Analysis of Heterogeneous Catalytic Reactions by Nonlinear Estimation

LEON LAPIDUS and T. I. PETERSON

International Business Machines Corporation, New York, New York

Integral conversion catalytic data have been studied by nonlinear estimation to determine whether the Langmuir-Hinshelwood heterogeneous catalytic models are more valid than the simpler homogeneous noncatalytic forms, and whether it is possible to discriminate among models. The experimental system chosen for this study was the vapor phase dehydration of ethanol. Data for a similar hypothetical catalytic system were also generated and analyzed to highlight the effects arising in the ethanol dehydration case.

Nonlinear estimation is shown to be a valuable adjunct in the study of reaction kinetics and mechanism. For the hypothetical system, the correct model was found to provide the best fit to the data. Linear regression, coupled with the usual criterion that models are acceptable if all parameter estimates are positive, did not provide this discrimination. For the experimental system of ethanol dehydration, as corroborated by the hypothetical system, such problems as degeneracy of the heterogeneous rate equations and comparatively large errors in measurement are shown to be obscuring factors in determining the best model. When these effects are involved, it is doubtful that the Langmuir-Hinshelwood equations are more warranted than the simpler homogeneous forms.

The Hinshelwood model based upon the Langmuir theory of adsorption has provided a pragmatic approach to the correlation of experimental rate data for heterogeneous catalytic reactions. Hougen and Watson (5) pioneered in this approach and showed the significant features of the method. However, there exists in the literature, as illustrated by Weller (*11*) and Boudart (*2*), considerable discussion of the need for analyzing data with the resulting relations. It has been proposed that many heterogeneous reactions, in particular, gas-solid catalytic reactions, can be represented by the simpler homogeneous (noncatalytic) rate equations instead of the more complex heterogeneous (catalytic) forms.

In the analysis of data, one or more heterogeneous mechanisms are proposed initially to represent the chemical reaction system. It is then necessary to verify the validity of the corresponding rate equations. The conventional scheme for verification, as first outlined by Hougen and Watson and later amplified by Yang and Hougen (*12*), consists of transforming the nonlinear rate equations into relations which are linear in the system parameters. Standard linear regression analysis can be applied to each postulated model and a criterion used such as accepting only those models which yield positive parameter estimates. Among others, Chou (*4*) has criticized this approach and suggested other criteria.

With the availability of computer techniques for nonlinear estimation, the analysis need not be constrained to the transformed rate equations. Instead it is possible to treat the rate equations in their original nonlinear form. The present paper makes use of the nonlinear estimation procedure in an attempt to answer several questions that arise in kinetic and mechanism studies. Work such as that of Box (*3*) and Peterson (*9*) has shown that for homogeneous reactions a distinct advantage is to be realized from the use of nonlinear estimation. Since the procedure is also applicable to heterogeneous mechanistic models, it may reveal features that linear estimation when used alone tends to obscure. Further, the nonlinear estimation procedure may be able to resolve the question as to whether the theoretically based Hinshelwood-Langmuir models are more justified than the empirical approach of the *pseudo order of reaction.*

In addition to studying these questions, this paper indicates a method for analyzing data from an integral reactor, a type of reactor that is sometimes dismissed in preference to a differential reactor to avoid the resulting computational analysis. On the other hand, considerable problems are often encountered in experimental work with a differential reactor, due to the small concentration changes measured. While integral reactors do not suffer from this difficulty, a penalty is, of course, paid for the possibly complex system resulting from competing reactions.

NONLINEAR ESTIMATION

Computer Program

The experimental computer program (*10*) for the present computations is a revised and expanded version of the nonlinear estimation program previously made available (*1*). Basically, nonlinear estimation (NLE) treats functional relations of the form

$$\eta = f(\mathbf{x}, \theta)$$

where x is a vector of independent variables, θ is a vector of parameters, and η, the dependent variable, may be linear or nonlinear in x and θ. If y_n is the nth experimental observation ($n = 1, 2, \ldots, N$) of the dependent variable, $\mathbf{x}_n$ the corresponding vector of values of x, and e_n the associated random error with zero mean, the relation may be represented by

$$y_n = f(\mathbf{x}_n, \theta_n) + e_n \quad n = 1, 2, \ldots, N$$

The NLE method provides a set of parameter estimates that satisfies the least-squares criterion, that is, that minimizes

$$S(\theta) = \sum_{n=1}^{N} (y_n - \eta_n)^2$$

Leon Lapidus is also with Princeton University, Princeton, New Jersey.

where $S(\theta)$ is the sum of squares of residuals. The only restriction is that η be real and single-valued for assigned values of $\mathbf{x}$ and θ. Thus, functions are acceptable when defined either by implicit differential or explicit integrated equations. Accordingly, the procedure can be extended under appropriate assumptions to cases where there are several dependent variables, such as those that occur in the rate equations that arise in the study of reaction kinetics and mechanisms.

