Application of Multiresponse Estimation to a Wetted Wall Column Model

Derrick K. Rollins
Dept. of Chemical Engineering, Dept. of Statistics, Iowa State University, Ames, IA 50011

James F. Davis
Dept. of Chemical Engineering, Ohio State University, Columbus, OH 43210

Jennifer Walker
Dept. of Statistics, Iowa State University, Ames, IA 50011

Chemical engineers frequently try to describe the physical and chemical behavior of processes using mathematical models. These models frequently contain unknown coefficients called parameters. Often these models, due to either phenomenological considerations or heuristic reasoning (such as artificial neural networks), are nonlinear in the unknown parameters. Model complexities are increasing as computer and measurement technologies advance, allowing scientists and engineers to measure and analyze simultaneously several physical quantities (called responses) that share parameter dependencies.

Parameters are typically estimated from experimental data by the minimization of a criterion called the *objective function*. For complex models, the objective functions are very elaborate, and estimation of model parameters can be a very time-consuming step in model development.

Least squares is one of the most, if not the most, widely used objective functions for parameter estimation. However, when observations on several response variables (characteristics of process performance) are taken for the same set of experimental conditions (runs), two critical error term assumptions required by least squares may not be true (Box et al., 1973). These are constant variances and independence between responses. For multiresponse data of this type, the multiresponse model described by Box and Draper (1965) is not limited by these two assumptions. It should be noted that weighted least squares, which minimizes a weighted sums of squares, can be used to account for differences in error variances. However, no version of least squares addresses correlation among response variables in the sense of statistical inference (Ziegel and Gorman, 1980).

Despite the advantages of multiresponse modeling over least squares, multiresponse estimation (ME) has not gained wide acceptance and use. Historically, this inconsistency can be attributed to the lack of "special-purpose algorithms" to minimize the objective function (Bates and Watts, 1985) as well as to a lack of understanding of the advantages over least squares just described.

Not too long ago, Stewart et al. (1992) discussed in detail some theoretical issues of ME, various properties of four ME algorithms, used an example from chemical kinetics to demonstrate estimation strategy using their algorithm, and reviewed ME applications to several other chemical engineering problems that have appeared in the literature since the late 1960s. The main purposes of this article are to extend the review of Stewart et al. (1992) by comparing the performance of two algorithms, DMRCVG (Bates and Watts, 1984) and GREG (Sorensen, 1982; Caracotsios, 1986; Stewart et al., 1992), and to demonstrate model assessment in this context using residual plots.

In order to accomplish these objectives a suitable process and model had to be selected. We chose the three-response, fractionation (wetted wall column), model of Tung et al. (1986) that followed the work of Davis et al. (1984). This model was selected not only because of familiarity and the availability of the data, but also because we could define a physical parameter with an assumed known value that we wanted to confirm by estimation. Therefore, one expected result was the confirmation of this value. It should be understood that the Tung et al. model served as an example to accomplish the objectives stated earlier (compare the algorithms and illustrate model assessment). Our analysis of this model should be considered as a first-step evaluation, and in this note, our evaluation stops here. Hence, any recommendations that we make to specifically improve this model or others of similar type will not be followed through in this note. Obtaining an adequate multiresponse model is not the main objective of the study, which is usually a several-step iterative process. It should also be noted that the modeling objective of Tung et al. (1986) did not consist of conformation to the multiresponse modeling assumptions but an ade-

Correspondence concerning this article should be addressed to D. K. Rollins.

quate fit in terms of closeness to the experimental data. Thus, our evaluation of their model meeting the assumptions of the multiresponse model should not be taken as inadequate modeling on their part. One of our objectives in evaluating their model against the multiresponse model assumptions was to provide further insight to model improvement, which is an important reason that we encourage the use of this technique.

