y = coefficient, dimensionless
θ = specific loading, lb. solids/lb. fluid
ρ_d = dispersed density of solids, lb./cu.ft. [as given by Equation (2)]
ρ_f = density of fluid media, lb./cu.ft.
χ = surface shape factor of solid particles, dimensionless

LITERATURE CITED

1. Allendorf, H. D., M.S. thesis, Pennsylvania State Univ., University Park (1960).
2. Belden, D. H., and L. S. Kassel, *Ind. Eng. Chem.*, **41**, 1174 (1949).
3. Cramp, W., and A. Priestly, *Engineer*, **137**, 89 (1924).
4. Frantz, J. F., *Chem. Eng.*, **69**, No. 19, 161; No. 20, 89; No. 22, 103 (1962).
5. Gopichand, T., K. J. R. Sarma, and M. N. Rao, *Ind. Eng. Chem.*, **51**, 1449 (1959).
6. Hariu, O. H., and M. L. Molstad, *ibid.*, **41**, 1148 (1949).
7. Hinkle, B. L., Ph.D. thesis, Georgia Inst. Technol. (1953).
8. Jones, J. H., W. G. Braun, T. E. Daubert, and H. D. Allendorf, *A.I.Ch.E. J.*, **12**, 1070 (1966).
9. Korn, A. H., *Chem. Eng.*, **57**, No. 3, 108 (1950).
10. Leva, Max, "Fluidization," McGraw-Hill, New York (1959).
11. Mehta, N. C., J. M. Smith, and E. W. Comings, *Ind. Eng. Chem.*, **49**, 986 (1957).
12. Razumov, I. M., *Intern. Chem. Eng.*, **2**, No. 4, 539 (1962).
13. Uspenskii, V. A., *Ekon. Topliva Za.*, **8**, No. 3, 26 (1951).
14. Zenz, F. A., *Petrol. Refiner*, **36**, No. 4, 173; No. 5, 261; No. 6, 133; No. 7, 175; No. 8, 147; No. 9, 305; No. 10, 162; No. 11, 321 (1957).
15. Zenz, F. A., and D. F. Othmer, "Fluidization and Fluid Particle Systems," Reinhold, New York (1960).

The Use of Residual Analysis for Building Mechanistic Models

WILLIAM J. HILL

Allied Chemical Corporation, Buffalo, New York

REIJI MEZAKI

Yale University, New Haven, Connecticut

The use of residuals of a diagnostic parameter for kinetic modeling was developed and demonstrated in two previous papers (*1, 2*). In the second paper (*2*) the importance of the selection of an initial model was emphasized for the successful use of the proposed technique. The aims of this communication are threefold. First, by means of mathematical statistics, we wish to prove the existence of definite correlation between the residuals associated with an inadequate model and the concentration level of an improperly described component. Second, we wish to illustrate that an appropriate choice of an experimental design can simplify the analysis. Third, we will illustrate a situation where a residual analysis may be misleading when not interpreted properly.

RESIDUALS AND CONCENTRATION LEVELS

In the earlier papers, Hougen-Watson reaction models were postulated for the complete catalytic oxidation of methane and the following model was found to be appropriate: gaseous methane and adsorbed oxygen react and produce both adsorbed carbon dioxide and adsorbed water. In the true case, the value of the diagnostic parameter C_1 can be given by

$$C_1 = b_o + b_2 x_2 + b_3 x_3 + b_4 x_4 + \epsilon \qquad (1)$$

Suppose instead we consider, as an initial trial model, a model for which the value of the calculated diagnostic parameter is

$$\hat{C}_1 = \hat{b}_o + \hat{b}_2 x_2 + \hat{b}_4 x_4 \qquad (2)$$

Note that this model implies that gaseous methane reacts with adsorbed oxygen, producing nonadsorbed carbon dioxide and adsorbed water.

Now Equation (1) may be simplified to the matrix form

$$\mathbf{C}_1 = \mathbf{X}_1 \underline{\theta}_1 + \mathbf{X}_3 \underline{\theta}_3 + \underline{\epsilon} \qquad (3)$$

where

$$\mathbf{X}_1 = [\mathbf{1}, \mathbf{x}_2, \mathbf{x}_4]$$
$$\mathbf{X}_3 = [\mathbf{x}_3]$$
$$\underline{\theta}_1 = \begin{bmatrix} b_o \\ b_2 \\ b_4 \end{bmatrix}$$
$$\underline{\theta}_3 = [b_3]$$

$\mathbf{C}_1$ is the $n \times 1$ vector of the diagnostic parameters (n is the number of observations), $\mathbf{X}_1$ is an $n \times 3$ matrix, and $\mathbf{X}_3$ is an $n \times 1$ vector. $\underline{\theta}_1$ and $\underline{\theta}_3$ are 3×1 and 1×1 vectors, respectively.

