Robust Error-in-Variables Estimation Using Nonlinear Programming Techniques

For systems described by algebraic or differential equation models where all variables are subject to error, the error-in-variables method (EVM) for parameter estimation has been shown to be superior to standard least-squares techniques. Previous EVM algorithms were developed assuming linear (or linearized) model equations. Unfortunately, many chemical engineering processes operate in strongly nonlinear regions where linear approximations may be inaccurate. In this paper, new algorithms using nonlinear programming techniques for the error-in-variables methods are proposed. In addition, a method for discerning when these methods are necessary is discussed. The proposed algorithms are compared to the least-squares method and traditional error-in-variable approaches. Improved parameter estimates for several steady-state nonlinear processes are demonstrated.

In-Won Kim
Michael J. Liebman
Thomas F. Edgar
Department of Chemical Engineering
University of Texas
Austin, TX 78712

Introduction

Parameter estimation is often a key step in the verification and subsequent use of a mathematical model in many fields of science and engineering. In chemical engineering problems it is frequently necessary to estimate the parameters in nonlinear algebraic or differential models where all variables are subject to error. In such problems there is no statistical distinction between independent and dependent variables. In this case the best estimates of the parameters can be obtained from a formalized method that treats the errors associated with all experimental observations.

In a classical regression approach the measurements of the independent variables are assumed to be free of error, while the observations of the dependent variables, the responses of the system, are subject to errors. However, in some engineering problems observations of the independent variables also contain errors. Kendall and Stuart (1973) and Moran (1971) have shown that errors in observations of the independent variables of a simple regression model can cause the least-squares estimator to be biased.

In the error-in-variables method (EVM), measurement errors in all variables are treated in the calculation of regression coefficients. EVM provides both parameter estimates and reconciled data estimates that are consistent with respect to the model. Many systems can be described by a model of the form

$$f(z, \theta) = \mathbf{0} \tag{1}$$

where

f = vector of k functions that comprise the model
z = vector of n measured variables
θ = vector of p parameters

Deming (1943) originally formulated the general problem of parameter estimation in models, by taking into account the errors in all measured variables. He was primarily concerned with obtaining approximate solutions because an exact solution was impossible at that time. Subsequently, exact solutions were proposed by other investigators, but until the beginning of the 1970s these methods were concerned only with fitting data by straight lines or polynomials (York, 1966; Williams, 1968; O'Neil et al., 1969; Southwell, 1969).

Britt and Luecke (1973) presented general algorithms for EVM; the objective function was optimized using Lagrange multipliers and the constraints were successively linearized with respect to the parameters and the measured variables. In this approach, both the parameter estimates and the reconciled measurements were obtained simultaneously. The methods of Peneloux et al. (1976) and Reilly and Patino-Leal (1981) differ from the previous methods in one important respect: for a given

Correspondence concerning this paper should be addressed to T. F. Edgar.

set of parameters, θ, the constraint equations are linearized or successively linearized about the measured variables, z. The data reconciliation routine (variable correction step) for z is nested within the parameter estimation routine. The main advantage of this method is that explicit calculation of the reconciled measurements is unnecessary and the size of estimation problem is dramatically reduced.

The latest advances for EVM have been made by Schwetlick and Tiller (1985) and Valko and Vajda (1987). The basic algorithms are substantially the same as those presented by Reilly and Patino-Leal. The only difference is the separation of the parameter estimation step and the data reconciliation step, resulting in a two-stage calculation. The constraints are linearized successively as done by Reilly and Patino-Leal.

All previously proposed EVM algorithms use linear approximations of the model equations in order to reduce computational effort. On many systems, such techniques should perform adequately. Unfortunately, these efficient, statistically based estimation methods may not be capable of handling the more severe nonlinearities present in many chemical engineering problems. In addition, the above methods do not provide a means for incorporating bounds or nonlinear constraints on the estimate values or handle differentially constrained models. Historically, application of the previous methods has been limited to vapor–liquid equilibrium systems. Recently these methods have been applied to an adsorption equilibrium model (Higher and Danner, 1986) and a single-stage flash (MacDonald and Howat, 1988).

The data reconciliation portion of the EVM calculation is especially sensitive to linearization errors. Liebman and Edgar (1988) have demonstrated markedly improved reconciliation estimates using nonlinear programming (NLP).

The objective of this work is to investigate the use of NLP techniques to perform the data reconciliation (variable correction) calculations as opposed to the use of successive linear approximation. Any improvement in the data reconciliation step should result in more reliable parameter estimates.

Measures of nonlinearity for EVM

When the model is nonlinear, a number of results that are true for the linear case no longer apply. A measure of the "amount of nonlinearity" in nonlinear problems is needed to quantify the nonlinear behavior. Such a measure helps determine when linearized results provide acceptable approximations to the nonlinear functions.

Two measures of nonlinearity were developed by Bates and Watts (1980, 1981): intrinsic nonlinearity and parameter-effects nonlinearity. The intrinsic nonlinearity (IN) is a measure of nonplanarity of the expectation (response) surface, which does not change with parameterizations (transformations of parameter). The parameter-effects nonlinearity (PE) is a measure of the nonuniformity of the parameter lines on the tangent plane. High intrinsic nonlinearities mean that the planar assumption (linear approximation of parameter lines) is no longer valid. The projections of the parameter curves on the tangent plane would be uniform with small PE. Large curvature means that inference based on the linear approximation is potentially problematic.

