Unbiased Estimation in Dynamic Data Reconciliation

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

A computationally fast technique accurately estimates process variables when conditions are dynamic due to changes in steady states. The process variable estimators are unbiased and have known distributions. Thus, confidence intervals for true values of process variables are provided. The formulation of this technique was motivated by a recursive, dynamic data reconciliation technique that obtains very accurate estimators. These two techniques are compared in terms of computational speed and accuracy of estimators. The proposed technique is computationally faster, but not as accurate when variances of process measurements are large. However, the accuracy of the proposed estimators is shown to approach that of the recursive technique by iteratively recalculating estimates and when measurement variances decrease.

Introduction

Darouach and Zasadzinski (1991) presented a dynamic online estimation algorithm for reconciling process variables. Their algorithm involves a recursive solution technique in weighted least squares. Darouach and Zasadzinski (1991) showed that their technique can improve the accuracy of process variables considerably when conditions are changing from one steady state to another steady state. Although not shown in Darouach and Zasadzinski (1991), their estimators have the attractive property of being unbiased.

In this article, we present a new data reconciliation technique for dynamic processes that is computationally simpler than the recursive technique of Darouach and Zasadzinski (1991). The essence of the proposed technique for the accumulation terms is a backward difference approximation. A constrained least-squares approach is used to obtain an optimal solution that guarantees improved values for process variables.

In our formulation of the optimization problem, two estimates for the accumulation variables are obtained for all time instants except the first and last. Hence, estimation accuracy is further improved by averaging when two estimates exist for one variable at one time instant. Improvement in addition to averaging is achieved by reestimating these process variables using the averaged estimates as if they are the original measurements. This is possible because the averaged estimates do not satisfy physical (material balance) constraints. Note, however, that nonaveraged estimates at each iteration do satisfy physical constraints, and, thus, can be used if this property is needed.

A potential drawback to our iterative approach is an increase in computational time with each additional iteration. Hence, in this article we present a simulation study to compare the computational speed and estimation accuracy of the Darouach and Zasadzinski (1991) technique with the proposed technique at various iteration steps. The process network that we used for this study was taken from Darouach and Zasadzinski (1991). Before presenting the theory of the proposed technique and the results of the simulation study, a summary of the major findings are given below.

Summary

We did two studies to compare the two approaches. The first one involved large measurement variances, and we varied the number of iterations for the Rollins and Devanathan technique. In this study, the Rollins and Devanathan estimator variances for the first iteration were about three times larger than the Darouach and Zasadzinski values. However, the Rollins and Devanathan algorithm was about nine times faster. The computation times for both techniques were about equal for the fourth iteration of the Rollins and Devanathan technique. Here the Rollins and Devanathan estimator variances were about 1.5 times larger. In the second study, we varied

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

measurement variances and determined results for the first iteration of the Rollins and Devanathan estimators only. In this study these Rollins and Devanathan estimator variances approached the Darouach and Zasadzinski estimator variances as the measurement variances decreased. Hence, our overall conclusion is that, if measurement variances are not too large, the Rollins and Devanathan estimators can have competitive accuracy (with the Darouach and Zasadzinski estimators) and greater computational speed.

Models

The process models are mathematical expressions for the material balances and the measurements. The physical model, presented first, represents material balance constraints. This is followed by the measurement model, a statistical expression for the measurements.