A number of modifications have increased the efficiency and usefulness of the original program. While these changes are, in general, too detailed to describe here, it is worthwhile to mention a few. As with all iteration processes, such as NLE, the starting values chosen for the parameters that are to be estimated are extremely important in assuring convergence. For kinetic systems of the complexity under investigation here, such starting estimates must be readily obtained. The modified NLE program incorporates a feature that permits estimating a single parameter at a time, holding all others fixed. By cycling from one parameter to another all parameters can be estimated to improve the starting estimates. This cyclic single-parameter estimation yields a set of starting parameter values that is suitable for the subsequent estimation of all parameters simultaneously.

Two major routines have also been developed in conjunction with the NLE program. The first routine is a special input program that allows the chemical reaction model and additional data to be entered in a manner most familiar to the chemist or chemical engineer. The user, whether he has a command of computer programming or not, is thus able to analyze rapidly a variety of reaction mechanisms.

The second routine has the feature of converting the input model and data into an appropriate set of ordinary differential equations and integrating these equations to yield theoretical values of the dependent variables. It is these theoretical values, obtained by integrating the equations of the assumed reaction model with its estimated parameters, that the NLE program uses to find least-squares parameter estimates.

The entire experimental program was run on an IBM 7090 and 7094 using the Fortran monitor system. Typical time for a run with three-parameter estimation and a single heterogeneous reaction mechanism was 2 to 5 min., depending on the number of iterations required to reach convergence.

Application of Nonlinear Estimation

Many investigators have used the technique of nonlinear estimation to elucidate theoretical kinetic mechanisms for experimental systems. For example Box (3) has used this treatment, and Peterson (9) has applied the method to the naphthalene oxidation system where homogeneous kinetics was assumed to hold. Also, Hunter and Mezaki (6) have used the procedure to analyze heterogeneous mechanisms.

The present work extends the general nonlinear estimation treatment from homogeneous (noncatalytic) to heterogeneous (catalytic) reactions for integral conversion data. In the context of this paper, a reversible homogeneous reaction mechanism may be represented by a model of the form

$$r = k_1 \left\{ \sum_{1}^{n} c_i L_i^{P_i} \right\} - k_2 \left\{ \sum_{1}^{m} c_j L_j^{P_j} \right\} \qquad (1)$$

The terms in brackets are composed of activities L, which will be interpreted as concentrations, the stoichiometric coefficients c, the orders of reaction P, and two reaction velocity constants k_1 and k_2. Thus, for example, given experimental data for a particular reaction system, nonlinear estimation can be used to find those values for k_1 and k_2 that fit the data in the least-squares sense.

Equation (1) can, of course, be extended to many components, rate constants, and reaction steps, as in the case of complex reaction mechanisms. If the ratio of the reaction velocity constants is identified with an equilibrium constant $K_{equil.}$, Equation (1) can also be written as

$$r = k \left[\left\{ \sum_{1}^{n} c_i L_i^{P_i} \right\} - \frac{1}{K_{equil.}} \left\{ \sum_{1}^{m} c_j L_j^{P_j} \right\} \right] \qquad (2)$$

where $k = k_1$ and $K_{equil.} = k_1/k_2$. Further, this form can be extended to the heterogeneous case. For example, for a heterogeneous reaction in which surface reaction is controlling, Equation (2) becomes, when based on the Langmuir-Hinshelwood approach

$$r = \frac{k \left[\left\{ \sum_{1}^{n} c_i L_i^{P_i} \right\} - \frac{1}{K_{equil.}} \left\{ \sum_{1}^{m} c_j L_j^{P_j} \right\} \right]}{[1 + (K_A L_A)^{\alpha} + K_B L_B + \ldots]^{\beta}} \qquad (3)$$

In the form shown above there exists the possibility of estimating parameters k, $K_{equil.}$, K_A, K_B, ..., α and β, where K_A, K_B, ... are the equilibrium adsorption coefficients. Explicit forms of this equation are quite varied, but many such models are often admissible for analysis when seeking that specific model that provides the best fit to the data. Alternate formulations sometimes represent the parameter k by another term such as kK_A. Experience in this work has indicated parameter estimates obtained with these formulations are not significantly different from a relation such as that given by Equation (3).

LINEAR ESTIMATION ANALYSIS

Experimental System

The physical and chemical system selected to illustrate the applicability of nonlinear estimation is the vapor phase dehydration of ethanol as studied by Kabel (7) and presented in the more concise article by Kabel and Johanson (8). The overall reaction is relatively simple, since it involves ethanol dehydration to diethyl ether and water

$$2C_2H_5OH = C_2H_5OC_2H_5 + H_2O$$

or, in symbolic form

$$2A = B + C$$

A sulfonated copolymer of styrene and divinyl benzene in the acid form was used as the catalyst.

Of primary interest are the experimental data that were collected under carefully controlled conditions in the apparent absence of other effects such as mass transfer. This latter point is especially important, since the rate-determining steps are then associated with the chemical kinetics and are not contaminated by transport phenomena.