Bates and Watts' parameter-effects and intrinsic nonlinearities can be used to indicate when linearization methods may produce less than satisfactory results. Unfortunately, their derivation is restricted to problems where the optimal solution can be written as an explicit function of θ. In order to use these measures of nonlinearity on the EVM, an explicit solution formulation for the optimal z needs to be obtained. One such form is the analytical solution to the linearly constrained weighted least-squares data reconciliation problem (Kuehn and Davidson, 1961) evaluated at the optimal z and θ:

$$\eta = z^* - VJ^T(JVJ^T)^{-1}[Jz^* - f(z^*, \theta^*)] \tag{2}$$

where

z^* = vector of optimal estimates obtained through EVM
η = vector of optimal estimates obtained from a linearized solution based at z^*
J = partial derivatives of constraint equations with respect to z evaluated at z^*
V = covariance matrix

Using the above explicit formulation, approximate derivatives can be calculated. Specifically, $\partial\eta/\partial\theta$ and $\partial^2_\eta/\partial\theta\partial\theta^T$ can be approximated using central finite-differencing. Specific details regarding the calculation of PE and IN can be found in Bates and Watts. Theoretically, PE and IN measures should be computed at each iteration of the linearization solution. In this work, however, the nonlinearity measures were only calculated at the optimum in order to minimize computational effort.

If we accept a curvature deviation from the tangent plane of less than 15% (Bates and Watts, 1988), then we will declare an analysis as having unacceptable curvature if

$$\text{PE}^* \sqrt{\text{F}(p, n-p: 1-\alpha)} > 0.30$$

or

$$\text{IN}^* \sqrt{\text{F}(p, n-p: 1-\alpha)} > 0.30$$

where p is the number of parameters, n is the number of variables and $\text{F}(p, n-p: 1-\alpha)$ is the appropriate value of the F statistic.

Derivation of Nonlinear Error-in-Variables Estimation Algorithms

An implicit model comprised of k equations can be described as

$$f(\zeta, \theta) = \mathbf{0} \tag{3}$$

where $f = (f_1, \ldots, f_k)^T$ represents the equations that connect the variables $\zeta = (\zeta, \ldots, \zeta_n)^T$ and parameters $\theta = (\theta_1, \ldots, \theta_p)^T$, where no distinction is made between independent and dependent variables.

In the experiment we observed the measurement values, z, of the variables and allow for errors in all of them. Thus,

$$z_i = \zeta_i + \epsilon_i \qquad i = 1, \ldots, m \tag{4}$$

Assume that the error ϵ_i has zero mean and a known positive definite covariance matrix V_i that can be estimated from replicate measurements. In practice the covariance matrix is usually assigned on the basis of the accuracy of the measurement devices.

Table 1. Parameter Estimates for X-ray Image Model

Parameter	Starting Value	Nested		Two-stage		SPEDR
		SL*	NLP	SL**	NLP	
Θ_1	−0.57	−0.9994	−0.9992	−0.9996	−0.9994	−0.5432
Θ_2	−3.4	−2.9240	−2.9315	−2.9308	−2.9311	−3.3863
Θ_3	0.1	0.0876	0.0876	0.0876	0.0876	0.1174
Θ_4	5.7×10^{-4}	0.0162	0.0162	0.0162	0.0162	0.0037
Θ_5	0.082	0.0800	0.0798	0.0798	0.0798	0.1064
CPU, s***		79.8	313.2	32.5	39.3	39.0

*Estimates from Reilly and Patino-Leal (1981)
**Estimates from Valko and Vajda (1987)
***Computing time in seconds on a DEC VAXstation II, Fortran 77

According to Eq. 4, $\epsilon_i = z_i - \zeta_i$, hence it is natural to estimate the parameters θ in Eq. 3 by minimizing the error norm function with respect to $\hat{z}$ and $\hat{\theta}$

$$\Psi(\hat{z}_1, \ldots, \hat{z}_m, \hat{\theta}) = \sum_{i=1}^{m} (z_i - \hat{z}_i)^T V_i^{-1} (z_i - \hat{z}_i)$$
$$\text{subject to} \quad f(\hat{z}_i, \hat{\theta}) = \mathbf{0}, \qquad i = 1, 2, \ldots, m \quad (5)$$

The following sections describe several new algorithms proposed for efficiently solving the above optimization problem.

Simultaneous parameter estimation and data reconciliation

The most straightforward approach for solving the nonlinear EVM problems is to use nonlinear programming to simultaneously estimate θ and z. In the traditional weighted least-squares (WLS) parameter estimation formulation there are only p optimization variables corresponding to the number of unknown parameters. In contrast, the simultaneous parameter estimation and data reconciliation (SPEDR) formulation for EVM has $nm + p$ optimization variables, as shown in Eq. 5, so its dimensionality increases directly with the number of observations m and the number of variables n. The sheer size of such an NLP problem precludes the direct use of NLP algorithms except for small values of n and m. In this paper, Lasdon's generalized reduced gradient code, GRG2 (Lasdon and Waren, 1986), was used to optimize $\hat{z}$ and $\hat{\theta}$ simultaneously.

Two-stage nonlinear EVM

The second proposed approach involves separating the parameter estimation and data reconciliation steps, as done by Valko and Vajda. Specifically, the two-stage algorithm can be written as:

1. At $j = 1$, $\hat{z} = z$ and $\theta = \theta_o$ where z are the experimental measurements
2. Find the minimum of the function for $\hat{\theta}$

$$\Psi_1 = \min_{\theta} \sum_{i=1}^{m} (\xi_i - z_i)^T V_i^{-1} (\xi_i - z_i)$$

subject to

$$\xi_i = [I - VA_i^T (A_i V A_i^T)^{-1} A_i] z_i + [VA_i^T (A_i V A_i^T)^{-1}] b_i \quad (6)$$

where $A_i z_i = b_i$ is the linear approximation to $f(z_i, \hat{\theta}) = 0$ at $z_i = \hat{z}_i$. ζ_i's represent an approximation to the optimal solution of the data reconciliation problem.