At the *i*th time instant, a total mass balance on each of the *n* nodes gives

$$\begin{aligned}
-\boldsymbol{W}_i^{*\,n\times 1}+\boldsymbol{W}_{i-1}^{*}+\boldsymbol{M}^{n\times v}\boldsymbol{Q}_i^{*\,v\times 1}
&= -[\boldsymbol{I}^{n\times n}\,|\,-\boldsymbol{M}^{n\times v}]\begin{bmatrix}\boldsymbol{W}_i^{*}\\ \boldsymbol{Q}_i^{*}\end{bmatrix}+[\boldsymbol{I}^{n\times n}\,|\,\boldsymbol{0}^{n\times v}]\begin{bmatrix}\boldsymbol{W}_{i-1}^{*}\\ \boldsymbol{Q}_{i-1}^{*}\end{bmatrix}\\
&= -\boldsymbol{E}\boldsymbol{X}_i^{*}+\boldsymbol{B}\boldsymbol{X}_{i-1}^{*}=[-\boldsymbol{E}\,|\,\boldsymbol{B}]\begin{bmatrix}\boldsymbol{X}_i^{*}\\ \boldsymbol{X}_{i-1}^{*}\end{bmatrix}\\
&= \boldsymbol{\Phi}\boldsymbol{X}_i^{*}=\boldsymbol{0}, \qquad (1)
\end{aligned}$$

where $i=2, \ldots, k$ and

$$\boldsymbol{E}^{n\times(n+v)}=[\boldsymbol{I}\,|\,-\boldsymbol{M}], \qquad (2)$$

$$\boldsymbol{B}^{n\times(n+v)}=[\boldsymbol{I}\,|\,\boldsymbol{0}], \qquad (3)$$

$$\boldsymbol{X}_i^{*}=\begin{bmatrix}\boldsymbol{W}_i^{*}\\ \boldsymbol{Q}_i^{*}\end{bmatrix}^{(n+v)\times 1}, \qquad (4)$$

$$\boldsymbol{\Phi}^{n\times 2(n+v)}=[-\boldsymbol{E}\,|\,\boldsymbol{B}], \qquad (5)$$

$$\boldsymbol{X}_i^{*2(n+v)\times 1}=\begin{bmatrix}\boldsymbol{X}_i^{*}\\ \boldsymbol{X}_{i-1}^{*}\end{bmatrix}, \qquad (6)$$

$\boldsymbol{W}_i^{*}$ is a $n\times 1$ vector of "true" and unknown total mass in the n nodes at time instant i, $\boldsymbol{Q}_i^{*}$ is a $v\times 1$ vector of "true" and unknown total mass flow rates for the v streams at time instant i, and k = current time instant.

The measurement model that applies to Eq. 1 is

$$\boldsymbol{X}_i^{2(n+v)\times 1}=\boldsymbol{X}_i^{*}+\boldsymbol{E}_i, \qquad (7)$$

where $i=2, \ldots, k$, and

$$\boldsymbol{E}_i=\begin{bmatrix}\boldsymbol{w}_i\\ \boldsymbol{q}_i\\ \boldsymbol{w}_{i-1}\\ \boldsymbol{q}_{i-1}\end{bmatrix}=\begin{bmatrix}\boldsymbol{\epsilon}_i\\ \boldsymbol{\epsilon}_{i-1}\end{bmatrix}, \qquad (8)$$

$$\mathrm{Var}\ (\boldsymbol{E}_i)=\boldsymbol{\Sigma}^{2(n+v)\times 2(n+v)} \qquad (9)$$

$$\boldsymbol{\epsilon}_j\sim N_{(n+v)}(\boldsymbol{0},\boldsymbol{V}), \qquad (10)$$

$$\boldsymbol{V}^{(n+v)\times(n+v)}=\begin{bmatrix}\boldsymbol{V}_W & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{V}_Q\end{bmatrix}, \qquad (11)$$

$$\mathrm{Var}\ (\boldsymbol{w}_j)=\boldsymbol{V}_W, \qquad (12)$$

$$\mathrm{Var}\ (\boldsymbol{q}_j)=\boldsymbol{V}_Q, \qquad (13)$$

$j=1, \ldots, k$. Note that $\boldsymbol{V}_W$ and $\boldsymbol{V}_Q$ are assumed to be known although this is not a restrictive assumption. In addition, the sample size is assumed to be one for convenience.

