ELSEVIER

Computers and Chemical Engineering 24 (2000) 2367–2383

Computers & Chemical Engineering

www.elsevier.com/locate/compchemeng

Comparison of steady state and integral dynamic data reconciliation

Miguel J. Bagajewicz *, Qiyou Jiang

School of Chemical Engineering and Material Science, University of Oklahoma, 100 East Boyd Street, T-335, Norman, OK 73019-0628, USA

Received 16 November 1999; received in revised form 27 March 2000; accepted 28 March 2000

Abstract

This paper is devoted to the comparison of the performance of integral approach to dynamic data reconciliation and steady state data reconciliation. It is shown that in the absence of biases and leaks, the performance of both approaches is similar. Moreover, it is proven that once the appropriate variance is chosen, both methods are identical in the absence of accumulation terms. Finally, an analysis is made on how large the discrepancies are when there are accumulation terms. © 2000 Elsevier Science Ltd. All rights reserved.

Keywords: Integral approach; Dynamic data reconciliation; Steady state data reconciliation

1. Introduction

Steady state data reconciliation in process plants was first proposed more than 35 years ago (Kuehn & Davidson, 1961). This seminal paper started a new area of research in process systems engineering that later changed the paradigm of plant monitoring and production accounting. Data reconciliation has been used to adjust the measurements in a process so that they comply with conservation laws. In assuming that the system is at steady state, all process variability was lumped into one single averaged value for each stream. Thus data reconciliation using the steady state assumption uses these averaged values together with some variance matrix. The variance of the measurements comes from the intrinsic nature of the variance of each measurement, and from other covariances that arise from the fact that the instruments are attached to a single power source. In addition, it has been proposed that process variability should also be included in such variance to capture the process dynamics. Although a few papers have addressed the issue of the determination of the variance using field data, it is not known well how effective these methods are in practice. The reason why this is such an elusive issue is because averaged values include process variability. Thus, the variance of these values has contributions from the instruments and from the process unsteady state.

The present commercially available technology is based on steady-state data reconciliation. This technology has been successful in improving the accuracy of production accounting. To use this software, it is common practice to use straightforward averages of several measurements of the different plant variables taken throughout a certain period of time. We know of no study investigating the effect of this averaging practice in the accuracy of the resulting estimates. Basically, the system fluctuations are lumped with instrument biases and inherent instrument white noise. Thus, under the assumption that averaged measurements can be used in the context of steady state data reconciliation, several methods have been proposed to perform the detection and elimination of biases and leaks.

* Corresponding author. Tel.: +1-405-3255811; fax: +1-405-3255813.
E-mail address: bagajewi@mailhost.ecn.uoknor.edu (M.J. Bagajewicz).

PII: S0098-1354(00)00498-1

Dynamic data reconciliation models were first presented by Stanley and Mah (1977), who adapted Kalman filtering in a quasi steady state condition; Darouach and Zasadzinski (1991) used a backwards difference approximation and recursive technique to solve the constrained least squares optimization problem. Rollins and Devanathan (1993) improved on the estimation accuracy using a maximum likelihood function and proposing two estimators that are later averaged. Narasimhan and Mah (1988) proposed to apply the generalized likelihood ratio (GLR) to dynamic situations with small departures from steady state values. Kao, Tamhane and Mah (1992) studied the effect of serially correlated data on gross error detection. They proposed composite test procedures based on window averages, pre-whitening procedures and the generalized likelihood ratio. Other methods have been presented for dynamic data reconciliation: Ramamurthi, Sistu and Bequette (1993) proposed a technique based on a successively linearized horizon, Liebman, Edgar and Lasdon (1992) used orthogonal collocation. Karjala and Himmelblau (1996) rely on neural networks and Albuquerque and Biegler (1996) on a discretization of the system of DAE using Runge Kutta methods. Finally, Bagajewicz and Jiang (1997) proposed an integral method for dynamic data reconciliation. As the name describes, the method relies on the integration of the process differential equations, a polynomial representation in time and data reconciliation based on adjusting the polynomial coefficients.

We know of no study where a careful comparison of steady state and dynamic data reconciliation is performed. Practitioners, especially software vendors, claim that dynamic data reconciliation is computationally too intensive and the effort implementing it in practice will not produce a meaningful improvement over the current steady state technology. In addition, the power of dynamic data reconciliation and gross error detection methods has not been thoroughly evaluated and concern exists that it might be not better than those based on steady state techniques. To clarify this issue, this paper presents a comparison between the classical steady state reconciliation and the integral dynamic data reconciliation. Dynamic data, subject to averaging and steady state reconciliation will be compared with the average of the results of dynamic data reconciliation and the averages of the true values. This paper focuses on a comparison of the integral approach to dynamic data reconciliation (Bagajewicz & Jiang, 1997, 1998) and the aforementioned steady state techniques to perform data reconciliation.

The issue of the proper use of steady state data reconciliation in the context of the current practice of data averaging is explored first. Then, a review of the integral dynamic data reconciliation method is presented. A theoretical comparison between the two is presented next, followed by a discussion on variance determination. Examples are then presented.

2. Average-based steady state reconciliation

Let the system of equations for material balances be represented by

$$\frac{\mathrm{d}v}{\mathrm{d}t} = Af \quad (1)$$

$$Cf = 0 \quad (2)$$

where v and f are the tank hold-ups and flows, respectively. Let f_i^+, $i = 1, \ldots, n$ be the vector of measurements of all stream flows for n instances of time and v_i^+ be the vector of tank hold-up measurements for the same times.

Steady state data reconciliation, as performed by several commercial packages (Datacon from Simsci and Sigmafine from OSI, among others) is based on

1. The average of the flow measurements $\bar{f}^+$
2. A new artificial flowrate f_i^{v+} defined as follows:

$$f_i^{v+} = \frac{v_i^+ - v_0^+}{(t_i - t_0)} \quad (3)$$

3. The average of the artificial flowrates $\bar{f}^{v+}$

Thus the system equations are now represented by:

$$Dy = 0 \quad (4)$$

where

$$D = \begin{bmatrix} A & -I \\ C & 0 \end{bmatrix} \quad (5)$$

and

$$y=\begin{pmatrix} f \\ f^{v} \end{pmatrix} \tag{6}$$

Let R be the variance of all measurements. If all the flows are measured, then the solution to the problem is:

$$\tilde{y}=[I-RD^{T}(DRD^{T})^{-1}D]\bar{y}^{+} \tag{7}$$

where

$$\bar{y}^{+}=\begin{pmatrix} \bar{f}^{+} \\ \bar{f}^{v+} \end{pmatrix}.$$

This solution is obtained by posing a problem as a least-squares minimization, assumption that has been proposed by the seminal paper by Kuehn and Davidson (1961) and substantiated by the maximum likelihood concepts (see Mah, 1990), using Bayesian theory (Johnston & Kramer, 1995) and information theory (Crowe, 1996).