3. If $j > 1$ and $\|\hat{\theta}^j - \hat{\theta}^{j+1}\| < \epsilon$ then stop. Otherwise, proceed to step 4.
4. At fixed $\hat{\theta}^{j+1}$, solve the data reconciliation problem.

$$\Psi_2 = \min_{\hat{z}} \sum_{i=1}^{m} (\hat{z}_i - z_i)^T V_i^{-1} (\hat{z}_i - z_i)$$
$$\text{subject to} \quad f(\hat{z}_i, \hat{\theta}) = \mathbf{0}, \qquad i = 1, 2, \ldots, m \quad (7)$$

5. Return to step 2.

The NLP in step 4 is actually solved as a set of decoupled NLPs. The difference between this algorithm and Valko and Vajda's algorithm is that the successive linearization solution is replaced with the nonlinear programming problem in step 4. Through the use of NLP in solving the data reconciliation problem any inaccuracies in the linearization approximation will be overcome. The main advantages of this method were its

Table 2. Parameter Estimates for Vapor–Liquid Equilibrium Model

Parameter	Starting Value	Nested		Two-stage		SPEDR
		SL	NLP	SL	NLP	
A/R	323.15	617.79	617.79	617.80	617.80	618.65
B/R	323.15	519.68	519.68	519.68	519.68	520.17
PE $\sqrt{F}$*		0.17	0.17	0.17	0.17	0.17
IN $\sqrt{F}$**		0.04	0.04	0.04	0.04	0.04
CPU, s		17.6	78.9	5.1	6.3	17.7

*Parameter-effects nonlinearity
**Intrinsic nonlinearity

Table 3. Parameter Estimates for Hyperbolic Model I

	Nested		Two-stage			
Parameter	SL	NLP	SL	NLP	SPEDR	WLS*
Θ	1.320	127.81	127.78	127.78	127.76	0.508
SSRES**	7907.6	719.31	719.32	719.31	719.32	—
PE $\sqrt{F}$	3.76	0.35	0.34	0.34	0.42	—
IN $\sqrt{F}$	4.84	0.22	0.25	0.22	0.37	—
CPU, s	16.1	88.2	10.8	15.5	28.0	0.33

*x variable assumed to be an independent variable
**Sum of squared residuals

ability to handle nonlinear constraints and its decomposition approach, yielding smaller optimization subproblems. In this work, a successive quadratic programming code, GREG, by Caracotsios (1986) was used to solve the parameter estimation problem in step 2, while GRG2 was used to solve the data reconciliation problem in step 4.

Nested nonlinear EVM

The final proposed nonlinear EVM approach is similar to that of Reilly and Patino-Leal. As in the two-stage EVM described above, the parameter estimation is decoupled from the data reconciliation problem. In this formulation, however, the data reconciliation problem is optimized at each iteration of the parameter estimation problem. The steps of the algorithm are

1. At $j = 1, \hat{z} = z$ and $\hat{\theta} = \theta_o$
2. Find the minimum of the function for $\hat{\theta}$ and $\hat{z}$

$$\Psi_1 = \min_{\hat{\theta}} \sum_{i=1}^{m} (\hat{z}_i - z_i)^T V_i^{-1} (\hat{z}_i - z_i) \tag{8}$$

subject to

$$\Psi_2 = \min_{\hat{z}} \sum_{i=1}^{m} (\hat{z}_i - z_i)^T V_i^{-1} (\hat{z}_i - z_i) \tag{9}$$

subject to $f(\hat{z}_i, \hat{\theta}) = \mathbf{0}, \quad i = 1, 2, \ldots, m$

The difference between this algorithm and Patino-Leal's algorithm is that the successive linearization solution is replaced with the nonlinear programming problem in Eq 9. As in the two-stage EVM method, the nested NLP is solved as a set of decoupled NLPs and the size of the largest optimization problem is reduced to the order of n. GREG and GRG2 were used to solve Eq. 8 and 9 respectively.

Test Problems

The three proposed algorithms were tested on several test problems discussed in the literature by other investigators. Comparisons are made in computer time and the converged values of both $\hat{z}$ and $\hat{\theta}$ for different methods.

X-ray image model

The nonlinear x-ray image model and data discussed by Reilly and Patino-Leal (1981) were used to compare the proposed algorithms with previous work. The data were fitted using an elliptical model of the form

$$(z_1 - \theta_1, z_2 - \theta_2) \begin{bmatrix} \theta_3 & \theta_4 \\ \theta_4 & \theta_5 \end{bmatrix} \begin{bmatrix} z_1 - \theta_1 \\ z_2 - \theta_2 \end{bmatrix} = 1.0 \tag{10}$$

Both z_1 and z_2 are independent and have known variances of 0.0001. All five methods discussed previously were used to obtain parameter estimates for $\theta_1, \theta_2, \theta_3, \theta_4$, and θ_5. The resulting estimates, shown in Table 1, were the same for all methods except for simultaneous parameter estimation and data reconciliation, which converged to a local minimum. For other starting points, the SPEDR approach provided the same estimates as other approaches. The two-stage successive linearization (SL) method proposed by Valko and Vajda was the fastest computationally, followed closely by the two-stage NLP approach. The nested methods and the simultaneous parameter estimation and data reconciliation approach required significantly more computational effort.