The Estimator

The estimator that we are proposing for $\boldsymbol{X}_i^{*}$ is a maximum likelihood estimator and is found by minimizing

$$(\boldsymbol{X}_i-\boldsymbol{X}_i^{*})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{X}_i-\boldsymbol{X}_i^{*}) \qquad (14)$$

with respect to $\boldsymbol{X}_i$ and subject to Eq. 1. The result of this optimization is the following estimator for $\boldsymbol{X}_i^{*}$ (see Mardia et al., 1979):

$$\begin{aligned}
\hat{\boldsymbol{X}}_i&=\boldsymbol{X}_i-\boldsymbol{\Sigma}\boldsymbol{\Phi}^T(\boldsymbol{\Phi}\boldsymbol{\Sigma}\boldsymbol{\Phi}^T)^{-1}\boldsymbol{\Phi}\boldsymbol{X}_i\\
&=\boldsymbol{D}\boldsymbol{X}_i, \qquad (15)\\
i&=2, \ldots, k
\end{aligned}$$

with

$$\mathrm{Var}\ (\hat{\boldsymbol{X}}_i)=\boldsymbol{\Sigma}-\boldsymbol{\Sigma}\boldsymbol{\Phi}^T(\boldsymbol{\Phi}\boldsymbol{\Sigma}\boldsymbol{\Phi}^T)^{-1}\boldsymbol{\Phi}\boldsymbol{\Sigma} \qquad (16)$$

where

$$\hat{\boldsymbol{X}}_i^{2(n+v)\times 1}=\begin{bmatrix}\hat{\boldsymbol{X}}_i\\ \hat{\boldsymbol{X}}_{i-1}\end{bmatrix}. \qquad (17)$$

Note that Eq. 15 satisfies the constraint equation, that is, $\boldsymbol{\Phi}\underline{\hat{\boldsymbol{X}}}_i=\boldsymbol{0}$.

Equation 17 shows that at each time instant i ($i=2, \ldots, k$), an estimate for the value of each variable at time instant i and $i-1$ is determined. Thus, for $i=2, \ldots, k$, one estimate is determined for each variable at time instant 1, two for each variable at time instants $2, \ldots, k-1$, and one for each variable at time instant k. For time instants $2, \ldots, k-1$, we recommend that the two values for each variable be averaged since the averages will have smaller variances.

The proposed estimators for each time instant will now be derived from Eq. 15 using averages where possible. First, let

$$\boldsymbol{D}_1^{(n+v)\times 2(n+v)}=[\boldsymbol{I}^{(n+v)\times(n+v)}\,|\,\boldsymbol{0}^{(n+v)\times(n+v)}] \qquad (18)$$

and

$$\boldsymbol{D}_2^{(n+v)\times 2(n+v)}=[\boldsymbol{0}^{(n+v)\times(n+v)}\,|\,\mathbf{I}^{(n+v)\times(n+v)}]. \qquad (19)$$

Now by application of Eqs. 15, 18 and 19, we propose the following estimators:

$$\hat{X}_1' = D_2\hat{X}_2 = D_2DX_2 = D_{2D}X_2, \tag{20}$$

$$\begin{aligned}\hat{X}_i' &= \frac{1}{2}D_1\hat{X}_i + \frac{1}{2}D_2\hat{X}_{i+1}\\ &= \frac{1}{2}D_1DX_i + \frac{1}{2}D_2DX_{i+1}\\ &= \frac{1}{2}D_{1D}X_i + \frac{1}{2}D_{2D}X_{i+1},\\ i &= 2, \ldots, k-1\end{aligned} \tag{21}$$

$$\hat{X}_k' = D_1\hat{X}_k = D_1DX_k = D_{1D}X_k. \tag{22}$$

Therefore, the variance-covariance matrices for $\hat{X}_1'$, $\hat{X}_i'$ ($i=2, \ldots, k-1$) and $\hat{X}_k'$ are

$$\Sigma_1 = D_{2D}\Sigma D_{2D}^T, \tag{23}$$

$$\Sigma_i = \frac{1}{4}D_{1D}\Sigma D_{1D}^T + \frac{1}{4}D_{2D}\Sigma D_{2D}^T + \frac{1}{2}D_{1D}\begin{bmatrix}0 & V\\ 0 & 0\end{bmatrix}D_{2D}^T, \tag{24}$$

$$i = 2, \ldots, k-1$$

$$\Sigma_k = D_{1D}\Sigma D_{1D}^T, \tag{25}$$

respectively.

