pdf is,

$$\phi_s((Y_j - f_j(\underline{X}))|\underline{X}^*) \simeq \frac{1}{\sqrt{2\pi}\sigma_{(\Delta Y_j)}} \exp\left(-\tfrac{1}{2}\left(\frac{Y_j - f_j(\underline{X})}{\sigma_{(\Delta Y_j)}}\right)^2\right) \quad \text{(A5)}$$

In Eq. A4 $f_j'(X_k)$ is used instead of the strict $f_j'(X_k^*)$. For fixed values of $\underline{X}$ Eq. A5 can be considered to be a conditional pdf;

$$\phi_s(Y_j|\underline{X}) = \phi_s((Y_j - f_j(\underline{X}))|\underline{X}^*) \quad \text{(A6)}$$

The density functions $\phi_s(Y_j|\underline{X}^*)$ are independent of each other for different Y_j variables while the $\phi_s(Y_j|\underline{X})$ density functions depend on each other since the random variables $\underline{X}$ are contained in each conditional pdf. For instance, the pdf for two dependent variables is given by,

$$\phi_s(Y_{1i}, Y_{2i}|\underline{X}_i) = \frac{1}{2\pi\sqrt{\sigma^2_{(\Delta Y_1)i}\sigma^2_{(\Delta Y_2)i} - (\sigma_{(\Delta Y_1, \Delta Y_2)i})^2}}$$
$$\exp\left\{\left[-\frac{\sigma^2_{(\Delta Y_1)i}\sigma^2_{(\Delta Y_2)i}}{2\sigma^2_{(\Delta Y_1)i}\sigma^2_{(\Delta Y_2)i} - (\sigma_{(\Delta Y_1, \Delta Y_2)i})^2}\right]\right.$$
$$\left[\left(\frac{Y_{1i} - f_1(\underline{X}_i)}{\sigma_{(\Delta Y_1)i}}\right)^2 - 2\sigma_{(\Delta Y_1, \Delta Y_2)i}\frac{(Y_{1i} - f_1(\underline{X}_i))(Y_{2i} - f_2(\underline{X}_i))}{\sigma^2_{(\Delta Y_1)i}\sigma^2_{(\Delta Y_2)i}}\right.$$
$$\left.\left. + \left(\frac{Y_{2i} - f_2(\underline{X}_i)}{\sigma_{(\Delta Y_2)i}}\right)^2\right]\right\} \quad \text{(A7)}$$

The likelihood function is the product of the density functions. Its maximization is in practice equivalent to minimization of the function

$$\sum_{i=1}^{n} \frac{\sigma^2_{(\Delta Y_1)i}\sigma^2_{(\Delta Y_2)i}}{\sigma^2_{(\Delta Y_1)i}\sigma^2_{(\Delta Y_2)i} - (\sigma_{(\Delta Y_1, \Delta Y_2)i})^2}\left[\left(\frac{Y_{1i} - f_1(\underline{X}_i, \hat{\underline{\theta}})}{\sigma_{(\Delta Y_1)i}}\right)^2\right.$$
$$- 2\sigma_{(\Delta Y_1, \Delta Y_2)i}\frac{(Y_{1i} - f_1(\underline{X}_i, \hat{\underline{\theta}}))(Y_{2i} - f_2(\underline{X}_i, \hat{\underline{\theta}}))}{\sigma^2_{(\Delta Y_1)i}\sigma^2_{(\Delta Y_2)i}}$$
$$\left. + \left(\frac{Y_{2i} - f_2(\underline{X}_i, \hat{\underline{\theta}}f)}{\sigma_{(\Delta Y_2)i}}\right)^2\right] \quad \text{(A8)}$$

where we have introduced the vector $\hat{\underline{\theta}}$ of the model parameters in the functions f_1 and f_2.