Table 4. Parameter Estimates for Hyperbolic Model II

	Nested		Two-stage			
Parameter	SL	NLP	SL	NLP	SPEDR	WLS*
Θ	127.78	127.78	127.78	127.78	127.77	0.451
SSRES**	719.32	719.32	719.32	719.32	719.32	—
PE $\sqrt{F}$	0.0	0.0	0.0	0.0	0.0	—
IN $\sqrt{F}$	0.0	0.0	0.0	0.0	0.0	—
CPU, s	32.86	102.82	12.66	20.60	17.71	0.39

*x variable assumed to be an independent variable
**Sum of squared residuals

Vapor-liquid equilibrium model

Evaluation of parameters for nonlinear thermodynamic models, especially vapor–liquid and gas–solid equilibrium models, has been of great interest in the chemical engineering literature in recent years, with many regression techniques used (Fabries and Renon, 1975; Anderson et al., 1978; Cholinski et al., 1981). The vapor–liquid equilibrium model presented by Valko and Vajda was used to compare the algorithms. The data consist of five experiments and four measured variables (P, T, x, y). Van Laar equations with two parameters (A and B) were used to fit the observed data. The model equations are

$$\gamma_1 x_1 p_1^0(T) - y_1 P = 0 \tag{11}$$

$$\gamma_2(1 - x_1) p_2^0(T) - (1 - y_1)P = 0 \tag{12}$$

$$p_1^0(T) = \exp\left[23.4803 - \frac{3{,}626.55}{T - 34.29}\right] \tag{13}$$

$$p_2^0(T) = \exp\left[21.0692 - \frac{2{,}927.17}{T - 50.22}\right] \tag{14}$$

$$\gamma_1 = \exp\left[\frac{A}{RT}\left(1 + \frac{A}{B}\frac{x_1}{1 - x_1}\right)^{-2}\right] \tag{15}$$

$$\gamma_2 = \exp\left[\frac{B}{RT}\left(1 + \frac{B}{A}\frac{1 - x_1}{x_1}\right)^{-2}\right] \tag{16}$$

The variances of the four measured variables were chosen in accordance with Cholinski et al. (1981):

$$\sigma_{x_i}^2 = 2.5 \times 10^{-5} \qquad \sigma_{y_i}^2 = 2.25 \times 10^{-4}$$
$$\sigma_P^2 = 1 \times 10^{-4}\, Pa^2 \qquad \sigma_T^2 = 1 \times 10^{-2} K^2$$

Equilibrium data for the binary system of methanol and 1,2-dichloroethane were taken from Reid et al. (1977).

The resulting parameter estimates are shown in Table 2. All of the EVM methods provided similar estimates. It is clear from the PE and IN measures of nonlinearity that there is only a slight deviation from linearity at the optimal parameter estimates.

Hyperbolic models

A hyperbolic model, $x^2y^2 = \theta$, discussed by Britt and Luecke (1973) was chosen to compare alternative methods for nonlinear EVM. Ten measurement sets of x and y were taken from Britt and Luecke. To avoid singularity of the measurement matrix, x_1 was changed from 0.0 to 0.1. An initial guess of $\theta_o = 1.0$ was used, along with constant variances of $\sigma_x = \sigma_y = 0.1$.

Two types of equation forms were tested to examine the effects of model formulation:

Model I

$$y = \sqrt{\theta}/x \tag{17}$$

Model II

$$x^2y^2 = \theta \tag{18}$$

The resulting parameter estimates are summarized in Table 3 for model I and in Table 4 for model II. The PE and intrinsic nonlinearity (IN) for model II were degenerate, since the second derivative of model II with respect to θ is 0. In this case, the stated criteria for acceptance of the linear approximation cannot be used. The estimate presented by Britt and Luecke was 127.78.

The successive linearization method for the data reconciliation step in the nested EVM of Reilly and Patino-Leal failed to provide optimal estimates for model I, but succeeded in finding optimal estimates for model II, as shown in Table 4. This dependence of estimate quality on model form indicates that even for the simple, one-parameter model, successive linearization methods may fail. For this particular example, the application of successive linearization gives rise to chaotic behavior (see Appendix), and convergence is never achieved.

Although not apparent in the results in Tables 3 and 4, the two-stage successive linearization EVM algorithm can suffer similar problems. In order to test the robustness of the SL method in two-stage EVM, measurement data for x and y were simulated for $\theta = 0.5$. Parameter estimates were obtained using various levels of measurement noise and model I. In all cases, an initial guess of $\theta_o = 0.001$ was used.