An estimator $\hat{\theta}$ is an unbiased estimator for θ, if $E[\hat{\theta}] = \theta$. By taking the expected value of Eq. 15 we show $\hat{X}_i$ to be an unbiased estimator of X_i^* as follows:

$$\begin{aligned}E[\hat{X}_i] &= E[X_i] - \Sigma\Phi^T(\Phi\Sigma\Phi^T)\Phi E[X_i]\\ &= X_i^*\end{aligned} \tag{26}$$

since $\Phi E[X_i] = \Phi X_i^* = 0$ by Eqs. 1 and 7.

Thus, Eqs. 20-22 are also unbiased estimators. Moreover, they are distributed as $(n+v)$ variate normal random variables with unbiased means, and variances given by Eqs. 23-25. Therefore, simultaneous $100(1-\alpha)\%$ confidence intervals at any time instant can be determined. For example, simultaneous $100(1-\alpha)\%$ confidence intervals for the process variables at time instant 1 are

$$\begin{gathered}\hat{X}_{1j}' \bullet z_{\alpha/(n+v)}\sqrt{e_j^T\Sigma_1 e_j}\\ j = 1, \ldots, n+v\end{gathered} \tag{27}$$

where $\hat{X}_{1j}'$ is the jth element of $\hat{X}_1'$, e_j is a $(n+v)\times 1$ vector with a one for the jth element and zeros elsewhere, and $z_{\alpha/(n+v)}$ is the upper $100(\alpha/\{n+v\})$th percentile of the standard normal distribution. Hence, since confidence intervals can be determined for these estimators, one can know, with a level of confidence, how "good" they really are without resorting to a simulation study.

Comparative Study and Background

The proposed approach is evaluated in this section by examining its estimator accuracy and computational speed against the technique of Darouach and Zasadzinski (1991). The accuracy of an estimator is measured by its mean-squared error. Note that, since the estimators in this study are unbiased, their mean-squared errors are equivalent to their variances. Hence, we present estimator variance as a measure of accuracy in this study. Exact variances for both techniques were determined by mathematical formulas (theoretically) and checked by Monte Carlo simulations.

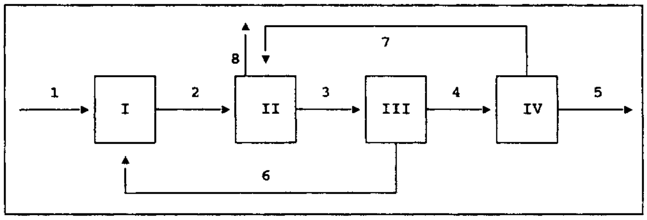

Figure 1. Process network.

Figure 1 shows the process network used in this evaluation. This network was taken from Darouach and Zasadzinski (1991). Measurement data were generated from normally distributed random errors with zero means. The measurement variances for the Q's were also taken from Darouach and Zasadzinski (1991) and are presented below:

$$V_Q = \text{Diag}\ (1, 1.96, 1.21, 0.49, 0.36, 0.16, 0.09, 0.25) \tag{28}$$

For the W's, the measurement variances changed throughout this study and are presented later. The true values for the measured variables do not affect the Darouach and Zasadzinski and Rollins and Devanathan mean-squared errors because their estimators are unbiased. That is, any set of true values satisfying the constraint equations will give the same the mean-squared errors. Thus, for this reason, and space limitation, we chose not to present the true values of the measured variables used in this investigation.