In this study we found both algorithms to be efficient and effective multiresponse parameter estimation tools and appear to be excellent candidates for estimation problems involving complex nonlinear multiresponse models. Thus, we highly recommend their use. In terms of multiresponse model adequacy the fractionation model was judged inadequate because of severe departure from error term assumptions. However, vapor temperature showed much promise as an acceptable modeling response over vapor composition and wall temperature. Therefore, in the current state of the Tung et al. model, for statistical inference, we recommend that only the single-response estimation results for vapor temperature be used. From the insight that this study provided, it appears that it may be possible to generally improve fractionation models by allowing each measured variable along the length of the column to be a different variable, for example, to treat each vapor temperature at each measured location as an individual variable. Multiresponse modeling allows this recommendation since it allows variables on the same run to be correlated. Finally, this study also suggests that one might be able to obtain the correct residual pattern for fractionation models by empirically modeling a variable (which could be the wetted ratio) as a function of column position.

Multiresponse Model

The general multiresponse model is given as (Box and Draper, 1965):

$$Y_{ui} = F_{ui}(X_u, \theta) + \epsilon_{ui},$$
$$u = 1, \ldots, N \quad \text{and} \quad i = 1, \ldots, K. \qquad (1)$$

The error-term assumptions are given by Eqs. 2 to 5 below:

$$E[\epsilon_{ui}] = 0 \qquad \text{all } i,u \qquad (2)$$
$$E[\epsilon_{ui}\epsilon_{vj}] = 0 \qquad \text{all } i,j,u \neq v \qquad (3)$$
$$E[\epsilon_{ui}\epsilon_{uj}] = \sigma_{ij} \quad i \neq j \qquad \text{all } u \qquad (4)$$
$$E[\epsilon_{ui}^2] = \sigma_{ii} \qquad \text{all } u. \qquad (5)$$

The error-term assumptions can be described as follows. First, Eq. 2 states that the expected value of the error terms for all responses is zero (the errors are unbiased). Second, Eq. 3 states that error terms for responses on different runs cannot be correlated (that is, they must be independent). Third, Eq. 4 states that the error terms for different responses on the same run can be correlated (that is, they can be dependent). Finally, Eq. 5 states that each response can have a different error-term variance.

There are two main differences between the multiresponse model and the least-squares model. First, least squares requires $E[\epsilon_{ui}\epsilon_{uj}] = 0$ for $i \neq j$ *and* all u ($\sigma_{ij} = 0$ in Eq. 4 for all u). Second, least squares requires $\sigma_{ii} = \sigma_{jj}$ for all i, j in contrast to Eq. 5 where each response can have a different error-term variance.

Wetted Wall Column Model

The physical process that Davis et al. (1984) (see their article for details) modeled was a laboratory fractionation column with a thin liquid falling film in laminar flow countercurrent to a rising vapor in turbulent flow. An extension of the Davis et al. model (see Molloy, 1986) that includes the wetted ratio (θ) (Tung et al., 1986) is given below:

$$\frac{dV}{dz} = (N_1 + N_2)\theta \qquad (6)$$

$$\frac{dL}{dz} = -\frac{dV}{dz} \qquad (7)$$

$$\frac{d(Vy)}{dz} = N_1\theta = \left[K_g(y_i - y) + \frac{dV}{dz}y_i\right]\theta \qquad (8)$$

$$\frac{d(Lx)}{dz} = -\frac{d(Vy)}{dz} \qquad (9)$$

$$\frac{d(VH)}{dz} = [H_g(T_i - T)]\theta \qquad (10)$$

$$\frac{d(Lh)}{dz} = -\frac{d(VH)}{dz} + Q\theta \qquad (11)$$

$$\frac{dT_w}{dz} = \frac{U_L(T_w - T_i)}{MC_p}\theta. \qquad (12)$$

Along the length of the column (z), Davis et al. collected data on y, T, and T_w. Hence, in this study z is considered to be an explanatory variable (an element of X_u in Eq. 1), and y, T, and T_w are considered to be response variables (Y's in Eq. 1). Equations 6 to 12 are solved by integrating from the top of the column. Physical properties were determined empirically as described in Davis et al.