Table 5. Estimates for Hyperbolic Model I with Simulated Data

σ_x	σ_y	Two-stage SL	Two-stage NLP
0.6	0.2	0.528	0.528
0.6	0.3	0.478	0.478
0.7	0.1	failed	0.589
0.7	0.2	failed	0.550
0.7	0.3	failed	0.496
0.7	0.5	failed	0.411
1.0	0.5	failed	0.463

Table 6. Physical Data for CSTR Model

$\tau = 100$ s	$C_1 = 5{,}000\ s^{-1}$
$Q_1 = 83{,}600$ J/mol	$\Delta H_r = -4{,}180$ J/mol
$\rho = 1.0$ g/l	$C_p = 4.18$ J/giK

Table 7. Simulated Noise-free Measurement Data for CSTR Model

No.	T_o K	T K	A_o mol/L	A mol/L	B mol/L
1	547.10	665.09	1.0000	0.8820	0.1180
2	531.70	675.09	1.0000	0.8566	0.1434
3	512.84	685.09	1.0000	0.8277	0.1723
4	490.59	695.09	1.0000	0.7955	0.2045
5	465.15	705.09	1.0000	0.7601	0.2399
6	436.88	715.09	1.0000	0.7218	0.2782
7	406.28	725.09	1.0000	0.6812	0.3188
8	373.95	735.09	1.0000	0.6389	0.3611
9	340.56	745.09	1.0000	0.5955	0.4045
10	306.84	755.09	1.0000	0.5517	0.4483

Table 8a. Parameter Estimates for CSTR Model

Case 1: $\sigma_T = \sigma_{To} = 1$; $\sigma_{Ao} = \sigma_A = \sigma_B = 0.01$; $\Theta_1 = 0.01717 s^{-1}$; $\Theta_2 = 12.58$

	Nested		Two-stage			
Parameter	SL	NLP	SL	NLP	SPEDR	WLS*
Θ_1	3.243×10^{-3}	0.463	1.688×10^{-2}	1.689×10^{-2}	3.904×10^{-2}	1.669×10^{-2}
Θ_2	0.451	1.982×10^{-2}	12.489	12.490	3.722	12.352
SSRES	32,105	107,033	34.7	33.0	81,498	—
PE $\sqrt{F}$	**	**	0.09	0.08	**	—
IN $\sqrt{F}$	**	**	0.06	0.05	**	—
CPU, s	395.0	90.4	60.4	77.9	73.64	1.04

*T_o and A_o assumed to be error-free
**Larger than 1.0×10^6

Table 8b. Parameter Estimates for CSTR Model

Case 2: $\sigma_T = \sigma_{To} = 5$; $\sigma_{Ao} = \sigma_A = \sigma_B = 0.05$; $\Theta_1 = 0.01717 s^{-1}$; $\Theta_2 = 12.58$

	Nested		Two-stage			
Parameter	SL	NLP	SL	NLP	SPEDR	WLS*
Θ_1	2.750×10^{-4}	0.160	1.569×10^{-2}	1.569×10^{-2}	1.570×10^{-2}	1.453×10^{-2}
Θ_2	0.451	3.286×10^{-2}	12.074	12.074	12.073	11.149
SSRES	1,413.6	43.6	34.6	34.6	34.6	—
PE $\sqrt{F}$	**	**	0.42	0.42	0.47	—
IN $\sqrt{F}$	**	**	1.36	0.94	1.00	—
CPU, s	386.0	122.98	57.31	69.04	133.91	1.05

*T_o and A_o assumed to be error-free
**Larger than 1.0×10^6

Table 8c. Parameter Estimates for CSTR Model

Case 3: $\sigma_T = \sigma_{To} = 10$; $\sigma_{Ao} = \sigma_A = \sigma_B = 0.1$; $\Theta_1 = 0.01717 s^{-1}$; $\Theta_2 = 12.58$

	Nested		Two-stage			
Parameter	SL	NLP	SL	NLP	SPEDR	WLS*
Θ_1	2.414×10^{-3}	4.045×10^{-2}	1.417×10^{-2}	1.417×10^{-2}	1.431×10^{-2}	1.246×10^{-2}
Θ_2	0.451	0.265	11.457	11.457	11.529	9.874
SSRES	393.8	1,001	34.4	34.4	34.4	—
PE $\sqrt{F}$	**	**	0.67	0.79	0.99	—
IN $\sqrt{F}$	**	**	4.25	3.11	3.65	—
CPU, s	391.7	537.1	59.8	75.3	107.7	1.02

*T_o and A_o assumed to be error-free
**Larger than 1.0×10^6

As shown in Table 5, the two-stage SL EVM did not converge to a value near the true parameter value for σ_x greater than 0.6, while using two-stage NLP EVM provided a reliable estimate for all noise levels. The parameter estimation failed due to a lack of convergence in the successive linearization data reconciliation step (see Appendix).

As expected, the weighted least-squares approach gave biased estimates for both model forms. These results were coincident with relatively large values of the parameter-effects nonlinearity (PE) for hyperbolic model I. The hyperbolic model and data set differ significantly from linearity, so techniques incorporating linearization were unable to converge to the true minimum.

CSTR model

The proposed methods were evaluated using simulated measurements for a common chemical engineering system: a steady-state, adiabatic CSTR with an irreversible, first-order reaction.

$$A \xrightarrow{k_1} B \tag{19}$$

The simulation for this system involved five simulated measurements (inlet temperature, outlet temperature, inlet concentration of A, outlet concentrations of A and B), two parameters

(c_1 and Q_1) and three algebraic constraints derived from a differential equation model for the system. The steady-state model for the irreversible reaction system is

$$\frac{1}{\tau}(A_o - A) - k_1 A = 0 \tag{20}$$

$$\frac{1}{\tau}(B_o - B) + k_1 A = 0 \tag{21}$$

$$\frac{1}{\tau}(T_o - T) + \frac{-\Delta H_r}{\rho C_p}(k_1 A) = 0 \tag{22}$$

The reaction rate constant was expressed as

$$k_1 = c_1 \exp\left(\frac{-Q_1}{RT}\right) \tag{23}$$

Transformed parameters were used, resulting in rate equations of the form

$$k_1 = c_1 \exp\left(\frac{-Q_1}{RT_r}\right) \exp\left[\frac{-Q_1}{RT_r}\left(\frac{T_r}{T} - 1\right)\right]$$

$$= \theta_1 \exp\left[-\theta_2\left(\frac{T_r}{T} - 1\right)\right] \tag{24}$$

with

$$\theta_1 = c_1 \exp\left(\frac{-Q_1}{RT_r}\right)$$

$$\theta_2 = \frac{Q_1}{RT_r}$$

where T_r is a reference temperature. The estimation of θ_1 and θ_2 should be more robust than for c_1 and Q_1 since the derivatives of Eq. 23 with respect to the parameters tend to be collinear, while the derivatives of Eq. 24 with respect to θ_1 and θ_2 are more nearly orthogonal (Bates and Watts, 1988; Ratkowsky, 1983).