A fair comparison of computational speed requires as much similarity as possible. Thus, we wrote algorithms for both techniques in the same language, made everything similar except for the equations to estimate the process variables and ran them on the same machine (a digital DECstation 5,000 workstation). In addition, we determined reported computational times by averaging 20 runs. We wrote the algorithm for the Rollins and Devanathan technique using the equations in the previous section. Similarly, we wrote the algorithm for the Darouach and Zasadzinski technique using their formulas to determine $\hat{X}_{i/i}$ (Eq. 22b, in Darouach and Zasadzinski, 1991), where $\hat{X}_{i/i}$ is the vector of variable estimators at time instant i based on the measurements up to time instant i. We verified correct performance of these algorithms by comparing variances the algorithms determined in Monte Carlo simulations with the variances calculated from the mathematical formulas. Each comparison was based on 1,000 simulated cases of data. For all comparisons at all time instants, agreement was excellent.

Reported estimator variances are overall variances. The overall variance for each process variable estimator was determined by averaging its variances for all time instants on a dynamic process run. Since the number of time instants for each dynamic process run in this study was 50, the overall

Table 1. Comparison of Estimator Variances When Measurement Variances Are Large

Method	W_1	W_2	W_3	W_4	Time 10^{-2} s	Time Ratio RD/DZ
Meas.	225	144	324	484	N/A	N/A
DZ	32	25	39	49	N/A	N/A
DZS	33	25	38	50	9.6	N/A
RDC^1	109	71	165	243	1.0	9.6
RDA^1	87	55	122	177	1.1	8.7
RD	84	54	121	181	N/A	N/A
RDC^2	92	72	129	193	3.3	2.9
RDA^2	64	42	92	135	3.6	2.7
RDC^3	81	68	114	172	5.5	1.7
RDA^3	54	35	77	114	6.1	1.6
RDC^4	73	68	103	154	8.0	1.2
RDA^4	49	32	68	100	8.8	1.1

The superscripts on RD and RDA represent the iterative trial number. The Meas. values are the variances of the measured variables. The DZ and DZS values are the Darouach and Zasadzinski (1991) estimator variances determined from theoretical equations and simulations, respectively. The RDC are variances for the nonaveraged estimators given by Eq. 15. The RDA are variances for the averaged estimators determined from Eqs. 20-22. The RD variances were determined theoretically from Eqs. 23 to 25. The RDC, RDA, and DZS values were determined from 1,000 simulated cases of data. N/A = not applicable.

Table 2. Comparison of Estimator Variances for Various Measurement Variances

Type	W_1	W_2	W_3	W_4
Meas.	225	144	324	484
DZ	33	25	38	49
RDA^1	84	54	121	181
Meas.	100	100	100	100
DZ	19	20	16	14
RDA^1	38	38	38	38
Meas.	50	50	50	50
DZ	12	12	10	8
RDA^1	20	20	20	19
Meas.	25	25	25	25
DZ	8	8	6	5
RDA^1	10	10	10	10
Meas.	10	10	10	10
DZ	4	4	3	3
RDA^1	4	4	4	4

The Meas. values are the variances of the measured variables. The DZ values are Darouach and Zasadzinski (1991) estimator variances determined theoretically. The RDA^1 values are variances determined from Eqs. 23 to 25 for one iteration.

variance for some estimator $\hat{\theta}_j$ of θ_j is represented by the equation below:

$$\sigma^2_{OA} = \sum_{j=1}^{50} \frac{E[(\hat{\theta}_j - \theta_j)^2]}{50} \tag{29}$$

where E is used for expected value. We also give the formula we used to obtain Darouach and Zasadzinski overall variances from Monte Carlo simulations:

$$\sigma^2_{DZ} = \sum_{l=1}^{1,000} \frac{\displaystyle\sum_{j=1}^{50} \frac{(\hat{X}_{l,i,j/j} - X^*_{i,j})^2}{50}}{1,000} \tag{30}$$

where $\hat{X}_{l,i,j/j}$ is the value of $\hat{X}_{i,j/j}$ for the lth simulation trial and $X^*_{i,j}$ is the true value of the ith variable at the jth time instant. We used a similar equation to calculate Rollins and Devanathan overall variances.