Reaction rate data on this chemical reaction system are available in the form of integral conversions. Concentrations in the effluent from the reactor were measured as a function of reciprocal space velocity. Pressure and temperature were maintained constant throughout the reactor.

Two types of experimental runs were performed. One series was made with the feed composition set at $A_o = 1.0$, $B_o = C_o = 0$, that is pure A in the feed. Another series of runs was made in which all three components were present to varying degrees in the feed stream, that is $A_o \pm 1.0$. From these runs only those taken under essentially identical isothermal and isobaric conditions of 120°C. and 1 atm. total pressure were selected for the present analysis.

Of particular note is the accuracy and precision of the data obtained. Kabel indicated that flow measurements had absolute errors in the third decimal place and that concentration measurements were accurate to the second and third decimal place. These estimates varied, depending on the conditions used during a given experiment. Since there are many other possible sources of error (for example, catalyst activity variation, temperature control and measurement), the most conservative estimate would be that the data were accurate to at least one part in the second significant figure.

Analysis of Data

In analyzing these experimental data, Kabel considered principally three possible heterogeneous mechanisms:

The reaction between adjacently adsorbed ethanol molecules, which yields a rate equation of the form

$$(a) \qquad r = \frac{k\,[A^2 - (BC/K_{equil.})]}{[1 + K_A A + K_B B + K_C C]^2}$$

The reaction between an adsorbed ethanol molecule and a gaseous ethanol molecule, which may be represented by

$$(b) \qquad r = \frac{k\,[A^2 - (BC/K_{equil.})]}{[1 + K_A A + K_B B + K_C C]}$$

The reaction between a gaseous ether molecule and a water molecule adsorbed on a single site, which gives

$$(c) \qquad r = \frac{k\,[A - \sqrt{BC/K_{equil.}}]}{[1 + \sqrt{BCK_A^2/K_{equil.}} + K_B B + K_C C]}$$

In the context of the present work Equations (*a*) and (*b*) may be referred to as surface-reaction controlling models (S.R.) with exponent β equal to 2 and 1, respectively. Equation (*c*) is a form of an adsorption-controlling mechanism (A.C.).

One of the objectives of the analysis carried out by Kabel was to determine the reaction mechanism that best represents the experimental data. It was thus necessary to find values for parameters k, K_A, K_B, and K_C, which when inserted into the reaction equation yielded the best fit to the experimental data. A least-squares analysis was applied to transformed linear forms of the initial rate equations. Parameter estimates so obtained were used to compute integral conversions, which were compared with observed conversions. From this analysis Kabel concluded that surface reaction mechanism (*a*) provided the best fit. At 120°C., by the use of an independently determined value for $K_{equil.} = 25.2$, the parameter values determined were

$k = 2.66 \times 10^{-3}$ $\qquad K_B = 7.0$ $\qquad$ (Columnar
$K_A = 3.4$ $\qquad K_C = 0$ $\qquad$ results)

In addition to these columnar experiments, Kabel performed a series of independent equilibrium adsorption experiments that yielded data suitable for determining equilibrium adsorption coefficients. The results of this work, showing excellent agreement with the columnar results, were

$K_A = 2.5$ $\qquad$ (Absorption
$K_B = 7.6$
$K_C = 0$ $\qquad$ results)

In summary, then, Kabel's work on the catalytic dehydration of ethanol represents extensive and very carefully obtained integral conversion data. Based on a linear least-squares fit to initial rate data and subsequent integration with the estimated parameters, a surface-reaction type of Langmuir-Hinshelwood model was selected. Independent adsorption measurements tended to confirm this result, since parameter values obtained by this method and columnar experiments were essentially identical.

NONLINEAR ESTIMATION ANALYSIS

Preliminary Analysis

For the analysis of this paper, an initial selection was made from the experimental data presented by Kabel. Four runs with $A_o = 1.0$, resulted in a total of forty-seven data points, while nine runs with $A_o \neq 1.0$, yielded a total of seventy data points. These runs are shown in Table 1 as data Sets *A* and *B*, respectively, with subdivisions A_1, A_2, A_3, A_4, indicating the four series comprising Set *A*. Corresponding run numbers assigned by Kabel are indicated. Table 1 also presents a partial summary of further subsets of the overall data. These subdivisions resulted from a preliminary study of the data and an examination of the residuals between the observed data and the computed values. Various models were used in this exploratory phase. By way of illustration, using the S.R. ($\beta = 2$) model and data Set *A*, Sets A_1 and A_2 exhibited essentially all negative residuals, while sets A_3 and A_4 had all positive residuals. On the other hand, Sets A_3 and A_4 combined as one set, displayed fairly random residuals. Thus, an appropriate subset selected for subsequent analysis was data Set *E*. Similarly, Set *G* formed from data Set *B* was chosen for further analysis.