Physical property data and simulated noise-free measurements are given in Tables 6 and 7. The ten measurement sets were simulated using $\theta_1 = 0.01717\ s^{-1}$ and $\theta_2 = 12.58$. Three different noise levels were chosen to investigate the effects of noise on the parameter estimates obtained by different methods. Initial guesses for the parameters were $\theta_1 = 10.0$ and $\theta_2 = 0.1$. The resulting estimates are summarized in Tables 8a–8c.

For this problem, only the two-stage EVM algorithms were able to find the optimal estimates for all three noise levels. The nested EVM algorithm using successive linearization and NLP failed with the initial guess, $\theta_o = (10, 0.1)^T$. But, at the smaller noise levels, the nested algorithm worked with initial guesses which were not far from the true values. Strangely, the simultaneous parameter estimation and data reconciliation approach gave suboptimal estimates for small noise levels, but gave the correct estimates for the largest noise level. SPEDR was sensitive to initial guesses. As expected, traditional weighted least-squares parameter estimation yielded biased estimates at the largest noise level.

To qualitatively examine the quality of fit for the resulting parameter estimates, plots of outlet A and B concentrations vs. outlet temperature were generated for the intermediate noise level. Figures 1 and 2 show the fit obtained using the two-stage EVM using NLP. The reconciled data generally follow along the curve representing the "true" values.

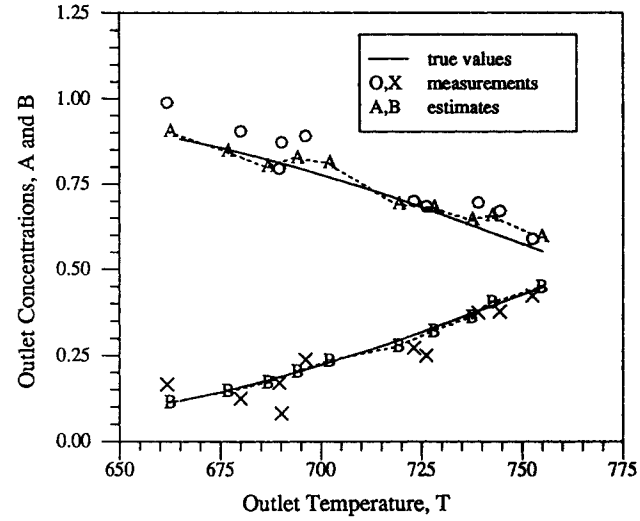

Figure 1. Reconciled data of outlet concentrations by two-stage EVM using NLP for case 2 of CSTR example.

Conclusions

In this paper, three new methods for nonlinear error-in-variables parameter estimation were investigated and compared to current techniques. The proposed methods were simultaneous parameter estimation and data reconciliation, two-stage error-in-variables using NLP, and nested error-in-variables using NLP. Of these, the two-stage nonlinear EVM using NLP outperformed all other techniques in terms of robustness and computa-

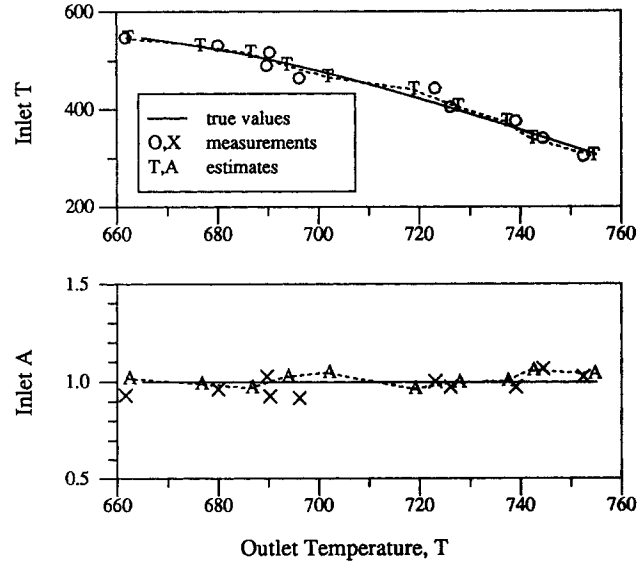

Figure 2. Reconciled data of inlet temperature and concentration by two-stage EVM using NLP for case 2 of CSTR example.

tional efficiency. Previous researchers have demonstrated the advantages of using EVM over classical WLS regression.

The four test problems illustrated the robustness of the proposed two-stage EVM using NLP. While other methods failed on one or more of the problems, the proposed algorithm succeeded in finding optimal parameter estimates for all problems in an efficient manner. The advantages of this method include the ability to handle nonlinear constraints (both equality and inequality) and the decomposition of the full SPEDR formulation into more manageable subproblems. The two-stage algorithm can easily be applied to chemical engineering problems.