When not all flows or tank hold-ups are measured, then matrix projection (Crowe, García Campos & Hrymak, 1983) or transformation of D to a canonical form (Madron, 1992) have been proposed to identify redundant measurements. These transformations are well known and are omitted here.

3. Integral model for dynamic linear reconciliation

The integral model has been presented in a recent paper (Bagajewicz & Jiang, 1997). After reducing D into its canonical form (Madron, 1992) the problem is reduced to

$$B_{R}\frac{dv_{R}}{dt}=A_{R}f_{R} \tag{8}$$

$$C_{R}f_{R}=0 \tag{9}$$

The model uses redundant measurements only. For this reason, subscript R is added. The result of cooptation is that linear combinations of hold-up variations need to be used (Bagajewicz & Jiang, 1997), hence matrix B_R. The following polynomial representation is proposed next:

$$f_{R}\approx\sum_{k=0}^{s}\alpha_{k}^{R}t^{k} \tag{10}$$

$$v_{R}\approx v_{R0}+\sum_{k=0}^{s}\omega_{k+1}^{R}t^{k+1} \tag{11}$$

where s is the polynomial order chosen.

Assuming that there are n measurements with normally distributed errors, a maximum likelihood formulation leads to a least squares problem given by:

$$\left.\begin{aligned}
&Min\sum_{i=0}^{n}\{(v_{Ri}-v_{Ri}^{+})^{T}S_{V}^{-1}(v_{Ri}-v_{Ri}^{+})+(f_{Ri}-f_{Ri}^{+})^{T}S_{F}^{-1}(f_{Ri}-f_{Ri}^{+})\}\\
&\text{s.t.}\\
&B_{R}(v_{Ri}-v_{R0})=A_{R}\int_{0}^{t}f_{R}(\xi)d\xi \quad (i=0,\ldots,n)\\
&C_{R}f_{Ri}=0 \quad (i=0,\ldots,n)
\end{aligned}\right\} \tag{12}$$

By substituting Eqs. (10) and (11) into Eq. (12), one obtains a linearly constrained quadratic problem. Thus, the procedure to obtain the solution is similar to the procedure used to obtain Eq. (7). The mathematical details are in Bagajewicz and Jiang (1997).

4. Variance in average-based steady state data reconciliation

To establish a comparison between steady state and dynamic data reconciliation in the case of time varying systems the value of variance used in the case of steady state data reconciliation should be discussed. It has been claimed in several papers (Almasy & Mah, 1984; Keller, Zasadzinski & Darouach, 1992; Chen, Bandoni & Romagnoli, 1997) that the variance used in steady state data reconciliation is different than the instrument variance. One of the main reasons stated for such a difference is the departure from steady state. Thus, before any comparison is performed a new variance needs to be determined. This has been for years the major problem for practitioners, as no well-established guidelines exist to determine this variance, except for the aforementioned few papers that recommend some strategies.

Let us assume first that the system is at steady state, i.e. the true values y are constant. Assume that the distribution of flow measurements is multivariate normal $N(y, R_F)$. Then:

1. The averages follow a normal distribution $N(y, R_F/n)$. Note that the variance of the mean is smaller as the number of measurements increases.
2. As a result of (1), $\lim_{n\to\infty} \bar{y} = y$, because the variance of the average tends to zero.

Since the system is at steady state, hold-up measurements are ignored.

The above findings lead us to the following conclusions:

> In a system that is at steady state and in the absence of biases and leaks, good estimates of the true values can be obtained from the averaging of a large number of data. Data reconciliation can improve the variance of these estimates and make them satisfy balance equations.

An additional important conclusion is also obtained from the above conclusions:

> Since the error associated with an average of measurements follows a normal distribution around the true value with variance R_F/n, steady state data reconciliation (assuming no outliers are present) should he used using a variance $R_s = R_F/n$.

Although the above conclusion looks straightforward, this has not been made very clear among practitioners.

5. Comparison between the integral dynamic model and the average-based steady state model based

The comparison procedure is as follows. Once Eq. (10) and Eq. (11) have been substituted into Eq. (12), the problem becomes:

$$\left.\begin{array}{l} \mathrm{Min}\{(Jv_{R0} + T_w\omega^R - v_R^+)^T R_V^{-1}(Jv_{R0} + T_\omega\omega^R - v_R^+) + (T_\alpha\alpha^R - f_R^+)^T R_F^{-1}(T_\alpha\alpha^R - f_R^+)\} \\ \text{s.t.} \\ D_m\omega^R = R_m\alpha^R \\ C_\alpha\alpha^R = 0 \end{array}\right\} \tag{13}$$

where

$$\alpha^R = \begin{bmatrix} \alpha_0^R \\ \alpha_1^R \\ \vdots \\ \alpha_s^R \end{bmatrix} \qquad \omega^R = \begin{bmatrix} \omega_1^R \\ \omega_2^R \\ \vdots \\ \omega_{s+1}^R \end{bmatrix} \qquad f_R^+ = \begin{bmatrix} f_{R0}^+ \\ f_{R1}^+ \\ \vdots \\ f f_{Rn}^+ \end{bmatrix} \qquad v_R^+ = \begin{bmatrix} v_{R0}^+ \\ v_{R1}^+ \\ \vdots \\ v_{Rn}^+ \end{bmatrix} \qquad J = \begin{bmatrix} I \\ I \\ \vdots \\ I \end{bmatrix}$$

$$D_m = \begin{bmatrix} B_R & 0 & \cdots & 0 \\ 0 & B_R & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & B_R \end{bmatrix} \qquad R_m = \begin{bmatrix} A_R & 0 & \cdots & 0 \\ 0 & A_R/2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_R/s+1 \end{bmatrix}$$