For the wetted wall column (WWC) model the number of responses is small, but the fitted response calculations are computationally intensive because of the coupled and highly nonlinear nature of this model. Since closed-form solutions are not known, the equations must be numerically integrated (details are given in Davis et al., 1984; Molloy, 1986).

Results

This section first compares the performance of the algorithms and then gives the estimation results. ME was performed first because, if successful, there would be no need for single-response estimation (SRE) to aid in determining model inadequacies. Since the ME results were not acceptable, single-response analysis follows this discussion.

For the WWC model, DMRCVG and GREG performed about the same. Table 1 is a comparison of results from the multiresponse and the three single-response analyses. The only difference worth noting is the failure of DMRCVG to converge in the multiresponse case. However, it is important

Table 1. Comparison of Estimation Results for GREG vs. DMRCVG

Algorithm	Variable	Runs	No. of Iter.	$\hat{\theta}$
GREG	y, T_w, T	23–26,33–38	4	1.040 ± 0.049
DMRCVB	y, T_w, T	23–26,33–38	4	1.051*
GREG	y	23–26,33–38	2	1.173 ± 0.024
DMRCVG	y	23–26,33–38	2	1.173
GREG	T_w	23–26,33–38	3	0.857 ± 0.039
DMRCVG	T_w	23–26,33–38	4	0.857
GREG	T	23–26,33–38	3	0.986 ± 0.018
DMRCVG	T	23–26,33–38	3	0.986

*This case did not reach the optimum. For the columns, the initial guess of θ is 1.4 in each case, $\hat{\theta}$ is the value of the parameter at the optimum, and 95% confidence intervals are given for θ (GREG results only).

to point out that the final value of the parameter estimate (1.051) was still very close to the optimum (1.040). In the multiresponse case, based on the approximate 95% confidence interval determined by GREG, $\theta = 1.0$ seems plausible. However, before accepting this conclusion, residual plots will be examined to evaluate model adequacy based on error-term assumptions. Note that residual plots support the error-term assumptions when the residuals are uniformly spread about zero in a random pattern (see Neter et al., 1983).

As shown by the residual plots for the ME case in Figures 1 to 3, the error term assumptions of unbiased errors and constant error variances appear to be violated. For all three cases (vapor composition, vapor temperature, and wall temperature) the model appears to be overpredicting since the residuals are more negative than positive. In addition, the residuals for wall temperature (Figure 3) appear to become more negative with increasing z until $z = 0.737$ m, and then they become more positive as z increases beyond 0.737 m. This convex pattern is an indication that at least one very important effect was omitted in the model. Since the predicted values are greater than the observed values for negative residuals, the column could have been losing heat near $z = 0.737$ m. If so, this loss was unaccounted for by the model. For vapor composition and vapor temperature, Figures 1 and 2, respectively, the variances of the error terms appear to

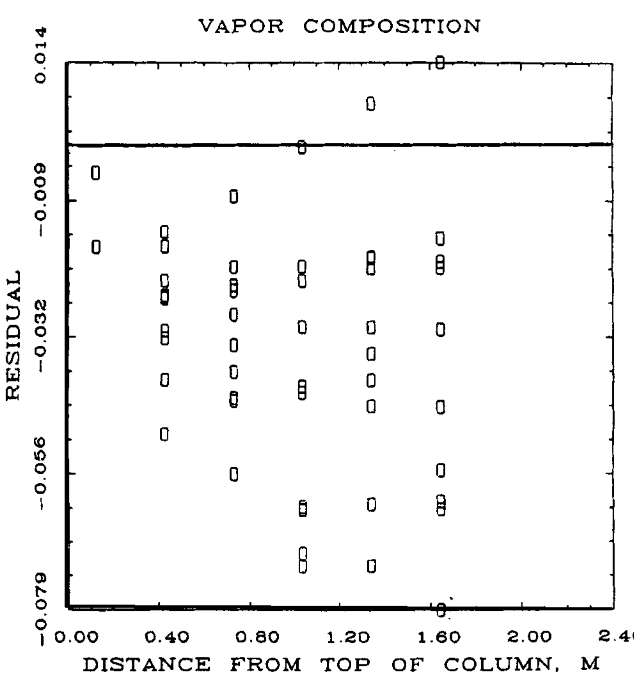

Figure 1. Temperature, concentration, and velocity in a falling film and adjacent vapor phase.
From Molloy (1986).