In this preliminary phase of the analysis, certain features are of interest. For example, given the forty-seven data points corresponding to $A_o = 1.0$ and surface reaction (S.R.) models such as (*a*) and (*b*), it is possible to determine the least-squares values for the parameters, k, K_A, K_B, $K_{equil.}$ in addition to the exponent β in the denominator term (K_C is assumed to be zero). Table 2 illustrates a path in such a calculation where four sequential stages of estimation were used. In stage 1, K_A, K_B, and β were held fixed, while k and $K_{equil.}$ (the values in brackets) were estimated. Shown with the resulting least-square estimates for k and $K_{equil.}$ is the standard deviation of the residuals, s, that is the square root of the residual mean square. Subsequent stages of the estimation

Table 1. Partitioning of Experimental Data Into Subsets

Original sets of data*

Set *A*

$A_o = 1.0$, 4 runs
$A_1 = 15$ data (run 1-3)
$A_2 = 14$ data (run 2-1)
$A_3 = 7$ data (run 3-1)
$A_4 = 11$ data (run 4-1)

47 total data points

Set *B*

$A_o \neq 1.0$, 9 runs
$B_1 = 6$ data (run 3-4)
$B_2 = 7$ data (run 3-5)
$B_3 = 8$ data (run 3-6)
$B_4 = 8$ data (run 3-7)
$B_5 = 8$ data (run 3-8)
$B_6 = 7$ data (run 3-9)
$B_7 = 9$ data (run 1-4)
$B_8 = 9$ data (run 3-2)
$B_9 = 8$ data (run 3-3)

70 total data points

Composite subsets of data

Set *E*
$A_o = 1.0$, 2 runs
$E_1 = A_3$
$E_2 = A_4$

18 total data points

Set *G*
$A_o \neq 1.0$, 3 runs
$G_1 = B_7$
$G_2 = B_8$
$G_3 = B_9$

26 total data points

* Run numbers correspond to those by Kabel (7).

process are evident from Table 2. The end result of this, or any alternate cycling procedure, is a best set of estimated parameters in the least-squares sense. As evident from Table 2, the initial estimates were so close to the final values that no large changes resulted. Of interest more generally, however, is the fact that it is common to experience more difficulty in estimating K_A and K_B than k or $K_{equil.}$.

Analysis of Data

The next step in the analysis was to compare the fits resulting from the use of nonlinear estimation when applied to different reaction models. In particular, the three models detailed previously by Equations (*a*), (*b*), and (*c*) were examined along with a desorption controlling (D.C.) model

$$(d) \qquad r = \frac{k\,[A^2 - BC/K_{equil.}]}{1 + K_A P_A + K_B\, K_{equil.} \dfrac{A^2}{C}}$$

and a homogeneous model

$$(e) \qquad r = k\,[A^2 - BC/K_{equil.}]$$

The data sets shown in Table 1, as well as others, were used in this analysis. Calculations were carried out for three different conditions: (1) The equilibrium constant $K_{equil.}$ was treated as a parameter to be estimated, although it is usually considered a thermodynamic constant and can be independently determined. (2) $K_{equil.}$ was fixed at 25.2 corresponding to the value determined by Kabel by equilibrium composition measurements. And (3) the adsorption coefficients K_A and K_B were fixed at various predetermined values and only k and $K_{equil.}$ were estimated. As in Kabel's work, K_C was set equal to zero.

Pertinent results are tabulated in Table 3. Detailed analysis is shown for data Set *A* and more abbreviated results displayed for data Sets *B*, *E*, and *G*.

TABLE 2. CYCLIC NONLINEAR ESTIMATION PROCEDURE FOR SURFACE REACTION MODEL. DATA SET *A*

Parameters

Stage	$k(\times 10^{-3})$	$K_{equil.}$	K_A	K_B	β	$s(\times 10^{-2})$
1	[2.77]	[1.25]	3.40	7.00	2.0	1.43
2	2.77	1.25	[3.44]	[6.63]	2.0	1.42
3	[2.62]	[1.18]	[3.32]	[6.24]	2.0	1.42
4	2.62	1.18	3.32	6.24	[2.0]	1.41

Numbers in brackets are parameter estimates obtained by nonlinear estimation; all other values held fixed.

TABLE 3. NONLINEAR ESTIMATION WITH VARIOUS REACTION MODELS

I. Data Set *A*

a. $K_{equil.}$ estimated
Parameters

Run	$k(\times 10^{-3})$	$K_{equil.}$	K_A	K_B	$s(\times 10^{-2})$
S.R., $\beta = 2$ [Surface reaction]	[2.62] (0.35)	1.18	[3.32] (0.28)	[6.24] (1.05)	1.43
S.R., $\beta = 1$ [Surface reaction]	2.77	1.24	[18.7] (0.6)	[35.2] (4.8)	1.42
A.C. [Adsorption]	0.166	2.52	[—2.69] (3.64)	[1.08] (0.26)	1.80
D.C. [Desorption]	0.314	0.260	[3.17] (0.42)	[0.363] (0.080)	4.35
Hom. [Homogeneous]	[0.142] (0.002)	[1.22] (0.17)	—	—	1.42