$$C_\alpha = \begin{bmatrix} C_R & 0 & \cdots & 0 \\ 0 & B_R & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & C_R \end{bmatrix} \qquad R_V^{-1} = \begin{bmatrix} S_V^{-1} & 0 & \cdots & 0 \\ 0 & S_V^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & S_V^{-1} \end{bmatrix}$$

$$R_F^{-1} = \begin{bmatrix} S_F^{-1} & 0 & \cdots & 0 \\ 0 & S_F^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & S_F^{-1} \end{bmatrix} \qquad T_\alpha = \begin{bmatrix} I & 0 & \cdots & 0 \\ I & t_1 I & \cdots & t_1^s I \\ \vdots & \vdots & \ddots & \vdots \\ I & t_n I & \cdots & t_n^s I \end{bmatrix} \qquad T_\omega = \begin{bmatrix} 0 & \cdots & 0 \\ t_1 I & \cdots & t_n^{s+1} I \\ \vdots & \ddots & \vdots \\ t_n I & \cdots & t_n^{s+1} I \end{bmatrix}$$

Once this problem is solved analytically (Bagajewicz & Jiang, 1997), the average of the reconciled flowrates is:

$$\bar{f}_R = \Gamma_t f \tag{14}$$

where Γ_t is an averaging (linear) operator. An average equivalent to the artificial flowrates can be obtained from

$$\overline{\frac{dv_R}{dt}} = \Gamma_t \frac{dv_R}{dt} \tag{15}$$

which will be compared with the result given by Eq. (7).

5.1. Comparison in the absence of accumulation terms

Assume there is no tank and only flowrates are involved. The data reconciliation problem at steady state becomes:

$$\text{Min}(f_R^* - \bar{f}_R^+)^T (n+1) S_F^{-1} (f_R^* - \bar{f}_R^+)$$

s.t.

$$C_R f_R{}^* = 0 \tag{16}$$

where f_R^* stands for the solution of the steady state model using the averages of the measurements. The solution is:

$$f_R{}^* = P\bar{f}_R^+ \tag{17}$$

where

$$P = [I - S_F C_R^T (C_R S_F C_R^T)^{-1} C_R] \tag{18}$$

This solution can be formalized in a similar way for a dynamic case. The problem is:

$$\left.\begin{array}{l} \text{Min} \sum_{i=0}^{n} (f_{Ri} - f_{Ri}^+)^T S_F^{-1} (f_{Ri} - f_{Ri}^+) \\ \text{s.t.} \\ C_R f_{Ri} = 0 \,\vdots\, i = 0, \ldots, n \end{array}\right\} \tag{19}$$

which has as solution the following:

$$f_{Ri} = P f_{Ri}^+ \quad i = 0, \ldots, n \tag{20}$$

Therefore we can conclude that the following holds

$$\bar{f}_R = P\bar{f}_R^+ \tag{21}$$

and

$$\bar{f}_R = f_R{}^* \tag{22}$$

Eq. (23) states that the averages of reconciled data obtained using data reconciliation are equal to the reconciled data using steady state reconciliation based on averages of dynamic measurements.

We will now show that this is also true when the integral dynamic data reconciliation model is used.

5.1.1. Proof

First note that under the no hold up assumption the solution of the integral dynamic data reconciliation problem is (See Appendix A)

$$f_R = T_\alpha[I - ZR_F C_\alpha^T (C_\alpha ZR_F C_\alpha^T)^{-1} C_\alpha] ZR_F (R_F^{-1} T_\alpha)^T f_R^+ \tag{23}$$

where Z is defined in the same appendix. However, the matrix within parenthesis can be expressed in terms of a new matrix P (Appendix B)

$$I - ZR_F(C_\alpha VR_F C_\alpha^T)^{-1} C_\alpha = \begin{bmatrix} I - S_F C_R^T (C_R S_F C_R^T)^{-1} C_r & & \\ & \ddots & \\ & & I - S_F C_R^T (C_R S_F C_R^T)^{-1} C_R \end{bmatrix}$$

$$= \begin{bmatrix} P & & \\ & \ddots & \\ & & P \end{bmatrix} \tag{24}$$

Therefore Eq. (23) becomes:

$$f_R = \begin{bmatrix} I & 0 & \cdots & 0 \\ I & I & \cdots & I \\ \vdots & \vdots & \ddots & \vdots \\ I & nI & \cdots & n^s I \end{bmatrix} \begin{bmatrix} P & & \\ & \ddots & \\ & & P \end{bmatrix} \begin{bmatrix} z_{00}I & z_{01}I & \cdots & z_{0s}I \\ z_{10}I & z_{11}I & \cdots & z_{1s}I \\ \vdots & \vdots & \ddots & \vdots \\ z_{s1}I & z_{s2}I & \cdots & z_{ss}I \end{bmatrix}$$

$$\times \begin{bmatrix} S_F & & \\ & \ddots & \\ & & S_F \end{bmatrix} \begin{bmatrix} I & I & \cdots & I \\ 0 & I & \cdots & nI \\ \vdots & \vdots & \ddots & \vdots \\ 0 & I & \cdots & n^s I \end{bmatrix} \begin{bmatrix} S_F^{-1} & & \\ & \ddots & \\ & & S_F^{-1} \end{bmatrix} f_R^+ \tag{25}$$

Thus

$$f_R = \begin{bmatrix} P & 0 & \cdots & 0 \\ P & P & \cdots & P \\ \vdots & \vdots & \ddots & \vdots \\ P & nP & \cdots & n^s P \end{bmatrix} \begin{bmatrix} z_{00}I & I\sum_{i=0}^{s} z_{0i} & \cdots & I\sum_{i=0}^{s} n^i z_{0i} \\ z_{10}I & I\sum_{i=0}^{s} z_{1i} & \cdots & I\sum_{i=0}^{s} n^i z_{1i} \\ \vdots & \vdots & \ddots & \vdots \\ z_{s0}I & I\sum_{i=0}^{s} z_{si} & \cdots & I\sum_{i=0}^{s} n^i z_{si} \end{bmatrix} f_R^+$$

