Nonlinear Least Squares for Model Screening

REIJI MEZAKI
Yale University, New Haven, Connecticut
and
JAMES R. KITTRELL
Chevron Research Company, Richmond, California

In a recent article (*1*) Lapidus and Peterson reported on the estimation of kinetic parameters of the Hougen-Watson rate models, which were postulated for the vapor phase catalytic dehydration of ethanol over an ion exchange resin by a cyclic nonlinear estimation procedure. They also attempted to discriminate among plausible rate models proposed in light of the values of the standard deviation of the residuals, and the signs of estimated parameters. From their analyses, however, it was not possible to discriminate clearly among three models; single, homogeneous, and dual site surface reaction rate controlling models.

It has been well accepted that a large number of plausible Hougen-Watson models can usually be postulated to describe a reaction system. As screening and discrimination of these rival models are troublesome tasks, some useful techniques have been developed especially for modeling in the field of applied kinetics (*5*). Nevertheless, it still requires considerable time and effort to properly perform the discrimination procedures on so many rival models. Thus it is sometimes quite desirable to take kinetic data to allow an initial screening of the inadequate models, as proposed by Yang and Hougen (*2*). Then more close attention can be directed toward the additional data required for a careful discrimination among the remaining more related rival models. Most of the discrimination techniques already developed are iterative in nature. Accordingly, these techniques may not be feasible to apply directly to the existing kinetic data, unless additional data can be supplied from the identical experimental system. Hence, for an existing reaction system some discrimination methods which are of noniterative nature should be recommended.

The purpose of this paper is to show that without gathering additional data, one can obtain a more conclusive result for the discrimination among three rival models of ethanol dehydration, adsorption controlling, single site, and dual site surface reaction rate controlling models. We will also point out some dangers that one may encounter when the discrimination of rival models is conducted simply on the information obtained from parameter estimation procedures.

RATE MODELS

Lapidus and Peterson (*1*) extracted the following four rival models from the study of Kabel and Johanson (*3, 4*).

(*a*) dual site surface reaction controlling model

$$r = \frac{kK_A\,[p_A{}^2 - (p_Bp_C/K_{eq})]}{[1 + K_Ap_A + K_Bp_B + K_Cp_C]^2}$$

(*b*) single site surface reaction controlling model

$$r = \frac{kK_A\,[p_A{}^2 - (p_Bp_C/K_{eq})]}{(1 + K_Ap_A + K_Bp_B + K_Cp_C)}$$

(*c*) adsorption controlling model

$$r = \frac{k\,[p_A - \sqrt{p_Bp_C/K_{eq}}]}{[1 + \sqrt{p_Bp_CK_A{}^2/K_{eq}} + K_Bp_B + K_Cp_C]}$$

(*d*) desorption controlling model

$$r = \frac{kK_{eq}K_B\,[p_A{}^2/p_C - p_B/K_{eq}]}{\left[1 + K_Ap_A + K_BK_{eq}\dfrac{p_A{}^2}{p_C} + K_Cp_C\right]}$$

In addition, Lapidus and Peterson examined a homogeneous reaction model (*e*) as a model for describing the system.

(*e*) $$r = k'\,[p_A{}^2 - p_Bp_C/K_{eq}]$$

According to their analysis, both adsorption and desorption controlling models are eliminated because of their large values of the standard deviation of the residuals, indicating the poorer fit of those models to experimental rate data. Furthermore, one of the parameter estimates for the adsorption model was negative. However further discrimination among the remaining three models was not possible with only the comparison of the sum of squares of residuals which can be obtained from nonlinear least squares estimation of parameters.