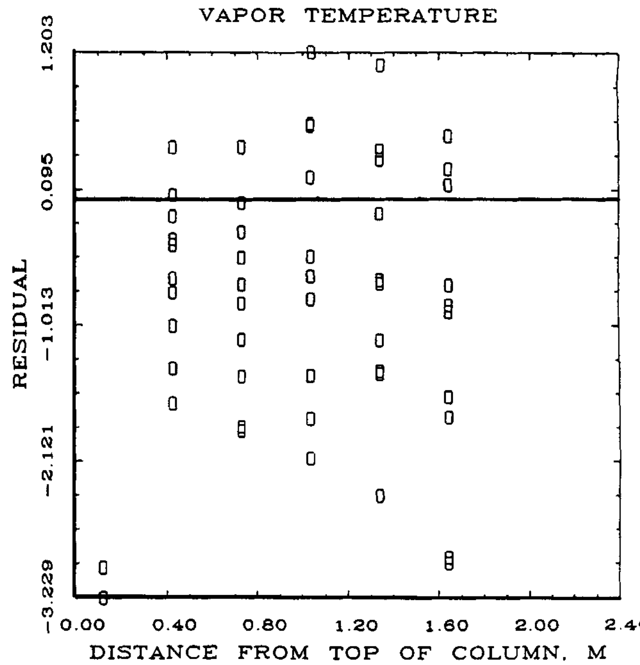

Figure 2. Vapor composition residual vs. distances from the top of the column for multiresponse case.

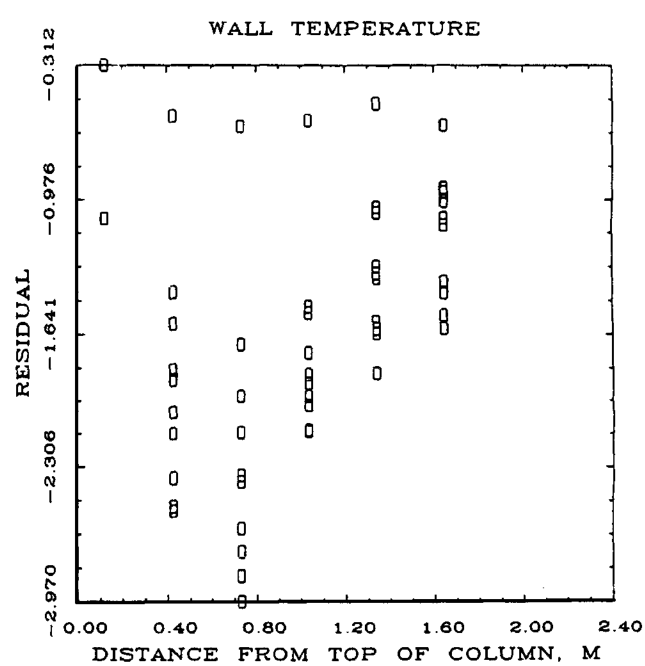

Figure 3. Vapor temperature residual vs. distances from the top of the column for multiresponse case.