If reduction of F_j transforms is desirable instead of reduction of Y_j values themselves ($\Delta\hat{F}_j$) must be introduced into into Eq. A8 instead of ($\Delta\hat{Y}_j$). It follows from the error propagation law (Taylor series truncated after the first order term) that

$$\sigma^2_{(\Delta F_1)i} = \left[\frac{\partial(\Delta F_1)}{\partial(\Delta Y_1)}\right]^2_i \sigma^2_{(\Delta Y_1)i} \quad \text{(A9)}$$

$$\sigma^2_{(\Delta F_2)i} = \left[\frac{\partial(\Delta F_2)}{\partial(\Delta Y_2)}\right]^2_i \sigma^2_{(\Delta Y_2)i} \quad \text{(A10)}$$

$$\sigma_{(\Delta F_1, \Delta F_2)i} = \left[\frac{\partial(\Delta F_1)}{\partial(\Delta Y_1)}\right]_i \left[\frac{\partial(\Delta F_2)}{\partial(\Delta Y_2)}\right]_i \sigma_{(\Delta Y_1, \Delta Y_2)i} \quad \text{(A11)}$$

Criterion A8 is now identical to Eq. 12. The multivariable Gaussian pdf for the mutual distribution of $\underline{Y}$ and $\underline{X}$ is,

$$\phi_s(Y, \underline{X}|\underline{X}^*) = \frac{1}{(2\pi)^{(r+s)/2}\prod_{j=1}^{r}\sigma_{Y_j}\prod_{k=1}^{s}\sigma_{X_k}}$$
$$\exp\left\{-\tfrac{1}{2}\left[\sum_{k=1}^{s}\left(\frac{X_k - X_k^*}{\sigma_{X_k}}\right)^2 + \sum_{j=1}^{r}\left(\frac{Y_j - f_j(\underline{X}^*)}{\sigma_{Y_j}}\right)^2\right]\right\} \quad \text{(A12)}$$

This density function leads directly to Eq. 7 by Anderson et al. (1978). The density function (Eq. A5) leads to Eq. 12. Since Eqs. A5 and A12 are identical within the order of magnitude of the errors caused by neglecting higher terms of the Taylor series the two criteria 7 and 12 are identical.

APPENDIX 2

Computer Programs

The random error generation is based on a program used in a work by Shakina and Kemény (1978). The program generates N data sets with n observations of x, T, y and P in each. These data sets are written into a file. Standard IBM SSP subroutines were used for the generation of random numbers. Two alternative prorams were used for the data reduction part, namely a program based on criterion 24 combined with the Hooke and Jeeves search method (Kemény, 1980) and a program based on criterion 7 combined with the Levenberg-Marquardt method for sums of squares (Skjold-Jørgensen, 1980). The computer programs used for the simulations are available from the authors.

NOTATION

A = see Eq. 39
a_{12}, a_{21} = interaction parameters of the UNIQUAC model
$\underline{\mathrm{Cov}}$ = covariance matrix
$\underline{\underline{C}}$ = see Eq. 15
D = see Eq. 25
E = expected value
f = some function
F = transform of the dependent variable
g = some function
n = number of data points
N = number of cycles
p = number of parameters
P = pressure

r = number of dependent variables
R = test quantity for Abbe's test, Eq. 39
R_G = gas constant
s = number of independent variables
S = see Eq. 39
T = temperature
$\mathrm{Var}(i)$ = variance (identical to σ_i^2)
w = see Eq. 15
x = liquid phase mole fraction
X = independent variable
y = vapor phase mole fraction
Y = dependent variable
Z = measured variable

Greek letters

γ = activity coefficient
δ = error in X
Δ = difference between an experimental value and the corresponding true value
ϵ = error in Y
$\underline{\theta}$ = parameter vector
λ = interaction parameter of the Wilson equation
ξ = parameter for systematic measurement errors
ρ = test quantity for the t-test
σ^2 = variance
ϕ = fugacity coefficient
ϕ_s = probability density function

Superscripts

jk = element j,k
T = transposed
o = reference state
-1 = inverse
$\tilde{}$ = systematic error
$\hat{}$ = estimated
∞ = infinite dilution
$*$ = true value

LITERATURE CITED

Anderson, T. F., D. S. Abrams, and E. A. Grens, "Evaluation of Parameters for Nonlinear Thermodynamic Models," *AIChE J.*, **24**, 20 (1978).

Bard, Y., "Nonlinear Parameter Estimation," Academic Press, New York and London (1974).