Similarly, Equation (2) in vector form is

$$\hat{\mathbf{C}}_1 = \mathbf{X}_1 \hat{\underline{\theta}}_1 \qquad (4)$$

where

$$\hat{\underline{\theta}}_1 = \begin{bmatrix} \hat{b}_o \\ \hat{b}_2 \\ \hat{b}_4 \end{bmatrix}$$

The expected value of $\hat{\mathbf{C}}_1$, $E(\hat{\mathbf{C}}_1)$, is obtained from

$$E(\hat{\mathbf{C}}_1) = E(\mathbf{X}_1 \hat{\underline{\theta}}_1) = \mathbf{X}_1 \; E(\hat{\underline{\theta}}_1) \tag{5}$$

Introducing the following relationship to Equation (5) for the least squares estimate of $\underline{\theta}_1$

$$\hat{\underline{\theta}}_1 = (\mathbf{X}_1' \; \mathbf{X}_1)^{-1} \; \mathbf{X}_1' \; \underline{C}_1 \tag{6}$$

we obtain

$$\begin{aligned} E(\hat{\mathbf{C}}_1) &= \mathbf{X}_1 \; (\mathbf{X}_1' \; \mathbf{X}_1)^{-1} \; \mathbf{X}_1' \; E(\mathbf{C}_1) \\ &= \mathbf{X}_1 \; (\mathbf{X}_1' \; \mathbf{X}_1)^{-1} \; \mathbf{X}_1' \; E(\mathbf{X}_1 \; \underline{\theta}_1 + \mathbf{X}_3 \; \underline{\theta}_3) \\ &= \mathbf{X}_1 \; \underline{\theta}_1 + \mathbf{X}_1 \; (\mathbf{X}_1' \; \mathbf{X}_1)^{-1} \mathbf{X}_1' \; \mathbf{X}_3 \; \underline{\theta}_3 \end{aligned} \tag{7}$$

Thus the expected values of the residuals for C_1 using the model given in Equation (2) are

$$\begin{aligned} E(\mathbf{C}_1 - \hat{\mathbf{C}}_1) &= \mathbf{X}_1 \; \underline{\theta}_1 + \mathbf{X}_3 \; \underline{\theta}_3 - \mathbf{X}_1 \; \underline{\theta}_1 \\ &\quad - \mathbf{X}_1 (\mathbf{X}_1' \; \mathbf{X}_1)^{-1} \; \mathbf{X}_1' \; \mathbf{X}_3 \; \underline{\theta}_3 \\ E(\mathbf{C}_1 - \hat{\mathbf{C}}_1) &= \underline{\delta} \; \mathbf{X}_3 \; \underline{\theta}_3 \end{aligned} \tag{8}$$

where

$\underline{\delta} = \mathbf{I} - \mathbf{X}_1 \, (\mathbf{X}_1' \, \mathbf{X}_1)^{-1} \, \mathbf{X}_1'$ and $\mathbf{I}$ is the $n \times n$ identity matrix.

If the experiments are designed such that $\mathbf{X}_1$ is orthogonal to $\mathbf{X}_3$, then $\mathbf{X}_1' \; \mathbf{X}_3 = \mathbf{O}$. (See reference *1* for an example of such a design.) Therefore

$$E(\mathbf{C}_1 - \hat{\mathbf{C}}_1) = \mathbf{X}_3 \; \underline{\theta}_3 \tag{9}$$

This relationship between the residuals and component $\mathbf{X}_3$ is clearly appreciated. That is, the residuals associated with the trial model are correlated with the concentration x_3, that of carbon dioxide. Notice that the choice of experimental design can simplify the analysis, making the relationship between the residuals and a certain variable more apparent.