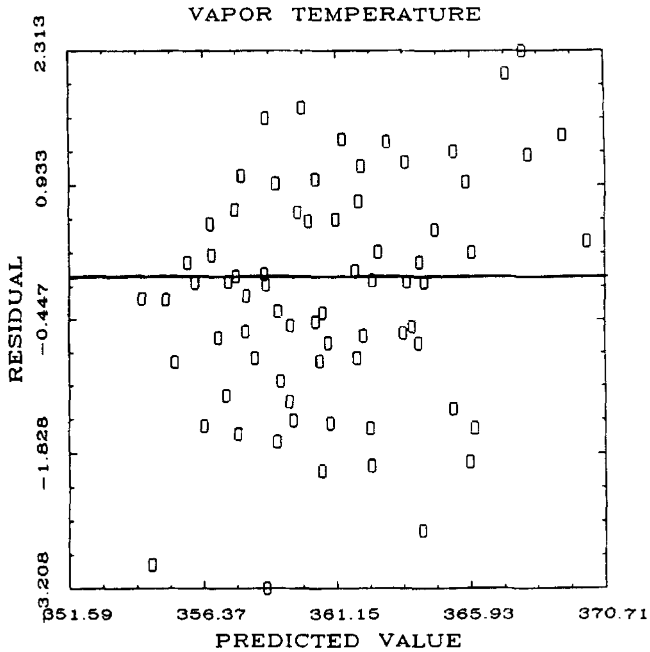

Figure 4. Vapor temperature residual vs. vapor temperature predicted value for single-response case.

increase with increasing distance from the top of the column (z). In contrast, Figure 3 shows just the opposite effect for wall temperature (the variance appears to be decreasing with increasing z). Other residual plots, which are not shown in this article for space considerations (Rollins, 1987), support these conclusions.

Since the ME results gave unacceptable residual plots, we decided to investigate SRE performance and use their results to assist in locating weaknesses in the model. In this study of the WWC model, none of the SRE analyses are completely acceptable. Except for the plot of the residuals of T vs. predicted T (Figure 4), all the single-response residual plots (not shown for space considerations) look very similar to the ones in the multiresponse case. Notice that in Figure 4, the biased error condition for vapor temperature (since the residuals are almost equally below and above zero) for this type of residual plot appears to have been eliminated (compared with Figure 2). (Also note that Figure 4 is very close to an ideal example of how a residual plot should look to support the error-term assumptions.) However, although the best of the three, we could not completely accept the fit for vapor temperature because a plot of vapor temperature residuals vs. z (not shown) showed violation of constant variance.

One way to possibly eliminate the heteroscedasticity (nonconstant variance of errors) problem observed earlier would be to consider each measured variable at each location to be a different measured variable. Thus, rather than three responses, there would be 24 responses (three variables times eight locations). Heteroscedasticity, as a function of z, would not be a concern because the multiresponse model allows different variances for different responses. The drawback with this recommendation is that at least 25 runs would be needed (the number of runs has to be greater than the number of responses; see Bates and Watts, 1988), which is much greater

Table 2. Parameter Estimation Results for Single-Response–Single-Position Cases*

Variable	z	$\hat{\theta}$
y	1.0	1.255
y	1.3	1.164
y	1.6	1.144
T_w	0.1	0.762
T_w	0.7	0.695
T_w	1.3	0.785
T_w	2.0	0.960
T	0.1	0.688
T	0.7	0.850
T	1.3	0.996
T	2.0	1.008

*$\hat{\theta}$ is the estimated value of the parameter at the optimum and z is in meters.

than used in this study. In addition, this recommendation does not address the biased error problems, which are solvable only by changing the mathematical model.

To provide further insight into possible deficiencies in the WWC model, we ran estimation analyses at fixed column positions. The single-response estimated values of θ using T and T_w are given in Table 2 by response type and location. Cases for y are not given because, at low values of z, optimal estimates could not be found before the algorithms tried to perform an illegal math operation. As Table 2 reveals, the estimated values of θ using both T and T_w support the conclusion that the column appears to be nearly completely wet at the bottom (θ is close to 1), with the wetness decreasing as you move upward (θ is about 0.7 at $z = 1.956$). This is, of course, only true if θ is truly the wetted ratio coefficient.