$$= \begin{bmatrix} P & & \\ & \ddots & \\ & & P \end{bmatrix} \begin{bmatrix} \left(z_{00} f_{R0}^+ + \sum_{i=0}^{s} z_{0i} f_{R1}^+ + \cdots + \sum_{i=0}^{s} n^i z_{0i} f_{Rn}^+\right) I \\ \left(\sum_{j=0}^{s} z_{j0} f_{R0}^+ + \sum_{j=0}^{s}\sum_{i=0}^{s} z_{ji} f_{R1}^+ + \cdots + \sum_{j=0}^{s}\sum_{i=0}^{s} n^i z_{ji} f_{Rn}^+\right) I \\ \left(\sum_{j=0}^{s} n^j z_{j0} f_{R0}^+ + \sum_{j=0}^{s} n^j \sum_{i=0}^{s} z_{ji} f_{R1}^+ + \cdots + \sum_{j=0}^{s} n^j \sum_{i=0}^{s} n^i z_{ji} f_{Rn}^+\right) I \end{bmatrix} \tag{26}$$

which finally renders

$$\sum_{i=0}^{n} f_{Ri} = P\Bigg[\left(z_{00} + \sum_{j=0}^{s} z_{j0} + \cdots + \sum_{j=0}^{s} n^j z_{j0}\right) f_{R0}^+ + \left(\sum_{i=0}^{s} z_{0i} + \sum_{j=0}^{s}\sum_{i=0}^{s} z_{ji} + \cdots + \sum_{j=0}^{s} n^j \sum_{i=0}^{s} z_{ji}\right) f_{R1}^+ + \cdots$$
$$+ \left(\sum_{i=0}^{s} n^i z_{0i} + \sum_{j=0}^{s}\sum_{i=0}^{s} n^i z_{ji} + \cdots + \sum_{j=0}^{s} n^j \sum_{i=0}^{s} n^j z_{ji}\right) f_{Rn}^+\Bigg]$$

$$= P\left(z_{00} + \sum_{k=1}^{n}\sum_{j=0}^{s} k^j z_{j0}\right) f_{R0}^{+} + \left(\sum_{i=0}^{s} z_{0i} + \sum_{k=1}^{n}\sum_{j=0}^{s} k^j \sum_{i=0}^{s} z_{ji}\right) f_{R1}^{+} + \cdots + \left(\sum_{i=0}^{s} n^i z_{0i} + \sum_{k=1}^{n}\sum_{j=0}^{s} k^j \sum_{i=0}^{s} n^i z_{ji}\right) f_{Rn}^{+} \quad (27)$$

But:

$$z_{00} + \sum_{k=1}^{n}\sum_{j=0}^{s} k^j z_{j0} = I \quad (28)$$

$$\sum_{i=0}^{s} z_{0i} + \sum_{k=1}^{n}\sum_{j=0}^{s} k^j \sum_{i=0}^{s} z_{ji} = I \quad (29)$$

and

$$\sum_{i=0}^{s} m^i z_{0i} + \sum_{k=1}^{n}\sum_{j=0}^{s} k^j \sum_{i=0}^{s} m^i z_{ji} = I \qquad 1 < m \le n \quad (30)$$

as it is shown in Appendices C and D, and from which we conclude that

$$\sum_{i=0}^{n} f_{Ri} = P[f_{R0}^{+} + f_{R1}^{+} + \cdots + f_{Rn}^{+}] \quad (31)$$

Thus we have

$$\bar{f}_R = P\bar{f}_R^{+} \quad (32)$$

and

$$\bar{f}_R = f_R{}^* \quad (33)$$

Therefore, the assertion that the averages of reconciled values from dynamic data reconciliation are the same as the reconciled values from steady state data reconciliation based on dynamic data averages is also true when the integral approach is used.

5.2. Comparison in the presence of accumulation terms

When tanks with hold-ups are involved, assume one can obtain the artificial flowrate for hold-ups ($f^v = \mathrm{d}v/\mathrm{d}t$) and the corresponding standard deviation (S_{f^v}). Then one gets the data reconciliation problem at steady state, which is similar to Eqs. (16) and (17):

$$\left.\begin{array}{l} \min (y_R^* - \bar{y}_R^{+})^T (n+1) S_y^{-1} (y_R^* - \bar{y}_R^{+}) \\ \text{s.t.} \\ G_R y_R{}^* = 0 \end{array}\right\} \quad (34)$$

where

$$y = \begin{pmatrix} f \\ f^v \end{pmatrix}, \qquad G_R = \begin{bmatrix} A_R & -I \\ C_R & 0 \end{bmatrix}, \qquad S_y = \begin{bmatrix} S_F & \\ & S_{f^v} \end{bmatrix}$$

For a dynamic system,

$$\left.\begin{array}{l} \min \sum_{i=0}^{n} (y_{Ri} - y_{Ri}^{+})^T S_y^{-1} (y_{Ri} - y_{Ri}^{+}) \\ \text{s.t.} \\ G_R y_{Ri} = 0 \quad i = 0, \ldots, n \end{array}\right\} \quad (35)$$

Obviously, from analyzing Eqs. (34) and (35), one can get the same conclusion as in the case when tanks are absent, that is,

$$\bar{y}_R = y_R{}^* \quad (36)$$

only if

1. There was a direct measurement of hold-up changes, or

2. there was a known relation between the variance one needs to use for the artificial flowrates and the hold-up variance.

Since

$$\bar{V}_R = \int \bar{f}^v_R \, dt \tag{37}$$

$$V_R{}^* = \int f_R{}^* \, dt \tag{38}$$

Thus, when hold-ups are involved, one can certainly expect:

$$\bar{f}_R = f_R{}^* \tag{39}$$

$$\bar{V}_R = V_R{}^* \tag{40}$$

In reality, however only the tank hold-ups, rather than the corresponding artificial flowrates, are directly measured. In applying steady state data reconciliation to a system with tank hold-ups, one has to estimate the standard deviations for the artificial flowrates from their hold-ups. This estimation can vary since different methods could be applied. Instead, dynamic data reconciliation has the advantage that makes use of tank hold-ups directly. Therefore, one can expect a difference between the results by applying steady state data reconciliation and dynamic data reconciliation to a system.

5.3. Variance determination

We restrict the analysis to systems were one assumes that the variance–covariance matrix is diagonal.