Screening of Inadequate Models

Kabel and Johanson (*4*) studied the vapor-phase dehydration of alcohol to ether over an ion exchange resin and reported ten data points at 120°C., one atmosphere total pressure and at a partial pressure of water equal to zero. These data are used for the analysis throughout this paper. The desorption rate controlling model can be conveniently simplified and linearized for initial rates as

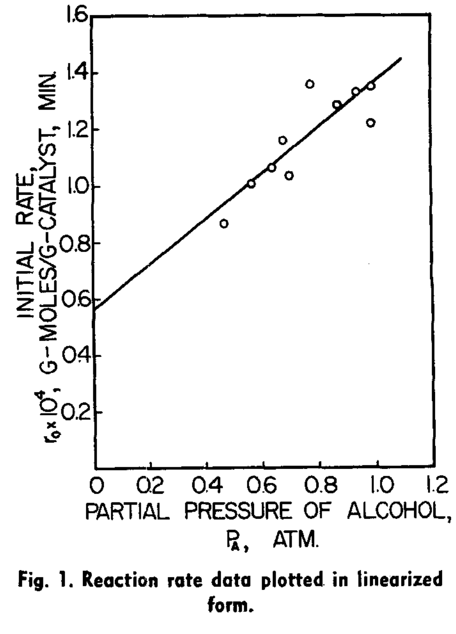

Fig. 1. Reaction rate data plotted in linearized form.

$$r_0 = k \tag{1}$$

This implies that the initial rates of the desorption controlling model are constant. As seen in Figure 1 this relationship was not fulfilled, suggesting that the desorption model should be discarded. The homogeneous rate model is thoroughly discussed by Kabel (6). Thus no attempt is made here to criticize the adequacy of the homogeneous model.

Under the condition of zero water partial pressure, the initial rate for the adsorption controlling model (c) is simplified to

$$r_0 = \frac{kp_A}{(1 + K_C p_C)} \tag{2}$$

Equation (2) suggests a plot of r_0 vs. p_A which should be linear with a zero intercept. Figure 1 shows the plot of r_0 vs. p_A. We analyzed these data and found, using all ten data points reported by Kabel and Johanson, that the 99.9% confidence interval of the intercept does not include zero. Note in Figure 1 that some curvature in the locus of points exists. However, the assumption that the data locus is straight and the test of the intercept is more sensitive than a test for the existence of curvature in the locus.

Discrimination among Surface Rate Controlling Models

The single site model shown by (b) may be linearized to

$$\frac{p_A^2}{r_0} = \frac{1}{kK_A} + \frac{p_A}{k} \tag{3}$$

if K_C is considered to be zero. Equation (3) shows that the plot of p_A^2/r_0 vs. p_A should be a straight line of a positive intercept. This relationship may also be obtained if the initial rate data were gathered for pure ethanol feed. The ten data points which were employed for the previous screening purpose were analyzed. Figure 2 presents the results. We estimated a value of -2.45×10^3 for the intercept with the 95% confidence limits of $\pm 1.46 \times 10^3$. This suggests that single site controlling model may be discarded because of the significantly negative values of the intercept. To make this result more conclusive, the parameter n of the generalized surface reaction rate controlling equations was determined through diagnostic parameters (7). For the dehydration of alcohol, the generalized surface rate controlling equation may be presented by

$$r = \frac{k'(p_A^2 - p_B p_C/K_{eq})}{(1 + K_A p_A + K_B p_B + K_C p_C)^n} \tag{4}$$

Introducing stoichiometric correlation for this reaction system and collecting terms, we obtain

$$\frac{\partial w}{\partial y} = \frac{[(I + K_A x_A + K_B x_B + K_C x_C) + (1/2\,K_B + 1/2\,K_C - K_A)x_A y]^n}{k'[(x_A^2(1-y)^2 - (x_B + 1/2\,x_A y)(x_C + 1/2\,x_A y)/K_{eq}]} \tag{5}$$

or

$$\frac{\partial w}{\partial y} = \frac{[\hat{C}_1 + \hat{C}_2 y]^n}{[x_A^2(1-y)^2 - (x_B + 1/2\,x_A y)(x_C + 1/2\,x_A y)/K_{eq}]} \tag{6}$$

where

$$\hat{C}_1 = b_0 + b_1 x_A + b_2 x_B + b_3 x_C \tag{7}$$

$$\hat{C}_2 = b x_A \tag{8}$$

Furthermore

$$b = 1/2\,K_B + 1/2\,K_C - K_A$$

$$b_0 = \frac{1}{k'}$$

$$b_1 = \frac{K_A}{k'}$$

$$b_2 = \frac{K_B}{k'}$$

$$b_3 = \frac{K_C}{k'}$$

When the reactor feed does not contain water, the observed value of C_2 is found to be:

$$C_2 = \frac{1}{nx_A^2 C_1^{n-1}} \left\{ x_A^2 \frac{\partial^2 w}{\partial y^2}\bigg|_{y=0} - C_1^n \left[2x_A^2 + \frac{1}{2}\frac{x_A x_B}{K_{eq}} \right] \right\} \tag{9}$$