b. $K_{equil.} = 25.2$
Parameters

Run	$k(\times 10^{-3})$	K_A	K_B	$s(\times 10^{-2})$
S.R., $\beta = 2$	[2.63] (0.05)	3.04	10.0	1.99
S.B., $\beta = 1$	0.254	[14.7] (0.7)	[62.7] (5.7)	1.89
A.C.	0.160	[82.6] (36.6)	[0.349] (0.260)	1.88
D.C.	0.0340	[0.276] (0.040)	[0.0014] (0.0001)	2.02
Hom.	[0.135] (0.003)	—	—	2.57

c. $K_{equil.}$ estimated
Parameters

Run	$k(\times 10^{-3})$	$K_{equil.}$	K_A	K_B	$s(\times 10^{-2})$
S.R., $\beta = 2$	2.69	2.17	3.04 [K_A, K_B fixed]	10.0	1.81
S.R., $\beta = 1$	2.37	1.37	3.00 [K_A, K_B fixed]	7.00	1.48

II. Data Set *E*

a. $K_{equil.} = 25.2$
Parameters

Run	$k(\times 10^{-3})$	K_A	K_B	$s(\times 10^{-2})$
S.R., $\beta = 2$	[2.33] (0.02)	3.26	6.91	0.403
S.R., $\beta = 1$	2.28	[16.6] (0.4)	[36.4] (4.4)	0.415
A.C.	[0.141] (0.001)	—202	1.49	0.465
A.C.	[0.133] (0.001)	35.4	0	0.472
Hom.	[0.128] (0.001)	—	—	0.408

b. $K_{equil.}$ estimated
Parameters

Run	$k(\times 10^{-3})$	$K_{equil.}$	K_A	K_B	$s(\times 10^{-2})$
S.R., $\beta = 2$	2.36	7.30	3.26 [K_A, K_B fixed]	6.91	0.416
Hom.	[0.129] (0.001)	[13.8] (0.1)	—	—	0.414

III. Data Set *B*

$K_{equil.} = 25.2$
Parameters

Run	$k(\times 10^{-3})$	K_A	K_B	$s(\times 10^{-2})$
S.R., $\beta = 2$	2.66	[4.20] (0.44)	[0.874] (0.960)	7.53
A.C.	0.208	[490] (140)	[—0.547] (0.230)	4.60
A.C.	0.236	[460] (113)	0	4.80
D.C.	0.0298	[0.306] (0.150)	[0.00117] (0.011)	4.40
Hom.	[0.155] (0.013)	—	—	5.12

IV. Data Set *G*

$K_{equil.} = 25.2$
Parameters

Run	$k(\times 10^{-3})$	K_A	K_B	$s(\times 10^{-2})$
S.R., $\beta = 2$	1.90	[—0.46] (2.11)	[27.6] (14.2)	8.54
Hom.	[0.130] (0.009)	—	—	3.90

The tabulated data of Table 3 lead to the following conclusions. Data Set *A* with the equilibrium constant estimated (*Ia*) shows that the desorption model is the poorest fit of all models. This is evident by the large value of the standard deviation of the residuals, 0.0435. Twice this standard deviation is an error of almost one part in the first significant decimal place. Since the conversion data were obtained with an accuracy of at least one part in the second decimal figure, the fit with the desorption model is outside a tolerable level. The adsorption controlling model is also eliminated because of the relatively large standard deviation of residuals and the negative value for one of the estimated adsorption coefficients. There remain as possible candidates for the best model the two surface reaction and the homogeneous models. As can be seen in Table 3 it is essentially impossible to discriminate among these remaining models, since the standard deviation of residuals and the standard deviations of parameter estimates reveal equally determined fits to the data. However, even in these cases, twice the standard deviation is about 0.03, which indicates a relatively poor fit. On the basis of these results it appears impossible to find that mechanistic model that is a unique representation of the data. Possible reasons for this difficulty and further discussion of this point will be deferred until later.

Also of interest is a comparison of these results with the case where the equilibrium constant is fixed at its independently determined value of 25.2 (*Ib*). The resulting standard deviations of residuals for the three most probable reaction models is higher than in the previous case. This would imply that the independently determined value of the equilibrium constant is not commensurate with the rate data; that is, there are experimental difficulties in obtaining a consistent value or the true heterogeneous mechanism is not one of those possibly isolated. Instead, the nonlinear estimation analysis indicates that a $K_{equil.}$ of 1.0 to 2.0 is better than 25.2. It must be recalled, however, that this conclusion should be considered tentative, since the fit to the experimental data is not good.

Analysis of data Sets *B* and *G* leads to the same conclusions; that is, none of the models provides a good fit to the data, and discrimination among models is difficult, if not impossible.