5.3.1. True variance

Although more elaborate estimates of the variances R_F and R_V of a system that is not at steady state can be constructed, we propose to obtain them by using the following procedure:

(1) Perform a fitting of the dynamic data. For the case of polynomial fitting of flowrates one obtains (Bagajewicz & Jiang, 1997)

$$\tilde{\alpha}^R = [U_F^T U_F]^{-1} U_F^T f_R^+ \tag{41}$$

Similarly, for tank hold-ups one has

$$\tilde{\omega}^R = [U_V^T U_V]^{-1} U_V^T v_R^+ \tag{42}$$

where

$$U_F = \begin{bmatrix} I & 0 & \cdots & 0 \\ I & t_1 I & \cdots & t_1^s I \\ \vdots & \vdots & \ddots & \vdots \\ I & t_n I & \cdots & t_n^s I \end{bmatrix} \qquad U_V = \begin{bmatrix} I & 0 & \cdots & 0 \\ I & t_1 I & \cdots & t_1^{s+1} I \\ \vdots & \vdots & \ddots & \vdots \\ I & t_n I & \cdots & t_n^{s+1'} I \end{bmatrix}$$

(2) Construct the following sets of data

$$\{p_{Ri}\} = \left\{ f_{Ri}^+ - \sum_{k=0}^{s} \tilde{\alpha}_k^R t_i^k \right\} \tag{43}$$

$$\{q_{Ri}\} = \left\{ v_{Ri}^+ - \left(v_{R0} + \sum_{k=0}^{s} \tilde{\omega}_{k+1}^R t_i^{k+1} \right) \right\} \tag{44}$$

(3) Estimate the variance R_{Fi} and R_{Vi} as the standard deviation of $\{p_{Ri}\}$ and $\{q_{Ri}\}$, respectively.

5.3.2. Variance of averages

One can assess the variance R_s by directly calculating the standard deviation of the time varying sample. Since data reconciliation is performed using the averages, once again, the variance R_s is obtained by dividing the variance of the sample by the number of measurements used.

5.4. Simple system example

To perform an assessment on how large are the discrepancies that one obtains by using a steady state model, the system of Fig. 1 was investigated first. The measured values are shown in Figs. 2–4 and the standard deviation of the dynamic data is shown in Table 1. To enhance the effects a comparatively large variance on hold-up measurement was used.

Table 2 shows the calculated standard deviations for the three streams used in averagebased steady state data reconciliation. Final results are shown in Table 3.

6. Results for a more complex system

Consider the system introduced by (Bagajewicz & Jiang, 1997) and depicted in Fig. 5. The calculated variances are given in Table 4.

Table 5 compares the reconciled flowrates (real and artificial) obtained performing steady state data reconciliation with the average of the results obtained using the integral dynamic data reconciliation. Deviations from true values are shown.

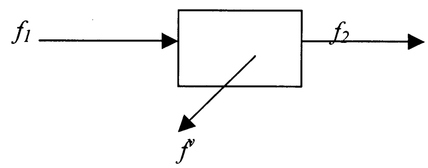

Fig. 1. A simple process.

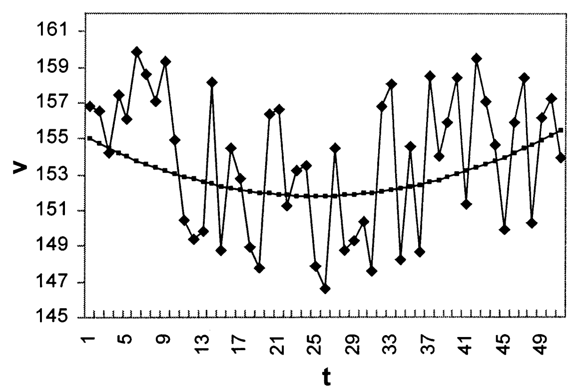

Fig. 2. Measurements and true values of tank hold-up.

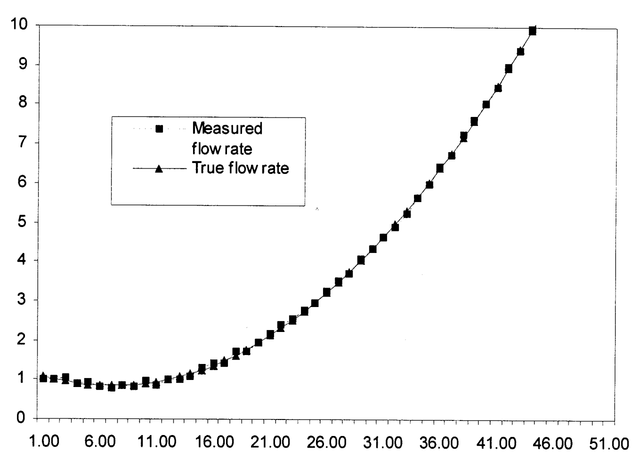

Fig. 3. Measurements and true values of f_1.

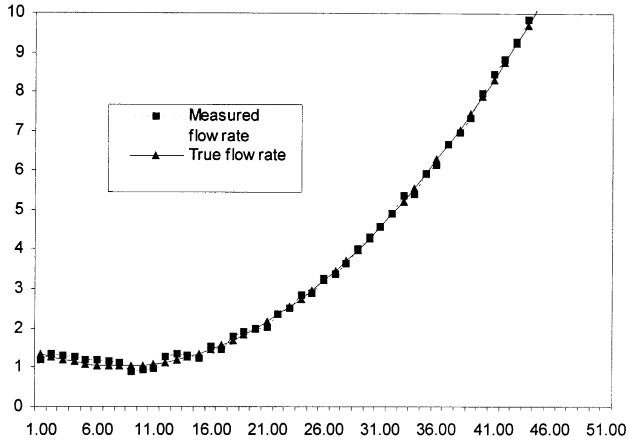

Fig. 4. Measurements and true values of f_2.

Table 1
Standard deviation used for dynamic state data reconciliation

f_1	f_2	V
0.04	0.08	3.2

Table 2
Standard deviation used for steady state data reconciliation

f_1	f_2	f^v
$0.04/\sqrt{51} = 0.0056$	$0.08/\sqrt{51} = 0.0112$	$3.2*\sqrt{2}/\sqrt{50} = 0.64$

Table 3
Comparison of steady state and dynamic state data reconciliation[a]

	f_1	f_2	f^v
Reconciled at steady state	4.6679	4.6595	0.0084
Reconciled dynamic state	4.6671	4.6629	0.0050
True value	4.6642	4.6542	0.01
Measured value	4.6679	4.6595	−0.0571

[a] Averages of indicated values.