Substituting Equation (8) into (9) we obtain

$$bx_A = \frac{1}{nx_A^2 C_1^{n-1}} \left\{ x_A^2 \frac{\partial^2 w}{\partial y^2}\bigg|_{y=0} - C_1^n \left[2x_A^2 + \frac{1}{2}\frac{x_A x_B}{K_{eq}} \right] \right\} \tag{10}$$

Now C_1 is estimated from the experimental data through

$$C_1 = \left\{ x_A^2 \frac{\partial w}{\partial y}\bigg|_{y=0} \right\}^{1/n} \tag{11}$$

The first and second derivatives of Equation (10) and (11) were estimated by analytically differentiating a second-order polynomial which was fitted to the conversion space time data. Employing experimental conditions for ten data sets reported (4) and these estimated derivatives, by nonlinear least squares techniques. The point estimates of n ranged from 1.89 to 2.41 for all ten sets of the data, indicating that the reaction favorably involves two adjacent sites. One can also approximately estimate the value of n, by using only the first derivative. The conversion space time data may be represented by a straight line at the neighborhood of zero conversion. In this case the second derivative of Equation (10) vanishes and the equation is simplified to

$$\ln\left(\frac{x_A^2}{r_0}\right) = n \ln (nb) + n \ln \left(\frac{x_A^2}{2x_A + \dfrac{1}{2K_{eq}} x_B} \right) \tag{12}$$

The value of n may be estimated by linear regression. The point estimate of n was 1.73 and the 95% confidence limits of the estimate were $\pm$ 0.44. This suggests that the dual site model is still superior to the single site models, although not precisely.

Finally as a supplementary test all parameters in both single and dual surface reaction models were estimated by nonlinear least squares. The data sets employed were nine out of the ten reported. We obtained the reaction rates by fitting the conversion data by

$$\ln(1-y) = a_1 w + a_2 w^2 + a_3 w^3 \tag{13}$$

This expression, then, was differentiated to obtain the differential reaction rates. The applicability of Equation (13) to the conversion space time data is discussed thoroughly elsewhere (9).

This numerical procedure of differentiation may introduce some errors. Accordingly the nonlinear estimates of parameters thus obtained may be considered to be questionable. However this paper does not attempt to discriminate between two surface rate controlling models with the results of nonlinear estimation. In order to eliminate the possible errors introduced by the numerical differentiation it would be best to use integral forms of rate equations. Total data employed were 37 points. For the purpose of references, the results of nonlinear estimation were shown in Table 1.

When fitting a large number of models by nonlinear least squares, it is necessary to examine in considerable detail, the sum of squares surface which is being minimized. Several important characteristics of sums of squares surfaces of typical Hougen-Watson models were presented elsewhere (8): they are the signs of converged parameter estimates, the joint confidence region of these estimates, the goodness of fits, and the nature of the sums of squares surfaces. One or two computer runs may not be able to supply the full information on these characteristics, except for an extremely simple case in which a system has only one or two parameters.

TABLE 1. PARAMETER ESTIMATES THROUGH NONLINEAR ESTIMATION TECHNIQUES

Parameter		Single site model	Dual site model
k	g.-moles/(g.-catalyst (atm.) (min.)	1.12×10^{-4}	7.06×10^{-4}
K_A	atm.$^{-1}$	-5.26	2.91
K_B	atm.$^{-1}$	-12.8	6.34
K_C	atm.$^{-1}$	0	0
Sum of squares of residual rates $\times 10^9$		6.4	2.3

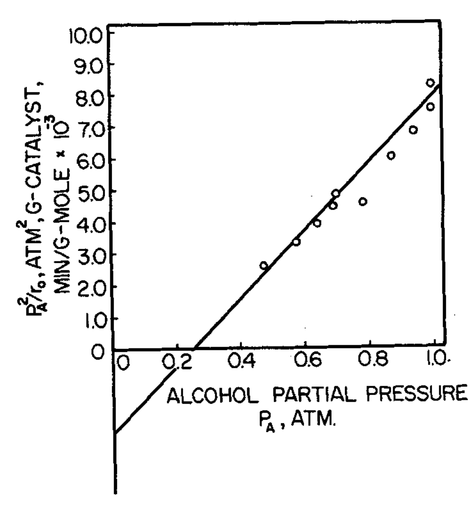

Fig. 2. Linear plot of single site controlling model.