Of the results obtained, only data Set *E* seemed suitable for obtaining a conclusive answer regarding the most probable model. Table 3 reveals that for data Set *E* the standard deviation of residuals is an order of magnitude smaller than for other data, the value being approximately 0.004. Reaction mechanisms for this set of data therefore provide an adequate fit. Unfortunately, however, there does not seem to be any way to discriminate as to which model is the best one, since several models result in about equally probable parameter estimates and equal standard deviations of residuals. It is interesting to note that the surface reaction model with $\beta = 2$ yields estimated values that largely agree with those of Kabel. By way of illustration, k by nonlinear estimation is 2.33×10^{-3} vs. 2.66×10^{-3} by Kabel, K_A 3.26 vs. 3.4 and K_B 6.91 vs. 7.0. However, instead of an equilibrium constant of 25.2, the estimated value is 7.3. Aside from these correspondences, it is not possible to assert that the surface reaction model with $\beta = 2$ is a better one than either the surface reaction with $\beta = 1$, or even the simple homogeneous model. In the next section a possible reason for this lack of discrimination will be discussed.

In summary, the results of this section indicate that a homogeneous type rate model is equally as probable for describing the reaction as the heterogeneous models. By using only heterogeneous models it is still impossible to conclude that any one model is more suitable than another unless external information is also considered. This conclusion, while not unknown to investigators of kinetics mechanisms, may seem rather surprising, since extensive data are available for estimating only a few numbers of parameters.

Simulated Data Study

It can, of course, be reasoned that the fact that there appears to be no heterogeneous mechanism better than the homogeneous one is due to a competition of heterogeneous mechanisms; that is, the experimental system under investigation does not have a single rate-determining heterogeneous step. This immediately raises the question whether, given a system that is known to have a single rate-determining step, it is possible by estimation techniques to recover that model. To elucidate the answer to this question a set of theoretical data was generated using a surface-reaction model, $\beta = 2$. Fourteen data points were used and these were subjected to random errors in the second and third decimal places. These data were then analyzed by nonlinear estimation to recover the parameter values. Thus by specifying numerical values of k, $K_{equil.}$, K_A, and K_B, the differential rate equation representing a surface-reaction mechanism was numerically integrated for different space velocities. This calculation yielded exact values of A at corresponding settings of space velocities. These values were then subjected to random deviations to generate simulated experimental data.

Table 4 presents the results of nonlinear estimation analysis for the two different hypothetical reaction systems. In the first system corresponding to $k = 20.0 \times 10^{-3}$, $K_A = 5.0$, $K_B = 10.0$, and $K_{equil.} = 30.0$, the standard deviation of the residuals is approximately 0.001 for the different models tried. This corresponds directly to the perturbation made in the third decimal place with the exact data.

The adsorption controlling model can be discarded because of the negative estimate of K_A, but all other models seem to be adequate. Thus the results of this series of calculations would appear to confirm the lack of discrimination among models; that is, the homogeneous model provides as good a fit to the data as the heterogeneous model.

This result is rather disturbing, since it is known in advance that the surface reaction $\beta = 2$ mechanism is the correct one. However, further examination provides a simple explanation of the results. Because of the material balance relations the following must always hold

$$B = \frac{1 - A}{2}$$

Now if the denominator in the surface reaction model is examined it can be seen that

$$[1 + K_A A + K_B B]^\beta = \left[\,1 + 5A + 10\left(\frac{1-A}{2}\right)\right]^\beta = 6^\beta$$

In other words the choice of $K_A = 5.0$ and $K_B = 10.0$ leads to a degenerate system in which the denominator is a constant rather than a function of A and B. Thus the surface reaction and homogeneous models should give essentially identical results except for the estimated values of k. The lack of model discrimination is due to this degeneracy condition.

To examine this point further, another series of runs was made with a second hypothetical case in which $k = 25.0 \times 10^{-3}$, $K_A = 4.5$, $K_B = 14.7$, and $K_{equil.} = 10.5$. The results are shown in Table 4. Considering that the

TABLE 4. NONLINEAR ESTIMATION FOR HYPOTHETICAL REACTION SYSTEMS

I. Data perturbed in third decimal place

A. $k = 20.0 \times 10^{-3}, K_A = 5.0, K_B = 10.0, K_{equil.} = 30.0$

Run	$k(\times 10^{-3})$	K_A	K_B	$s(\times 10^{-2})$
S.R., $\beta = 2$	[23.4] (0.1)	[5.49] (0.67)	[11.0] (1.5)	0.102
S.R., $\beta = 1$	[18.3] (0.0)	[31.8] (5.8)	[64.1] (13.0)	0.0996
A.C.	[0.560] (0.090)	[—3.44] (5.99)	[0.054] (0.091)	0.0963
Hom.	[0.556] (0.001)	—	—	0.0904

B. $k = 25.0 \times 10^{-3}, K_A = 4.5, K_B = 14.7, K_{equil.} = 10.5$

Run	$k(\times 10^{-3})$	K_A	K_B	$s(\times 10^{-2})$
S.R., $\beta = 2$	[22.4] (0.0)	4.23	13.8	0.0912
S.R., $\beta = 1$	[4.27] (0.01)	3.98	20.9	0.103
A.C.	0.899	[—11.3] (1.5)	[1.56] (0.03)	0.139
Hom.	[0.557] (0.013)	—	—	1.81