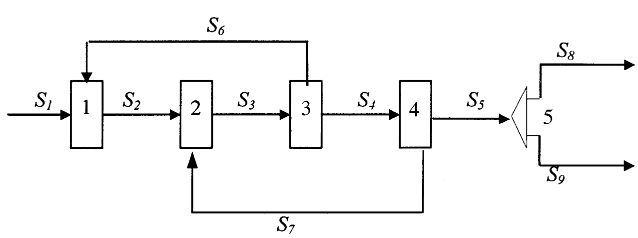

Fig. 5.

By simply comparing the values of the estimates of the instrument varialices (R_F and R_V) to the values of the variances used in steady state data reconciliation, one finds that the relative values from one to another vary. Thus, estimates of variances of flowrates that exhibit larger flow variations are also relatively larger. The importance of picking the proper variance when using data reconciliation is thus highlighted.

7. Conclusion

This paper has presented a comparison of steady, state data reconciliation and the integral dynamic data reconciliation. The theoretical and experimental results show that:

- When there is no tank and only flowrates are involved, the averages of reconciled values from dynamic data reconciliation are the same as the reconciled values from steady state reconciliation based on dynamic data averages, provided the right variance is used.
- When tank hold-ups are included in data reconciliation, the averages of reconciled flowrates/tank hold-ups from dynamic data reconciliation could theoretically be the same as the reconciled flowrates/hold-ups from steady state data reconciliation based on dynamic data averages, if a proper assessment of variance existed.
- Variances of measurements can be assessed using the dynamic data and can be used in steady state data reconciliation when averaged dynamic data is used.

The performance comparison of steady state and dynamic data reconciliation on gross error detection will be addressed in the future papers.

8. Notation

A balance matrix defined by Eq. (1)
A_R matrix used to denote linear combinations of redundant flows
B_R matrix used to denote the linear combination of tank hold-ups
C matrix defined by Eq. (2)
C_R matrix used to denote balances of redundant flows
C_α matrix used in model (13). Defined in the text
D matrix defined by Eq. (5)
D_m matrix used in model (34). Defined in the text
f vector of flowrates
G_R matrix used to denote balances of redundant flows
I identity matrix
J matrix used in model (13). Defined in the text
P matrix defined by Eq. (18)
S variance–covariance matrix
s polynomial order in Eqs. (10) and (11)
t time variable
U used in Eqs. (41) and (42)
R_F variance matrix of all flow measurements
R_V variance matrix of all volume measurements
R_m matrix used in model (13). Defined in the text
T_ω matrix used in model (13). Defined in the text
v vector of unit holdup
y vector of true values defined by Eq. (6)
$\tilde{y}$ vector of estimates of y given by Eq. (7)
Z auxiliary matrix
z_{ij} element of matrix Z

Greek letters
α^R vector used in model (13). Defined in the text
Γ_t averaging operator

Table 4
Variance of flowrates and holdups

	Variances of flowrates									Variances of Holdups			
	1	2	3	4	5	6	7	8	9	1	2	3	4
Original	0.04	0.08	0.2	0.16	0.12	0.2	0.16	0.04	0.08	3.2	15.0	6.0	3.0
Estimated	0.045	0.082	0.21	0.18	0.15	0.22	0.16	0.042	0.078	3.85	19.99	5.93	3.04

Table 5
Comparison of steady state and dynamic state data reconciliation

	Average flowerets (including real and artificial)[a]												
	1	2	3	4	5	6	7	8	9	10	11	12	13
Reconciled at steady state	1.469	2.467	3.696	4.469	3.369	5.136	4.373	1.431	1.938	4.139	3.144	−5.910	−3.273
Reconciled at dynamic state	1.470	2.464	3.701	4.451	3.373	5.150	4.388	1.432	1.941	4.164	3.158	−5.909	−3.316
True value	1.477	2.477	3.688	4.477	3.377	5.177	4.377	1.431	1.965	4.183	3.173	−5.973	−3.283

[a] Flowrates 10–13 correspond to the artificial flowrates corresponding to tanks 1–3.

ω^R vector used in model (13). Defined in the text

Subscripts and superscripts
F flowrate-related quantities
R redundant quantities
v hold-up-related quantities
$+$ measured quantities
$*$ estimates obtained using steady state models

Acknowledgements

Partial financial support from KBC Advanced Technologies for Q. Jiang is acknowledged.

Appendix A

In this appendix, Eq. (23) is derived. In the absence of accumulation terms, Eq. (12) becomes:

$$\left.\begin{array}{l}\min\{(T_\alpha\alpha^R - f_R^+)^T R_f^{-1}(T_\alpha\alpha^R - f_R^+)\} \\ \text{s.t.} \\ C_\alpha\alpha^R = 0\end{array}\right\} \quad (A1)$$

The substitution of Eq. (10) renders

$$\min\{(T_\alpha\alpha^R - f_R^+) + R_f^{-1}(T_\alpha\alpha^R - fr^{7i})\}$$

$$C_\alpha\alpha^{R=0} \quad (A2)$$

The solution to this problem is analytical and is given by:

$$\alpha^R = [I - QC_\alpha^T(C_\alpha QC_\alpha^T)^{-1}C_\alpha](-1/2Qw) \quad (A3)$$

However, Q, the inverse of $(T_\alpha^T, R_F^{-1}, T_\alpha)$ can be performed analytically. Then, according to Bagajewicz and Jiang (1997),

$$Q = (T_\alpha^T R_F^{-1} T_\alpha)^{-1} = ZR_F \quad (A4)$$

where

$$R_F = \begin{bmatrix} S_F & 0 & \cdots & 0 \\ 0 & S_F & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & S_F \end{bmatrix} \qquad Z = \begin{bmatrix} (n+1)I & I\sum_{i=1}^n j & \cdots & I\sum_{i=1}^n j^s \\ I\sum_{i=1}^n j & I\sum_{i=1}^n j^2 & \cdots & I\sum_{i=1}^n j^{s+1} \\ \vdots & \vdots & \ddots & \vdots \\ I\sum_{i=1}^n j^s & I\sum_{i=1}^n j^{s+1} & \cdots & I\sum_{i=1}^n j^{2s} \end{bmatrix}^{-1}$$

Therefore

$$\alpha^R[I - ZR_F C_\alpha^T(C_\alpha ZR_F C_\alpha^T)^{-1}C_\alpha]ZR_F(R_F^{-1}T_\alpha)^T f_R^+ \quad (A5)$$

Since $f_R = T_\alpha\alpha^R$, we have immediately Eq. (23).