On the other hand, if we select

$$\hat{C}_1 = \hat{b}_o + \hat{b}_2 \; x_2 + \hat{b}_3 \; x_3 \tag{10}$$

as trial prediction model, we can derive the following relationship:

$$E(\mathbf{C}_1 - \hat{\mathbf{C}}_1) = \mathbf{X}_4 \; \underline{\theta}_4 \tag{11}$$

Equation (12) indicates that the residuals associated with this model are correlated with the concentration x_4, that of water.

When the trial model is actually the true model we can expect

$$E(\mathbf{C}_1 - \hat{\mathbf{C}}_1) = \mathbf{O} \tag{12}$$

This implies that the residuals should be random and should show no systematic correlation with the concentration of any component. This is the most desirable situation.

Hougen-Watson types of rate equations frequently can be linearized and then can be arranged into the linear forms as shown in Equation (1). In cases where an efficient experimental design such as an orthogonal design is employed, the association between the residuals of an inadequate model and the concentration of a certain component is often obvious and hence corrections can be made in the model to eliminate this inadequacy.

MISLEADING RESULTS

A study like that described above can lead to misleading interpretations if certain things are not understood. For example, the converse of Equation (12) is not necessarily true. That is, if we find that the expected values of residuals associated with a particular model are zero, this does not imply that the model is the best representation of the true equation. In fact, models that have more parameters than necessary for the true case will satisfy Equation (12). To illustrate this suppose a calculated trial model is taken to be

$$\hat{\mathbf{C}}_1 = \mathbf{X}_1 \; \hat{\underline{\theta}}_1 + \mathbf{X}_2 \; \hat{\underline{\theta}}_2 \tag{13}$$

when the true model is actually

$$\mathbf{C}_1 = \mathbf{X}_1 \; \underline{\theta}_1 + \underline{\epsilon} \tag{14}$$

Note that the true model does not include the extra terms $\mathbf{X}_2 \; \hat{\underline{\theta}}_2$. It can be shown* that

$$E(\hat{\mathbf{C}}_1) = \mathbf{X}_1 \; \underline{\theta}_1 \tag{15}$$

and hence

$$E(\mathbf{C}_1 - \hat{\mathbf{C}}_1) = \mathbf{X}_1 \; \underline{\theta}_1 - \mathbf{X}_1 \; \underline{\theta}_1 = \mathbf{O} \tag{16}$$

Even though we find that the expected values of the residuals for a particular model are zero this does not imply that the model is the true model. In the example above we actually used more terms ($\mathbf{X}_2 \; \underline{\theta}_2$) in the trial model than was necessary to represent the true system. Physically this could mean that we are accounting for a certain component in the wrong manner. Actually in most cases $\underline{\theta}_2$ would be estimated to be zero or nearly zero if Equation (13) were fitted. One method to prevent the excessive addition of parameters is to start with the simplest model first. This was discussed in the previous paper (*2*).

NOTATION

b_i = coefficients of Equation (1)

$\hat{b}_i$ = estimated coefficients of b_i

C_1 = diagnostic parameter

$\hat{C}_1$ = predicted value of the diagnostic parameter

$\mathbf{C}_1$ = $n \times 1$ vector of the values of the diagnostic parameter

$\hat{\mathbf{C}}_1$ = $n \times 1$ vector of the predicted value of the diagnostic parameter

$\mathbf{I}$ = $n \times n$ identity matrix with 1's down the diagonal and 0's elsewhere

n = number of observations

x_1, x_2, x_3, x_4 = initial partial pressures of methane, oxygen, carbon dioxide, and water, respectively

$\mathbf{x}_i$ = $n \times 1$ vector of the n values of component i

$\mathbf{X}_i$ = matrix associated with the components

ϵ = experimental error

$\underline{\epsilon}$ = $n \times 1$ error vector

$\underline{\theta}_i$ = vector associated with the coefficients of components

$\hat{\underline{\theta}}_i$ = estimated value of vector θ_i

$\mathbf{1}$ = $n \times 1$ vector of 1's

$\mathbf{O}$ = $n \times 1$ vector of 0's

LITERATURE CITED

1. Hunter, W. G., and Reiji Mezaki, *A.I.Ch.E. J.*, **10**, 315 (1964).
2. Kittrell, J. R., W. G. Hunter, and Reiji Mezaki, *ibid.*, **12**, 1014 (1966).

* The derivation has been deposited as document 9407 with the American Documentation Service, Library of Congress, Washington, D. C. 20402, and may be obtained for $1.25 for photoprints or 35-mm. microfilm.