This study consisted of two parts. For the first part we used the large measurement variances given in Darouach and Zasadzinski (1991) and determined Rollins and Devanathan estimates from Eq. 15 and Eqs. 20-22 for four iterations. Recall that Eq. 15 gives estimates that satisfy the physical constraints and Eqs. 20-22 are averaged estimates (which are more accurate) that do not satisfy the physical constraints. As mentioned earlier, the latter estimates make the iterative process possible. The purposes of this part were to verify proper algorithmic performances and to compare estimator accuracy and computational speed at different iteration steps of the Rollins and Devanathan technique. Table 1 shows the results of this part of the study. The second part of this investigation consisted of evaluating the size of measurement variance on estimation accuracy. Variances for the Rollins and Devanathan technique were determined for the first iteration only. We varied measurement variances in Table 2 by starting with the large values in Table 1 and reducing them four times by approximately 50% each time. Note that in both tables only results for W's are given. Results for the Q's are not given, because their variances were too small to provide any interesting analyses or conclusions. However, for Q's with larger variances the conclusions of this study should be applicable to the Q's as well.

Comparative Study and Result

In Table 1, DZ and DZS denote the variances for the Darouach and Zasadzinski estimators determined theoretically and by Monte Carlo simulations, respectively. A comparison of these two methods they agree for all four W's. Hence, it appears that the Darouach and Zasadzinski algorithm gives the correct estimates. In Table 1 three Rollins and Devanathan estimates are represented. RD and RDA denote variances for averaged estimators determined theoretically (calculated by Eqs. 23-25) and by Monte Carlo simulations, respectively. The numerical superscript is used to identify the iteration number. Again the agreement between RD and RDA values is excellent, confirming correct operation of the Rollins and Devanathan algorithm. The third Rollins and Devanathan method, RDC, represents the nonaveraged Rollins and Devanathan estimators (calculated by Eq. 15) which satisfy the physical constraints. As expected, at every iteration step, the variances of this estimator are larger than the variances of the RDA estimators.

Table 1 also shows that both the Darouach and Zasadzinski and Rollins and Devanathan estimators improved the accuracy of the process variables over the measurements. The Darouach and Zasadzinski estimator makes the greatest improvement but the difference between the Darouach and Zasadzinski and Rollins and Devanathan variances decreases as the number of iterations for the Rollins and Devanathan estimators increases. In addition, it appears that the size of the improvement for both techniques increases as the measurement variances increase. As shown by Table 1, in this study, the Rollins and Devanathan algorithm was about 9 times faster for RDC^1

(where the Darouach and Zasadzinski estimators were 3.5 to 5 times more accurate) and only slightly faster for RDA[4] (where the Darouach and Zasadzinski estimators were 1.5 to 2 times more accurate).

Table 1 seems to indicate that relative improvement between the estimators and the measurements and between each estimator is a function of the measurement variances. Thus, in Table 2 the measurement variances are varied in order to study this effect on accuracy. As shown, when the measurement variances are the same for all four W's, the Darouach and Zasadzinski and RDA[1] variances are about the same. This table also shows that as the measurement variances decrease the RDA[1] variances approach the Darouach and Zasadzinski variances. Thus, depending on the size of the measurement variances, it may be possible to achieve accuracy close to the Darouach and Zasadzinski estimators but at a much greater computation rate.

Summary

In this work, a computationally fast dynamic data reconciliation technique was presented that can significantly improve the accuracy of measured process variables. In comparing this Rollins and Devanathan technique with the Darouach and Zasadzinski (1991) recursive technique, we found the Rollins and Devanathan technique to be much faster but less accurate when measurement variances are large. However, the accuracy of Rollins and Devanathan procedure quickly approached the accuracy of the Darouach and Zasadzinski technique as the number of iterations of the Rollins and Devanathan procedure increased (which also significantly increases computational time) and as the measurement variances decreased. The Darouach and Zasadzinski approach appears to take a giant step in improving accuracy but it also takes a big step in increasing computational time. In contrast, the Rollins and Devanathan technique takes smaller steps in improving accuracy and smaller steps in increasing computational time. Thus, in selecting between these two techniques, one should consider the size of the measurement variances and the importance of accuracy and speed for the particular application. Moreover, one must remember that as the size of process networks increase, small differences in computational speed can be critical.