Appendix B

In this appendix, we derive Eq. (24). Since

$$R_F C_\alpha^{\mathrm{T}} = \begin{bmatrix} S_F & & \\ & \ddots & \\ & & S_F \end{bmatrix} \begin{bmatrix} C_R^T & & \\ & \ddots & \\ & & C_R^T \end{bmatrix} = \begin{bmatrix} S_F C_R^T & & \\ & \ddots & \\ & & S_F C_R^T \end{bmatrix} \tag{B1}$$

To simplify notation, we assume that

$$Z = \begin{bmatrix} z_{00}I & z_{01}I & \cdots & z_{0s}I \\ z_{10}I & z_{11}I & \cdots & z_{1s}I \\ \vdots & \vdots & \ddots & \vdots \\ z_{s0}I & z_{s1}I & \cdots & z_{ss}I \end{bmatrix} \tag{B2}$$

where z_{ij} can be easily obtained by inverting the following matrix:

$$Z_R = \begin{bmatrix} (n+1) & \sum_{i=1}^{n} j & \cdots & \sum_{i=1}^{n} j^s \\ \sum_{i=1}^{n} j & \sum_{i=1}^{n} j^2 & \cdots & \sum_{i=1}^{n} j^{s+1} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{n} j^s & \sum_{i=1}^{n} j^{s+1} & \cdots & \sum_{i=1}^{n} j^{2s} \end{bmatrix} \tag{B3}$$

Therefore

$$Z R_F C_\alpha^T = \begin{bmatrix} z_{00}I & z_{01}I & \cdots & z_{0s}I \\ z_{10}I & z_{11}I & \cdots & z_{1s}I \\ \vdots & \vdots & \ddots & \vdots \\ z_{s0}I & z_{s0}I & \cdots & z_{ss}I \end{bmatrix} \begin{bmatrix} S_F C_R^T & & \\ & \ddots & \\ & & S_F C_R^T \end{bmatrix} \tag{B4}$$

which, after operating, results in:

$$Z R_F C_\alpha^T = \begin{bmatrix} z_{00}S_F C_R^T & z_{01}S_F C_R^T & \cdots & z_{0s}S_F C_R^T \\ z_{10}S_F C_R^T & z_{11}S_F C_R^T & \cdots & z_{1s}S_F C_R^T \\ \vdots & \vdots & \ddots & \vdots \\ z_{s0}S_F C_R^T & z_{s1}S_F C_R^T & \cdots & z_{ss}S_F C_R^T \end{bmatrix} \tag{B5}$$

Pre-multiplying by C_α, one obtains:

$$\left.\begin{aligned}
C_\alpha Z R_F C_\alpha^T &= \begin{bmatrix} C_R z_{00} S_F C_R^T & C_R z_{01} S_F C_R^T & \cdots & C_R z_{0s} S_F C_R^T \\ C_R z_{10} S_F C_R^T & C_R z_{11} S_F C_R^T & \cdots & C_R z_{1s} S_F C_R^T \\ \vdots & \vdots & \ddots & \vdots \\ C_R z_{s0} S_F C_R^T & C_R z_{s1} S_F C_R^T & \cdots & C_R z_{ss} S_F C_R^T \end{bmatrix} \\
&= \begin{bmatrix} C_R S_F C_R^T & & \\ & \ddots & \\ & & C_R S_F C_R^T \end{bmatrix} \begin{bmatrix} z_{00} I & z_{01} I & \cdots & z_{0s} I \\ z_{10} I & z_{11} I & \cdots & z_{1s} I \\ \vdots & \vdots & \ddots & \vdots \\ z_{s0} I & z_{s1} I & \cdots & z_{ss} I \end{bmatrix} \\
&= \begin{bmatrix} C_R S_F C_R^T & & \\ & \ddots & \\ & & C_R S_F C_R^T \end{bmatrix} Z
\end{aligned}\right\} \tag{B6}$$

Inverting:

$$(C_\alpha Z R_F C_\alpha^T)^{-1} = Z^{-1} \begin{bmatrix} C_R S_F C_R^T & & \\ & \ddots & \\ & & C_R S_F C_R^T \end{bmatrix}^{-1} = Z^{-1} \begin{bmatrix} (C_R S_F C_R^T)^{-1} & & \\ & \ddots & \\ & & (C_R S_F C_R^T)^{-1} \end{bmatrix} \tag{B7}$$

Similar manipulations lead to the following:

$$\left.\begin{aligned}
C_\alpha^T (C_\alpha Z R_F C_\alpha^T)^{-1} C_\alpha &= \begin{bmatrix} C_R^T & & \\ & \ddots & \\ & & C_R^T \end{bmatrix} Z^{-1} \begin{bmatrix} (C_R S_F C_R^T)^{-1} & & \\ & \ddots & \\ & & (C_R S_F C_R^T)^{-1} \end{bmatrix} \begin{bmatrix} C_R & & \\ & \ddots & \\ & & C_R \end{bmatrix} \\
&= \begin{bmatrix} C_R^T (C_R S_F C_R^T)^{-1} C_R & & \\ & \ddots & \\ & & C_R^T (C_R S_F C_R^T)^{-1} C_R \end{bmatrix} Z^{-1}
\end{aligned}\right\} \tag{B8}$$

Then,

$$Z R_F C_\alpha^T (C_\alpha Z R_F C_\alpha^T)^{-1} C_\alpha = \begin{bmatrix} S_F C_R^T (C_R S_F C_R^T)^{-1} C_R & & \\ & \ddots & \\ & & S_F C_R^T (C_R S_F C_R^T)^{-1} C_R \end{bmatrix} \tag{B9}$$

and finally,

$$I - Z R_F C_\alpha^T (C_\alpha Z R_F C_\alpha^T)^{-1} C_\alpha = \begin{bmatrix} I - S_F C_R^T (C_R S_F C_R^T)^{-1} C_R & & \\ & \ddots & \\ & & I - S_F C_R^T (C_R S_F C_R^T)^{-1} C_R \end{bmatrix}$$

$$= \begin{bmatrix} P & & \\ & \ddots & \\ & & p \end{bmatrix} \tag{B10}$$

Appendix C

Here we prove that Eq. (28) holds.
Since

$$ZZ^{-1}=I \tag{C1}$$

Then

$$(n+1)z_{00}+\sum_{j=1}^{n} jz_{10}+\cdots+\sum_{j=1}^{n} j^{s}z_{s0}=I \tag{C2}$$