Bard, Y. and L. Lapidus, "Kinetic Analysis by Digital Parameter Estimation," *Catalysis Reviews*, 2(1), 67 (1968).

Box, G. E. P. and N. R. Draper, "Bayesian Estimation of Common Parameters from Several Responses," *Biometrika*, **52**, 355 (1965).

Box, M. J., "Improved Parameter Estimation," *Technometrics*, **12**, 219 (1970).

Clutton-Brock, M., "Likelihood Distributions for Estimating Functions when Both Variables are Subject to Error," *Technometrics*, **9**, 261 (1967).

Fabries, J. and H. Renon, "Methods of Evaluation and Reduction of Vapor-Liquid Equilibrium Data of Binary Mixtures," *AIChE J.*, **21**, 735 (1975).

Gmehling, J. and U. Onken, "Vapor-Liquid Equilibrium Data Collections," Dechema Chemistry Data Series I/2a (1977).

Hunter, W. G., "Estimation of Unknown Constants from Multiresponse Data," *Ind. Eng. Chem. Fund.*, **6**, 461 (1967).

Hooke, R. and T. A. Jeeves, "Direct Search Solution of Numerical and Statistical Problems," *J. Assoc. Comput. Mach.*, **8**, 212 (1961).

Kemény, S., "Reduction of Binary VLE Data," In Hungarian, Thesis (1980).

Kemény, S. and J. Manczinger, "Treatment of Binary Vapour-Liquid Equilibrium Data," *Chem. Eng. Sci.*, **33**, 71 (1978).

Kowalik, J. and M. R. Osborne, "Methods for Unconstrained Optimization Problems," Elsevier (1968).

Linnik, J. W., "Die Metode der kleinsten Quadrate in moderner Darstellung," VEB Deutsche Verlag der Wissenschaften, Berlin (1961).

Péneloux, A., R. Deyrieux, E. Canals, and E. Neau, "The Maximum Likelihood Test and the Estimation of Experimental Inaccuracies: Application to Data Reduction for Vapor-Liquid Equilibrium," *J. de Chim. Phys.*, **73**, 706 (1976).

Péneloux, A., R. Deyrieux, and E. Neau, "Réduction des données experimentales et théorie de l'information," *J. de Chim. Phys.*, **72**, 1101 (1975,1).

Péneloux, A., R. Deyrieux, and E. Neau, "Réduction des données sur les équilibres liquide-vapeur binaires isothermes. Critères de précision et de cohérence. Analyse de l'information," *J. de Chim. Phys.*, **72**, 1107 (1975,2).

Schubert, E., "Vergleich verschiedener Metoden zuer Ermittlung der Parameter in der Wilson- und NRTL-Gleichung," *Chem.-Ing. Tech.*, **46**, 73 (1974).

Shakina, E. and S. Kemény, CHISA'78 Conference, Paper 3.9, "Extension of the Ridge Regression Method for VLE Data," Prague (1978).

Skjold-Jørgensen, S., "On the UNIFAC and UNIQUAC models," Ph.D. Thesis (1980).

Sutton, T. L. and J. F. MacGregor, "The Analysis and Design of Binary Vapour-Liquid Equilibrium Experiments, Part I: Parameter Estimation and Consistency Tests," *Can. J. Chem. Eng.*, **55**, 602 (1977).

Ulrichson, D. L. and F. D. Stevenson, "Effects of Experimental Errors on Thermodynamic Consistency and on Representation of Vapor-Liquid Equilibrium Data," *Ind. Eng. Chem. Fund.*, **11**, 287 (1972).

Van Ness, H. C., S. M. Byer, and R. E. Gibbs, "Vapor-Liquid Equilibrium, Part I: An Appraisal of Data-Reduction Methods," *AIChE J.*, **19**, 238 (1973).

Van Ness, H. C., F. Pedersen, and P. Rasmussen, "Vapor-Liquid Equilibrium, Part V: "Data Reduction by Maximum Likelihood," *AIChE J.*, **24**, 1055 (1978).

Manuscript received July 11, 1980; permission received February 2, and accepted February 6, 1981.