II. Data perturbed in second decimal place

A. $k = 25.0 \times 10^{-3}, K_A = 4.5, K_B = 14.7, K_{equil.} = 10.5$

Run	$k(\times 10^{-3})$	K_A	K_B	$s(\times 10^{-2})$
S.R., $\beta = 2$	[25.0] (0.4)	4.78	13.9	1.04
S.R., $\beta = 1$	[4.26] (0.63)	4.59	19.2	1.04
A.C.	[0.897] (0.002)	—22.1	1.82	1.09
Hom.	[0.556] (0.013)	—	—	1.77

data were perturbed in the third decimal place, it can be seen that the best model is the surface reaction case, $\beta = 2$. The standard deviation for the homogeneous model is at least ten times greater than that for the surface reaction case. Further, the adsorption controlling model yields a negative parameter value, while the standard deviation of residuals for the surface reaction $\beta = 1$ is somewhat larger than that for the $\beta = 2$ case. The results in Table 4, which correspond to perturbations in the second decimal place, confirm these results, but not with the same assurance as the hypothetical case with perturbations in the third decimal place. This trend is, of course, to be expected.

In summary, the nonlinear estimation technique permits not only an adequate fit to the experimental data, but also the ability to discriminate among alternate models. Further, the best fit is obtained with the correct model.

Returning now to the ethanol dehydration analysis, it can be seen that this experimental system is subject to the same degeneracy condition in that $K_A \simeq 3.4$ and $K_B \simeq 7.0$. Thus the lack of discrimination among models is probably due not to defects in the nonlinear estimation analysis, but rather to an accidental arrangement of the physical parameters. Raising the experimental temperature would probably remove the degeneracy, assuming that the ratio of K_A and K_B changed.

TABLE 5. LINEAR ESTIMATION FOR HYPOTHETICAL SYSTEM OBSERVED DATA PERTURBED IN THIRD DECIMAL PLACE

$k = 25.0 \times 10^{-3}, K_A = 4.5, K_B = 14.7, K_{equil.} = 10.5$

Run	$k(\times 10^{-3})$	$K_{equil.}$	K_A	K_B
S.R., $\beta = 2$	17.7	*	3.63	12.1
S.R., $\beta = 1$	16.8	*	1.51	5.75
A.C.	24.8	*	154	0.430
D.C.	0.358	*	—0.709	0.176
Hom. 1	0.552	5.00	—	—
Hom. 2	†	1.11	—	—

* $K_{equil.}$ fixed at 10.5.
† k fixed at 25.0×10^{-3}.

There is one further point that has been explored with these hypothetical data. An alternate analysis technique previously mentioned is to estimate parameters and discriminate among various models by applying linear regression. Thus in the surface reaction case, $\beta = 2$

$$(a-0) \quad r = \frac{k\,[A^2 - BC/K_{equil.}]}{[1 + K_A A + K_B B]^2}$$

This equation can be rearranged to

$$(a-1) \quad \left[\frac{A^2 - BC/K_{equil.}}{r}\right]^{1/2} = \frac{1}{\sqrt{k}} + \frac{K_A}{\sqrt{k}}A + \frac{K_B}{\sqrt{k}}B$$

Or, equivalently

$$(a-2) \quad y = a_0 + a_1 A + a_2 B$$

With the use of the homogeneous and heterogeneous models previously considered the hypothetical data with perturbations in the third decimal place were analyzed by linear regression. Table 5 shows the results for the nondegenerate case in which the nonlinear estimation analysis was previously able to make an excellent discrimination of models. Of the six models considered by linear regression only one can be discarded on the basis of negative parameter estimates. The remaining five models, including two homogeneous models, are all suitable for representing the data, if the criterion is used that models having positive parameter estimates should be retained. Of interest is the fact that the surface reaction case $\beta = 2.0$ yields parameter values in moderate agreement with the hypothetical values. Unfortunately, if the data came from an experimental system, this postiori information would be of little use. The standard deviations of residuals, while not shown in Table 5, reveal that the surface reaction models are superior to all other models. The homogeneous models are the poorest by far in fitting the data.

CONCLUSIONS

When the nonlinear estimation technique is applied to a consistent set of data, it can be used to discriminate among various possible reaction models to find the best model. Linear estimation when used alone may not be satisfactory for making this discrimination. However, such problems as degeneracy of the heterogeneous equations and errors in measurement are shown to be obscuring effects in determining the best model. When these difficulties are involved, it is questionable that the relatively complex heterogeneous rate equations are more warranted than the simpler homogeneous equations.

ACKNOWLEDGMENT

The authors wish to acknowledge the computer time on the IBM 7090 and 7094 made available by Princeton University (N.S.F. Grant G.P.-579) and the IBM Corporation.