We can conclude that for data sets in highly nonlinear regions of a model, the two-stage EVM using NLP is the only recommended solution method. In addition, the parameter effects and intrinsic nonlinearity measures of nonlinearity appear to be good indicators for whether linear approximations are valid, except in degenerate cases. Such an indicator can be utilized in determining whether a nonlinear programming EVM solution needs to be attempted.

Notation

A = outlet reactant concentration, mol/L
A = linear approximation coefficient matrix, Eq. 6
A_o = inlet reactant concentration, mol/L
b = linear approximation righthand side vector
B = outlet product concentration, mol/L
B_o = inlet product concentration, mol/L
c_1 = preexponential constant, s^{-1}
C_p = heat capacity, J/g · K
ΔH_r = heat of reaction, J/mol
f = vector of m functions that comprise the model
J = partial derivatives of constraint equations with respect to z evaluated at z^*
k = number of model equations
k_1 = reaction rate constant, s^{-1}
m = number of independent sets of measurements
n = number of variables in each set
p = number of parameters
P = pressure, Pa
p_i^o = pure component vapor pressure, Pa
Q_1 = activation energy, J/mol
R = gas constant, 8.314 J/mol · K
T = outlet temperature, K
T_o = inlet temperature, K
V = covariance matrix
x_i = liquid mole fraction
y_i = vapor mole fraction
z = vector of n measured variables
$\hat{z}$ = vector of estimates obtained through EVM
z^* = vector of optimal estimates obtained through EVM

Greek letters

γ_i = activity coefficient
η = vector of optimal estimates obtained from a linearized solution based at z^*, Eq. 2
θ = vector of p parameters
θ_g = vector of initial guess of θ
$\hat{\theta}$ = vector of estimates of θ
θ^* = vector of optimal estimates obtained through EVM
ρ = density, g/L
τ = time constant, s
ξ = vector of estimates obtained from linearized solution, Eq. 6
ζ = vector of n true variables

Acronyms

EVM = error-in-variables method
IN = intrinsic nonlinearity
NLP = nonlinear programming
PE = parameter-effects nonlinearity
SPEDR = simultaneous parameter estimation and data reconciliation
WLS = weighted least-squares

Literature Cited

Anderson, T. F., D. S. Abrams, and E. A. Grens, "Evaluation of Parameters for Nonlinear Thermodynamic Models," *AIChE J.*, **24**(1), 20 (1978).

Bates, D. M., and D. G. Watts, "Relative Curvature Measures of Nonlinearity," *J. Roy. Statist. Soc.*, **B42**(1), 1 (1980).

———, "Parameter Transformations for Improved Approximate Confidence Regions in Nonlinear Least Squares," *Annals Statistics*, **9**(6) 1152 (1981).

———, *Nonlinear Regression Analysis and Its Application*, Wiley, New York (1988).

Britt, H. I., and R. H. Luecke, "The Estimation of Parameters in Nonlinear, Implicit Models," *Technometrics*, **15**, 233 (1973).

Caracotsios, M. "Model Parameteric Sensitivity Analysis and Nonlinear Parameter Estimation. Theory and Applications," Ph.D. Diss., Univ. Wisconsin–Madison (1986).

Cholinski, J., M. Palczewska-Tulinska, A. Szefranska, and D. Wyrzykowskastankiewicz, "A New Method of Parameter Adjustment and Diagnostic Checks on Gamma Models used in Vapor–Liquid Equilibrium Calculations," *Chem. Eng. Sci.*, **36**, 173 (1981).

Deming, W. E., *Statistical Adjustment of Data*, Wiley, New York (1943).

Fabries, J. F., and H. Renon, "Method of Evaluation and Reduction of Vapor–Liquid Equilibrium Data of Binary Mixtures," *AIChE J.*, **21**(4), 735 (1975).

Finlayson, B. A., *Nonlinear Analysis in Chemical Engineering*, McGraw-Hill, New York (1980).

High, M. S., and R. P. Danner, "Treatment of Gas–Solid Adsorption Data by the Error-in-Variables Method," *AIChE J.*, **32**(7), 1138 (1986).

Kendall, M. G., and A. Stuart, *The Advanced Theory of Statistics*, Griffin, London, **2**(1973).

Kuehn, D. R., and H. Davidson, "Computer Control," *Chem. Eng. Prog.*, **57**(6), 44 (1961).

Lasdon, L. S., and A. D. Waren, *GRG2 User's Guide*, Dept. Computer Info. Sci., Cleveland State Univ. (1986).

Liebman, M. J., and T. F. Edgar, "Data Reconciliation for Nonlinear Processes," Preprint, AIChE Ann. Meet., Washington, DC (1988).

MacDonald, R. J., and C. S. Howat, "Data Reconciliation and Parameter Estimation in Plant Performance Analysis," *AIChE J.*, **34**(1), 1 (1988).

Moran, P. A. P., "Estimating Structural and Functional Relationships," *J. Multivariate Analysis*, **1**, 232 (1971).

O'Neil, M., I. G. Sinclair, and F. J. Smith. "Polynomial Curve Fitting when Abscissas and Ordinates Are Both Subject to Error," *Computer J.*, **12**, 52 (1969).

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

Ratkowsky, D. A., *Nonlinear Regression Modeling: A Unified Practical Approach*, Dekker, New York (1983).

Reid, R. C., J. M. Prausnitz, and J. K. Sherwood, *The Properties of Gases and Liquids*, McGraw-Hill, New York, 3ed. (1977).