In addition, although not shown in this article, the Rollins and Devanathan method is applicable to situations when the measurement variances are unknown. Finally, we are currently considering ways to extend this approach to gross error detection.

Notation

D = matrix given by Eq. 15
D_1 = matrix given by Eq. 18
D_2 = matrix given by Eq. 19
D_{1D} = matrix given by Eq. 22
D_{2D} = matrix given by Eq. 20
DZ = Darouach and Zasadzinski (1991)
e_j = $(n+v)\times 1$ vector with a one for the jth element and zeros elsewhere
E, B = constraint matrices in generalized dynamic system
$E[\hat{\theta}]$ = expected value of $\hat{\theta}$
I = $n\times n$ identity matrix
k = current time instant
M = incidence matrix
m_{ij} = element (i, j) of incidence matrix M
n = number of nodes
Q_i = vector of flow measurements at time instant i
Q_i^* = vector of true values at time instant i
V = covariance matrices of measurement errors
RD = proposed approach
V_Q = covariance matrix of measurement errors on flows Q
V_W = covariance matrix of measurement errors on volumes W
v = number of flows
v_i = vector of measurement errors on Q_i
w_i = vector of measurement errors on W_i
W_i^* = vector of true total mass values at time instant i
W_j = vector of measured total mass values at time instant j
$\hat{W}$ = vector of estimates for the W's
W_1 = measured value of total mass at node 1
$\hat{X}_i$ = $(n+v)\times 1$ vector of estimates for true values at time instant i, given by Eq. 17
$\hat{X}_i'$ = $(n+v)\times 1$ vector of the proposed estimates for true values at time instant i, given by Eqs. 20–22
$\hat{X}_{1j}'$ = the jth element of $\hat{X}_1'$
X_i^* = $(n+v)\times 1$ vector of true values of unknown variables at time instant i
$\hat{X}_{j/j}$ = Darouach and Zasadzinski estimator vector of variable estimators at time instant j based on the measurements up to time instant j
$X_{i,j}^*$ = true value of the ith variable at the jth time instant
X_i = $2(n+v)\times 1$ vector of measured values at time instant i, Eq. 7
X_i^* = $2(n+v)\times 1$ vector of true values given by Eq. 6
$\underline{X}_i$ = $2(n+v)\times 1$ vector of estimates for true values estimated at time instant i, given by Eq. 15
$z_{\alpha/2p}$ = $100(\alpha/2p)$th percentile of the normal distribution

Greek letters

ϵ_i = vector of measurement errors at time instant i, Eq. 10
E_i = vector of measurement errors given by Eq. 8
θ_j = true value at time instant j
$\hat{\theta}_j$ = estimated value at time instant j
θ = true value of a variable
$\hat{\theta}$ = estimator for the true value
σ^2_{OA} = averaged variance of an estimator over the 50 time instants given by Eq. 29
σ^2_{DZ} = averaged variance for a Darouach and Zasadzinski estimator over the 50 time instants determined by simulation; see Eq. 30
Σ_i = variance-covariance matrix of $\hat{X}_i'$, given by Eqs. 23–25
Σ = variance matrix, Eq. 9
Φ = constraint matrix, Eq. 5

Others

$\sim$ = distributed
$\approx$ = approximately equal to
$\mathbf{0}$ = $n\times v$ null matrix

Superscript

T = transpose

Literature Cited

Darouach, M., and M. Zasadzinski, "Data Reconciliation in Generalized Linear Dynamic Systems," *AIChE J.*, **37**(2), 193 (1991).

Mardia, K. V., J. T. Kent, and J. M. Biddy, *Multivariate Analysis*, Academic Press, New York (1979).

Rollins, D. K., and J. F. Davis, "Unbiased Estimation of Gross Errors in Process Measurements," *AIChE J.*, **38**, 563 (1992).

Manuscript received Apr. 17, 1992, and revision received Dec. 4, 1992.