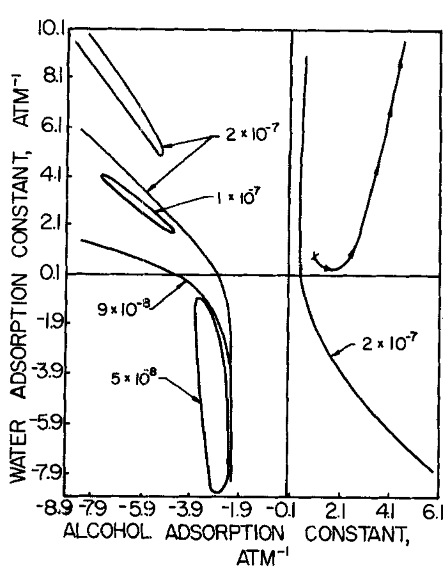

Fig. 3. Contours of sums of squares surface of single site model.

For the single site surface rate controlling model, a two dimensional section of the sums of squares surface (at $k = 4.8 \times 10^{-5}$ g.-moles/(min.) (atm.) (g.-cat.) and $K_C = 0$) was presented in Figure 3. The arrows shown in the figure indicate the path taken by subsequent iterations of a nonlinear least squares calculation. This clearly shows that no converged estimates were obtained with the initial estimates of parameters located in the first quadrant. The second quadrant possessed a relatively high minimum and the third quadrant contains an absolute minimum which gives the parameter estimates in Table 1. For any converged estimates of parameters, it would be advisable to investigate whether the minimum of the sum of squares of residuals was obtained. Due to the considerable expenditure of effort required for exploring the sum of squares surfaces, the use of nonlinear least squares for screening is not highly recommended. Some other method such as a linear plot and the use of a diagnostic parameter frequently yields more conclusive result for screening postulated models. For the system discussed here a linear plot could more quickly indicate that the dual site model is more adequate than the single site model.

NOTATION

a_i = constants
b, b_i = estimated value of a collection of adsorption and rate constants
C_1, C_2 = observed values of the diagnostic parameter
$\hat{C}_1, \hat{C}_2$ = estimated values of the diagnostic parameter C_1 and C_2
k = forward rate constant, g.-moles/(g.-cat.) (min.) (atm.)
k' = rate constant, g.-moles/(g.-cat.) (min.) (atm.^{2})
K_A, K_B, K_C = adsorption equilibrium constant of alcohol, water, ether, atm.$^{-1}$
K_{eq} = thermodynamic equilibrium constant
n = exponent of denominator of generalized rate expression
p_A, p_B, p_C = partial pressure of alcohol, water, ether, atm.
r, r_0 = rate and initial rate of reaction, g.-moles/(g.-cat.) (min.)
w = space time, (min.) (g.-cat.)/g.-mole
x_A, x_B, x_C = initial pressure of alcohol, water, ether, atm.
y = conversion of alcohol

LITERATURE CITED

1. Lapidus, Leon, and T. I. Peterson, *AIChE J.*, **11**, 891 (1965).
2. Yang, K. H., and O. A. Hougen, *Chem. Eng. Prog.*, **46**, No. 3, 146 (1950).
3. Kabel, R. L., Ph.D. thesis, Univer. Washington, Seattle (1961).
4. Kabel, R. L., and L. N. Johanson, *AIChE. J.*, **8**, 621 (1962).
5. Kittrell, J. R., Reiji Mezaki, and C. C. Watson, *Brit. Chem. Eng.* **11**, No. 1, 15 (1966).
6. Kabel, R. L., paper to appear in the *AIChE J.*
7. Kittrell, J. R., W. G. Hunter, and Reiji Mezaki, *ibid.*, **12**, 1014 (1966).
8. Mazaki, Reiji, and J. R. Kittrell, *Ind. Eng. Chem.*, **59**, No. 5, 63 (1967).
9. Kittrell, J. R., and Reiji Mezaki, *Brit. Chem. Eng.* **11**, 1538 (1966).