Examination of residual plots for the cases in Table 2 showed that T performed exceptionally well (an example plot is given in Figure 5 for $z = 1.3$ and is similar to the other values of z). In contrast, the T_w residual plots (not shown due to space constraints) showed systematic patterns, and thus, indicated a lack of fit. Based on these results it appears that some variable, which could likely be the wetted ratio, is changing with z and that the WWC model could be improved by incorporating this functional relationship into the model. For example, rather than a constant θ, one could use $\theta = \theta_0 + \theta_1 z$, or some other higher order relationship, where θ_0 and θ_1 are parameters that could be determined by ME. Ultimately, if the goal is an adequate multiresponse model, one should explore this recommendation simultaneously with other model changes to achieve this goal. Since the scope of our study only consisted of a first-step evaluation, we are not able to evaluate this recommendation in terms of this goal. However, we invite others who are attempting to develop multiresponse models of this type to explore the validity of this recommendation.

Although we cannot go into details because of space requirements, we will comment that the residuals plots were helpful in finding two unrelated mistakes in the data by showing extreme observations called *outliers*. Outliers are points that lie far beyond the scatter of the other residuals (Neter et al., 1983). Sometimes outliers are correct data and just represent large random errors or inadequacies in the model, but often they are incorrect data that can be cor-

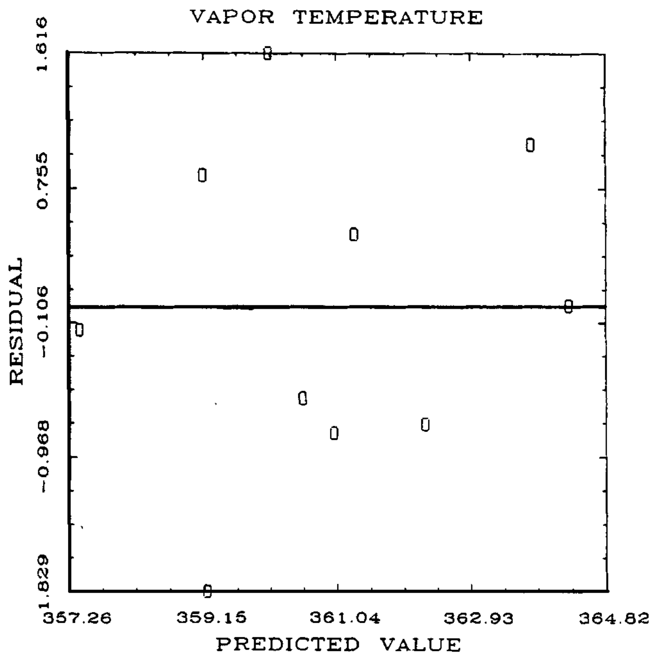

Figure 5. Vapor temperature residual vs. vapor temperature predicted value for single-response case: z = 1.346.

rected when traced back to the origin of the mistake(s). We are emphasizing the latter point because outlier analysis to find "bad" data is often not used. In this study, both mistakes were caused by incorrect computer data entries. One involved incorrect entry for a complete run (one line of data) and the other involved incorrect entry for one input of a run (one value on a line of data).

Concluding Remarks

The DMRCVG and GREG ME algorithms appear to be very valuable parameter estimation tools, and their use is highly recommended. Since GREG internally addresses over-parameterization and -response dependencies, it will probably provide the most complete study with the minimum amount of effort. For parameter estimation problems under rectangular constraints (such as $0 < \theta < 1$), GREG is recommended since DMRCVG does not address constrained estimation. Another reason to prefer GREG is that it generates important summary information. Users of DMRCVG must generate all their own summary information. Also, all users will have to generate their own graphical results when using either algorithm. However, if the algorithms are made available to users via a library file, adding some of the commonly used plots as part of the output options is recommended.

Multiresponse modeling provided insight on two ways to possibly improve the WWC model. The first one was to let the expected value of each response variable be different at each location. This could possibly eliminate the nonconstant variance problem but would require more runs to estimate parameters. The second one was to include a variable that varies as a function of column position, and estimate this relationship using ME.

Given the present state of the WWC model, vapor temperature conforms best to the ME model assumptions, and therefore has the most reliable estimate of θ. In addition, it appears that the only way to improve the ME fits of vapor composition and wall temperature is by making significant changes in the mathematical modeling equations.