This can be rewritten as

$$(n+1)z_{00}+\sum_{k=1}^{s}\sum_{j=1}^{n} j^{k}z_{k0}=(n+1)z_{00}+\sum_{j=1}^{n}\sum_{k=1}^{s} j^{k}z_{k0}$$

$$=(n+1)z_{00}+\left(\sum_{j=1}^{n}\sum_{k=0}^{s} j^{k}z_{k0}-nz_{00}\right)$$

$$=z_{00}+\sum_{j=1}^{n}\sum_{k=0}^{s} j^{k}z_{k0}=I \tag{C3}$$

Appendix D

Here we prove that Eqs. (29) and (30) hold.
Again from Eq. (C1) the following equations hold:

$$(n+1)z_{00}+\sum_{k=1}^{s}\sum_{j=1}^{n} j^{k}z_{0k}=I \tag{D1}$$

$$(n+1)z_{i0}+\sum_{k=1}^{s}\sum_{j=1}^{n} j^{k}z_{ik}=0 \quad i=1,\ldots,s \tag{D2}$$

Add up all equations in Eqs. (D1) and (D2) to obtain:

$$(n+1)\sum_{i=0}^{s} z_{i0}+\sum_{k=1}^{s}\sum_{j=1}^{n} j^{k}\sum_{i=0}^{s} z_{ik}=(n+1)\sum_{i=0}^{s} z_{i0}+\sum_{j=1}^{n}\sum_{k=1}^{s} j^{k}\sum_{i=0}^{s}\sum_{i=0}^{s} z_{ik}$$

$$=\sum_{i=0}^{s} z_{i0}+\sum_{j=1}^{n}\sum_{k=0}^{s} j^{k}\sum_{i=0}^{s} z_{ik}=I \tag{D3}$$

Multiply by m^{i} both sides of Eq. (D2)

$$(n+1)m^{i}z_{i0}+\sum_{k=1}^{s}\sum_{j=1}^{n} j^{k}m^{i}z_{ik}=0 \quad i=1,\ldots,s; \qquad m=2,\ldots,n \tag{D4}$$

Add up all equations from Eqs. (D2), (D3) and (D4) to obtain:

$$\sum_{i=0}^{s} m^{i}z_{i0}+\sum_{j=0}^{n}\sum_{k=0}^{s} j^{k}\sum_{i=0}^{s} m^{i}z_{ik}=I \tag{D5}$$

Q.E.D.

References

Almasy, G. A., & Mah, R. S. H. (1984). Estimation of measurement error variances from process data. *Industrial & Engineering Chemistry, Process Design & Development*, *23*, 779.

Albuquerque, J. S., & Biegler, L. T. (1996). Data reconciliation and gross-error detection for dynamic systems. *American Institute of Chemical Engineering Journal*, *42*, 2841.

Bagajewicz, M., & Jiang, Q. (1997). Integral approach to plant linear dynamic reconciliation. *American Institute of Chemical Engineering Journal*, *43*, 2546.

Bagajewicz, M., & Jiang, Q. (1998). Gross error modeling and detection in plant linear dynamic reconciliation. *Computers & Chemical Engineering* (submitted).

Chen, J., Bandoni, A., & Romagnoli, J. A. (1997). Robust estimation of measurement error variance/covariance from process sampling data. *Computers & Chemical Engineering*, *21*(6), 593–600.

Crowe, C. M. (1996). Formulation of linear data reconciliation using information theory. *Computers & Chemical Engineering*, *51*(12), 3359–3366.

Crowe, C. M., García Campos, Y. A., & Hrymak, A. (1983). Reconciliation of process flow rates by matrix projection. Part 1: linear case. *American Institute of Chemical Engineering Journal*, *29*, 881–888.

Darouach, M., & Zasadzinski, M. (1991). Data reconciliation in generalized linear dynamic systems. *American Institute of Chemical Engineering Journal*, *37*, 193.

Johnston, L. P. M., & Kramer, M. A. (1995). Maximum likelihood data rectification. Steady, state systems. *American Institute of Chemical Engineering Journal*, *41*, 11.

Kao, C. S., Tamhane, A. C., & Mah, R. S. H. (1992). Gross error detection in serially correlated process data. 2. Dynamic systems. *Industrial Engineering & Chemical Research*, *31*, 254.

Karjala, T. W., & Himmelblau, D. M. (1996). Dynamic rectification of data via recurrent neural nets and the extended Kalman filter. *American Institute of Chemical Engineering Journal*, *42*, 2225.

Keller, J. Y., Zasadzinski, M., & Darouach, M. (1992). Analytical estimator of measurement error variances in data reconciliation. *Computers & Chemical Engineering*, *16*, 185.

Kuehn, D. R., & Davidson, H. (1961). Computer control. II. Mathematics of control. *Chemical Engineering Programming*, *57*, 44.

Liebman, M. L, Edgar, T. F., & Lasdon, L. S. (1992). Efficient data reconciliation and estimation for dynamic process using nonlinear programming techniques. *CES*, *16*, 963.

Madron, F. (1992). *Process plant performance*. *Measurement and data processing for optimization and retrofits*. Chichester, UK: Ellis Horwood.

Mah, R. S. H. (1990). *Chemical process structures and information flows*. London: Butterworths.

Narasimhan, S., & Mah, R. S. H. (1988). Generalized likelihood ratios for gross error identification in dynamic processes. *American Institute of Chemical Engineering Journal*, *34*, 1321.

Ramamurthi, Y., Sistu, P. B., & Bequette, B. W. (1993). Control-relevant dynamic data reconciliation and parameter estimation. *Computers & Chemical Engineering*, *17*, 41.

Rollins, D. K., & Devanathan, S. (1993). Unbiased estimation in dynamic data reconciliation. *American Institute of Chemical Engineering Journal*, *39*, 1330.

Stanley, G. M., & Mah, R. S. H. (1977). Estimation of flows and temperatures in process networks. *American Institute of Chemical Engineering Journal*, *23*, 642.