Reilly, P. M., and H. Patino-Leal, "A Bayesian Study of the Error-in-Variables Model," *Technometrics*, **23**(3), 221 (1981).

Schwetlick, H., and V. Tiller, "Numerical Methods for Estimating Parameters in Nonlinear Models with Error in the Variables," *Technometrics*, **27**(1), 17 (1985).

Seydel, R., *From Equilibrium to Chaos, Practical Bifurcation and Stability Analysis*, Elsevier, New York (1988).

Southwell, W. H., "Fitting Experimental Data," *J. Comput. Phys.* **4**, 465, (1969).
Valko, P., and S. Vajda, "An Extended Marquardt-type Procedure for Fitting Error-in-Variables Models," *Comput. Chem. Eng.*, **11**(1), 37 (1987).
Williams, J. H., "Least-Squares Fitting of a Straight Line," *Can. J. Physics*, **46**, 1845 (1968).
York, P. "Least-Squares Fitting of a Straight Line," *Can. J. Physics*, **44**, 1709 (1966).

Appendix: Chaotic Behavior Arising from Successive Linearization

For some examples, such as hyperbolic model I, neither convergence nor divergence occurs when successive linearization is used for data reconciliation. A close examination of the steps involved provides an explanation for the observed behavior.

The nonlinear model equation is

$$f(\hat{x}, \hat{y}) = \hat{y} - \frac{\sqrt{\theta}}{\hat{x}} = 0 \tag{A1}$$

Using a first-order Taylor series approximation at (x_o, y_o), the equation can be written as

$$f(\hat{x}, \hat{y}) \simeq y_0 - \frac{\sqrt{\theta}}{x_0} + \frac{\sqrt{\theta}}{x_0^2}(\hat{x} - x_0) + (\hat{y} - y_0) = 0 \tag{A2}$$

Then the linearized data reconciliation problem becomes

$$\min_{\hat{x}\ \hat{y}} (\hat{x} - x)^2 + (\hat{y} - y)^2 \tag{A3}$$

subject to

$$\hat{y} = -\frac{\sqrt{\theta}}{x_0^2}\hat{x} + \frac{2\sqrt{\theta}}{x_0} \tag{A4}$$

Substitution of Eq. A4 into Eq. A3 yields

$$\min_{\hat{x}} F = (\hat{x} - x)^2 + \left(-\frac{\sqrt{\theta}}{x_0^2}\hat{x} + \frac{2\sqrt{\theta}}{x_0} - y\right)^2 \tag{A5}$$

For optimality, find $\hat{x}^*$ to satisfy $dF/d\hat{x} = 0$, so

$$\hat{x}^* = \frac{x x_0^4 + 2\theta x_0 - \sqrt{\theta} y x_0^2}{x_0^4 + \theta} \tag{A6}$$

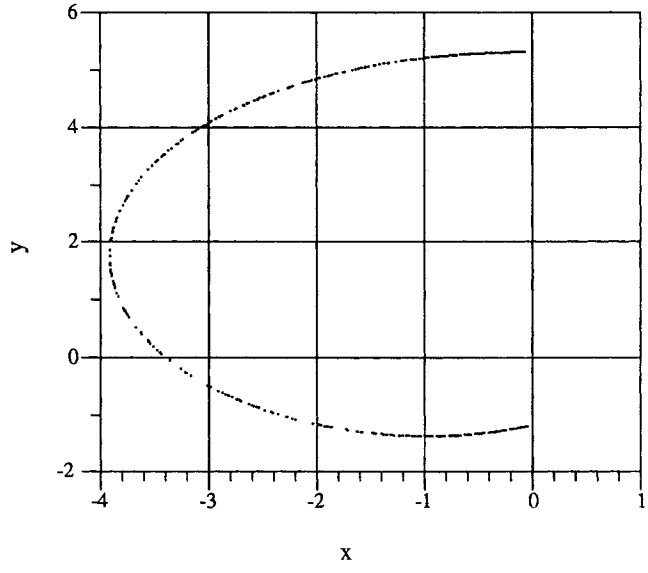

Figure A1. Optimal estimates $(\hat{x}_{k+1}^*, \hat{y}_{k+1}^*)$ calculated successively for hyperbolic model.

For successive application, this becomes

$$\hat{x}_{k+1}^* = \frac{x\hat{x}_k^4 + 2\theta\hat{x}_k - \sqrt{\theta} y \hat{x}_k^2}{\hat{x}_k^4 + \theta} \tag{A7}$$

and

$$\hat{y}_{k+1}^* = -\frac{\sqrt{\theta}}{\hat{x}_k^2}\hat{x}_{k+1}^* + \frac{2\sqrt{\theta}}{\hat{x}_k} \tag{A8}$$

For example, consider $x = 0.9$, $y = 5.4$ (the second data set of Britt and Luecke, 1973) and $\theta_o = 4.7$. Using $\hat{x} = 0.9$ and $\hat{y} = 5.4$, Eqs. A7 and A8 can be used to calculate $(\hat{x}_{k+1}^*, \hat{y}_{k+1}^*)$ by successive substitution. Figure A1 shows the result of more than 500 successive iterations. All $(\hat{x}_{k+1}^*, \hat{y}_{k+1}^*)$ are confined to lie on a C-shaped curve. Sequential iterates are randomly distributed on the curve and never converge nor diverge (Finlayson, 1980). In this case, a strange attractor (Seydel, 1988) has introduced chaotic behavior into the successive iterates.

Manuscript received Jan. 29, 1990, and revision received May 21, 1990.