Dimensional Considerations in Viscoelastic Flows

JOHN C. SLATTERY

Northwestern University, Evanston, Illinois

I would like to comment from the view of the Noll simple fluid on three topics which have been discussed recently: the definition of characteristic time for a fluid, the significance of the Weissenberg and Deborah numbers, and the solidlike behavior of viscoelastic fluids during rapid accelerations. Before examining these ideas, let me briefly review the concept of the Noll simple fluid. (*Viscoelastic* is used here in the sense that the materials obey neither of the classical linear relations, Newton's law of viscosity and Hooke's law of elasticity. Subclasses of such materials are fluids which show a finite relaxation time and fluids which exhibit normal stresses in viscometric flows (*1*). The term *viscoelastic* is commonly used in the literature in referring to these subclasses.)

THE NOLL SIMPLE FLUID

Let $\boldsymbol{\xi}$ be the place at time $t - s$ $(0 \leq s < \infty)$ of that material particle which at time t occupies the place $\mathbf{x}$

$$\boldsymbol{\xi} = \boldsymbol{\chi}_{(t)}\,(\mathbf{x}, t - s) \tag{1}$$

We call $\boldsymbol{\chi}_{(t)}$ the relative deformation function. By Equation (1) we may describe motion which took place in the material at all times $t - s$ prior to the time t. The gradient with respect to $\mathbf{x}$ of the relative deformation function is called the *relative deformation gradient*.

$$\mathbf{F}_{(t)}\,(t - s) = \nabla\, \boldsymbol{\chi}_{(t)}\,(\mathbf{x}, t - s) \tag{2}$$

The relative right Cauchy-Green strain tensor is defined as

$$\mathbf{C}_{(t)}\,(t - s) = \mathbf{F}_{(t)}\,(t - s)^T \cdot \mathbf{F}_{(t)}\,(t - s) \tag{3}$$

where $\mathbf{F}_{(t)}(t - s)^T$ represents the transpose of $\mathbf{F}_{(t)}(t - s)$.

Noll defines an incompressible simple fluid (*2*) as one for which the stress $\mathbf{T}$ at the position $\mathbf{x}$ and time t is specified within an indeterminate pressure p by the history of the relative right Cauchy-Grees strain tensor for the material which is within an arbitrarily small neighborhood of $\mathbf{x}$ at time t

$$\mathbf{T} + p\,\mathbf{I} = \frac{\mu_o}{s_o} \overset{\infty}{\underset{\sigma = o}{\mathbf{H}^*}} \left(\mathbf{C}_{(t)}\,(t - s_o\sigma)\right) \tag{4}$$

Here we follow Truesdell's discussion of the dimensional indifference of the definition of a simple material (*2, 3*). The quantity $\overset{\infty}{\underset{\sigma = o}{\mathbf{H}^*}}$ is a dimensionally invariant tensor valued functional. (By tensor valued functional we mean an operator which maps tensor valued functions into tensors.) The constants μ_o and s_o are, respectively, a characteristic viscosity and characteristic time or natural time lapse of the fluid.

DEFINITION OF CHARACTERISTIC TIME OF FLUID

Truesdell (*3*) proposes a definition for the characteristic time s_o which is based upon the three material functions which describe the behavior of incompressible Noll simple fluids in viscometric flows (such as axial flow in a tube or Couette flow). Two of these functions yield normal stress differences for each rate of shear; the third, the shear viscosity function, assigns to each rate of shear what is often referred to as the *apparent viscosity*. His definition is constructed so as to yield a zero characteristic time in the event that the fluid shows no normal stress differences in viscometric flows. Bird (*4, 5*) introduces a characteristic time defined in terms of the shear viscosity function alone. With respect to data correlations, the latter type of definition has a current advantage, since the two material functions which describe normal stress differences are more difficult to measure than the viscosity function. Astarita (*6*) raises the following questions based upon the differences in these definitions.