This work is taken in part from the dissertation of T. I. Peterson, submitted to the faculty of the Polytechnic Institute

of Brooklyn in partial fulfillment of the requirements for the degree of doctor of philosophy.

NOTATION

A,B,C = concentrations of components as function of time
A_o,B_o,C_o = initial (or feed) concentrations of components
c_i = stoichiometric coefficients in reaction kinetics formulation
k_1,k_2 = reaction velocity constants
k = k_1
$K_{equil.}$ = equilibrium constant = k_1/k_2
K_A,K_B = equilibrium adsorption coefficients
L_i = generalized components in reaction kinetics formulation
P_i = powers in reaction kinetics formulation
r = rate of reaction
s = standard deviation of residuals
$\mathbf{x}$ = vector of independent variables
y = observations on dependent variable
α,β = powers in denominator of Langmuir-Hinshelwood rate expressions
η = dependent variable
ϕ = vector of parameters

LITERATURE CITED

1. Booth, G. W., and T. I. Peterson, "A.I.Ch.E. Computer Program Manual No. 3," A.I.Ch.E., New York (1960); IBM Computer Program WL NL1, Distribution No. 687. Available from SHARE Library, IBM Corp., New York (1959).
2. Boudart, M., *A.I.Ch.E. J.*, **2**, 62 (1956).
3. Box. G. E. P., *Ann. N. Y. Acad. Sci.*, **86**, 792 (1960).
4. Chou, C., *Ind. Eng. Chem.*, **50**, 799 (1958).
5. Hougen, O. A., and K. M. Watson, "Chemical Process Principles, Part 3," Wiley, New York (1947).
6. Hunter, W. G., and R. Mezaki, *A.I.Ch.E. J.*, **10**, 315 (1964).
7. Kabel, R. L., Ph.D. thesis, Distr. No. 61-3991, University Microfilms, Inc., Ann Arbor, Michigan (1961).
8. ———, and L. N. Johanson, *A.I.Ch.E. J.*, **8**, 621 (1962).
9. Peterson, T. I., *Chem. Eng. Sci.*, **17**, 203 (1962).
10. Peterson, T. I., and Leon Lapidus, IBM Computer Program, Distribution No. T2IBM0014. Available from SHARE Library, IBM Corp., New York (1964).
11. Weller, S., *A.I.Ch.E. J.*, **2**, 59 (1956).
12. Yang, K. H., and O. A. Hougen, *Chem. Eng. Progr.*, **46**, 146 (1950).

Manuscript received November 24, 1964; revision received April 29, 1965; paper accepted May 3, 1965. Paper presented at A.I.Ch.E. Boston meeting.

Interaction Model for Critical Temperatures of Multicomponent Mixtures of Methane-Free Aliphatic Hydrocarbons

OKAN EKINER and GEORGE THODOS
Northwestern University, Evanston, Illinois

A mathematical treatment based on an interaction model has yielded an expression capable of establishing the critical temperatures of multicomponent aliphatic hydrocarbon mixtures. This model postulates an infinite-series expansion for the difference between the actual critical temperature and its corresponding pseudocritical value. For methane-free aliphatic hydrocarbon mixtures, this infinite series has been truncated beyond the third-order interactions. Explicit relationships are given for the second-order interaction coefficients A_{ij} and B_{ij} as a function of the binary-system parameter τ_{ij} and the third-order interaction coefficient β_{ijk} as a function of χ_{ijk}, the ternary-system parameter.

This interaction model has been applied to one hundred-sixteen different compositions of binary, ternary, quaternary, and quinary aliphatic hydrocarbon systems to produce an overall average deviation of 0.35%. Owing to the different interaction behavior of methane, systems containing methane have not been included in this study.

The prediction of the critical temperatures of binary and multicomponent hydrocarbon systems has been a subject of considerable speculation, treated for the most part from empirical approaches. These considerations utilize convenient parameters for the establishment of basic constants pertinent to each of these studies (*2, 3, 5–7, 13*). The present study attempts to employ a direct mathematical development which utilizes an interaction model as a basis to formulate the governing relationships for establishing the critical temperature of a multicomponent system.

MATHEMATICAL DEVELOPMENT

The model postulated in this study utilizes the results of the interactions of the different species to establish the critical temperature of an n component mixture. For this purpose, the following series expansion is proposed:

$$T_c = \sum_{i=1}^{n} \alpha_i n_i + \sum_{i,j}^{n} \alpha_{ij} n_i n_j + \sum_{i,j,k}^{n} \alpha_{ijk} n_i n_j n_k + \sum_{i,j,k,l}^{n} \alpha_{ijkl} n_i n_j n_k n_l + \dots \quad (1)$$

In Equation (1), α_i, α_{ij}, α_{ijk}, α_{ijkl} ... represent coefficients resulting from all the possible interactions of the species in the mixture. The subscripts i, j, k, l, ... denote the interacting species, and n_i, n_j, n_k, n_l ... are the mole fractions of components i, j, k, l, ... in the mixture.