Notation

C_p = heat capacity of the cooling water, J/(kg·K)
i = the response variable number
F = vector of expected responses
h = molar enthalpy of the liquid phase averaged over the bulk flow with the standard state taken at the interface, J/kmol
H = molar enthalpy of the vapor phase averaged over the bulk flow with the standard state taken at the interface, J/kmol
K = number of responses
K_G = vapor phase mass-transfer coefficient corrected for finite mass-transfer rate through the interface, kmol/(s·m^2)
L = liquid molar flow rate per unit length of wetted perimeter, kmol/(s·m)
M = cooling water flow rate per unit width, kmol/(s·m)
N_j = interfacial molar flux of from the liquid phase to vapor phase, for component j
P = vector of predicted responses
Q = wall heat flux, W/m^2
R = vector of residuals
Y_{ui} = the ith observed response for run u
T_i = interface temperature, K
T_w = wall temperature, K
u = the run number
U_L = liquid heat-transfer coefficient, W/(m·K)
V = the vapor molar flow rate per unit length of wetted perimeter
x = mole fraction of the more volatile component in the liquid phase
X_u = the vectors of operating variables for run u
y = mole fraction of the more volatile component in the vapor phase
y_i = vapor mole fraction at the interface
Y = vector of the observed responses

Greek letters

ϵ_{ui} = the error associated with response i for run u
θ = vector of true parameters
$\hat{\theta}$ = vector of estimated parameters, also used for the wetted ratio
σ_{ii} = the variance of the error for response i
σ_{ij} = the covariance of the error for responses i and j

Literature Cited

Bates, D. M., and D. G. Watts, *Nonlinear Regression Analysis and Its Applications*, Wiley, New York (1988).
Bates, D. M., and D. G. Watts, "Multiresponse Estimation with Special Application to Systems of Linear of Linear Differential Equations," *Technometrics*, **27**, 329 (1985).
Bates, D. M., and D. G. Watts, "A Multi-response Gauss-Newton Algorithm," *Commun. Stat.-Simula. Comput.*, **13**(5), 705 (1984).
Box, G. E. P., W. G. Hunter, J. F. McGregor, and J. Erjavec, "Some Problems Associated with the Analysis of Multiresponse Models," *Technometrics*, **15**, 33 (1973).
Box, G. E. P., and N. R. Draper, "The Bayesian Estimation of Common Parameters from Several Responses," *Biometrika*, **52**, 355 (1965).
Caracotsios, M., "Model Parametric Sensitivity Analysis and Nonlinear Parameter Estimation, Theory and Applications," PhD Diss., Univ. of Wisconsin-Madison (1986).
Davis, J. F., H. H. Tung, and R. S. H. Mah, "Fractionation with Condensation and Evaporation in Wetted-Wall Columns," *AIChE J.*, **30**, 328 (1984).
Molloy, B. L., "A Design for Fractionating Condensers," PhD Thesis,

University of Honors Program, College of Engineering, The Ohio State Univ.-Columbus (1986).
Neter, J., W. Wasserman, and M. H. Kutner, *Applied Linear Regression Models*, Irwin, Homewood, IL (1983).
Rollins, D. K., "Model Assessment and Multiresponse Estimation for a Fractionation Model With Coupled Ordinary Differential Equations," MS Thesis, The Ohio State Univ.-Columbus (1987).
Sorensen, J. P., "Simulation, Regression and Control of Chemical Reactors by Collocation Techniques," PhD Thesis, Danmarks tekniske Hojskole, Lyngby, Denmark (1982).
Stewart, W. E., M. Caracotsios, and J. P. Sorensen, "Parameter Estimation from Multiresponse Data," *AIChE J.*, **38**(5), 641 (1992).
Tung, H. H., J. F. Davis, and R. S. H. Mah, "Fractionating Condensation and Evaporation in Plate-Fin Devices," *AIChE J.*, **32**(7), 1116 (1986).
Ziegel, E. R., and J. W. Gorman, "Kinetic Modelling with Multiresponse Data," *Technometrics*, **22**(2), 139 (1980).

Manuscript received Oct. 22, 1993, and revision received Nov. 2, 1994.